

Etienne Lwamba
Shannon Shisler
Will Ridlehoover
Meital Kupfer
Nkululeko Tshabalala
Promise Nduku
Laurenz Langer
Sean Grant
Ada Sonnenfeld
et al.

Strengthening women's empowerment and gender equality in fragile contexts towards peaceful and inclusive societies

October 2021

Systematic
Review 47

Social protection

About 3ie

The International Initiative for Impact Evaluation (3ie) promotes evidence-informed equitable, inclusive and sustainable development. We support the generation and effective use of high-quality evidence to inform decision-making and improve the lives of people living in poverty in low- and middle-income countries. We provide guidance and support to produce, synthesise and quality assure evidence of what works, for whom, how, why and at what cost.

3ie systematic reviews

3ie systematic reviews appraise and synthesise the available high-quality evidence on the effectiveness of social and economic development interventions in low- and middle-income countries. These reviews follow scientifically recognised review methods, and are peer-reviewed and quality assured according to internationally accepted standards. 3ie is providing leadership in demonstrating rigorous and innovative review methodologies, such as using theory-based approaches suited to inform policy and programming in the dynamic contexts and challenges of low- and middle-income countries.

About this review

Strengthening women's empowerment and gender equality in fragile contexts towards peaceful and inclusive societies was made possible thanks to generous funding from the Transitional Development Assistance Programme of the German Federal Ministry for Economic Cooperation and Development (BMZ) and undertaken in cooperation with the German Corporation for International Cooperation GmbH (GIZ). 3ie is publishing this technical report as received from the authors; it has been formatted to 3ie style.

All content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors are also the sole responsibility of the authors. Comments or queries should be directed to the corresponding author, Etienne Lwamba, at elwamba@3ieimpact.org.

Suggested citation: Lwamba, E, Shisler, S, Ridlehoover, W, Kupfer, M, Tshabalala, N, Nduku, P, Langer, L, Grant, S, Sonnenfeld, A, Anda, D, Eyers, J and Snilstveit, B, 2021. *Strengthening women's empowerment and gender equality in fragile contexts towards peaceful and inclusive societies*. 3ie Systematic Review 47. New Delhi: International Initiative for Impact Evaluation (3ie). Available at: <https://doi.org/10.23846/SR00047>

3ie systematic review executive editor: Birte Snilstveit

Production manager: Tanvi Lal

Assistant production manager: Akarsh Gupta

Strengthening women's empowerment and gender equality in fragile contexts towards peaceful and inclusive societies

Etienne Lwamba
International Initiative for Impact Evaluation (3ie)

Shannon Shisler
3ie

Will Ridlehoover
3ie

Meital Kupfer
3ie

Nkululeko Tshabalala
Africa Centre for Evidence, University of Johannesburg (ACE)

Promise Nduku
ACE

Laurenz Langer
ACE

Sean Grant
3ie

Ada Sonnenfeld
3ie

Daniela Anda
3ie

John Evers
3ie

Birte Snilstveit
3ie

3ie Systematic Review 47

October 2021



Acknowledgements

We are grateful to 3ie and ACE colleagues who provided inputs at various stages of the review and supported on the administrative, communication and finance management. We are particularly grateful to Judith Müller-Gerold for her time, flexibility and inputs from BMZ TDA's perspective. We are grateful to the consultants and colleagues who contributed to the screening, retrieval and extractions processes: Ingunn Storhaug, Joshua Mwakalikamo, Shona Putuka, and Tomasz Kozakiewicz. We would also like to thank the members of our advisory group: Amber Fletcher, Molly Melching, Tania Paffenholz, Nyachangkuoth Tai, Lili Mottaghi, Laura Shepherd, Katie West-Allen, Anu Mundkur, Sarah Taylor, Caitlin Shannon, Bidisha Barooah, Ada Sonnenfeld, Naila Kabeer and Mohammad Naciri. The authors of this systematic review would like to specifically thank the following contributing authors: Adedayo Olofinyehun, Ashiqun Nabi, Bidisha Barooah, Cem Yavuz, Denise Ferris, Joshua Furgeson, Lina Khan, Sanghwa Lee and Umer Gupta. Finally, thanks to all the authors who shared unpublished and forthcoming papers and any additional data that helped improve the review.

Plain language summary

This 3ie systematic review aims to assess the effects of gender-specific and transformative interventions on women's empowerment and gender equality in fragile and conflict affected states (FCAS) with high levels of gender inequality and their contribution to building peaceful and inclusive societies.

Background

Across the globe, women face tremendous challenges when it comes to equitable access to resources, exercising meaningful agency and decision-making power and aspiring to and accomplishing achievements. Despite gains made in recent decades towards inclusion and empowerment, significant challenges still remain for women, particularly in fragile and conflict-affected settings where these challenges are exacerbated.

To help identify what approaches can improve women's empowerment in fragile settings with high levels of gender inequality, this systematic review synthesises the evidence on the effects of 14 different types of gender-sensitive and gender-transformative interventions.

What studies are included?

Our systematic review includes 104 unique studies covering 55 identified programmes and 32 linked impact evaluations papers, in addition to 90 linked qualitative and process evaluations. These studies took place in 29 countries identified as being particularly fragile and having a high level of gender inequality and all were published between 2009 and 2021.

Of the quasi-experimental studies included, almost half were deemed as having a high risk of bias, and we assessed nearly all remaining quasi-experimental studies as having some risk of bias concerns. Among the studies using an experimental design, only around one third were assessed as having a high risk of bias.

Do gender-specific or transformative interventions increase women's empowerment and gender equality?

Most of the identified interventions have significant positive effects on empowerment and gender equality outcomes closely related to the purpose of the intervention. Asset transfers and cash transfers show some of the largest effects on women's access to and ownership of assets, credit and income. Village savings and loan associations and institutional provisions of savings and loans have the largest effects on women's capacity to use and understand financial, banking and business services effectively. Life skills and capacity building programmes realise their largest effects on women having better life skills. Inclusive community-driven development has the largest effect on increased representation of women in local and sub-national civil and political processes.

Through our qualitative synthesis, we explored potential explanations for the lack of effects further along the causal chain. Despite interventions demonstrating a positive effect on the outcomes they specifically targeted, there is little evidence of trickle-down

effects on other empowerment outcomes, which includes resources, agency and achievements. Few interventions led to positive effects on outcomes such as the attitudes of men towards women, agreement with justifications for intimate partner violence, or women's participation in decision-making processes at the household and community levels.

The *Grading of Recommendations, Assessment, Development and Evaluations* (GRADE) methodology was used to summarise and rate certainty of bodies of evidence. Of the 190 included bodies of evidence (including those based on single studies), the majority of the evidence is from very low (57%) and low (27%) confidence, while 14% are of moderate and 2% are of high confidence. Summary of findings tables in the main report provide concise explanations of how certainty of evidence ratings is assessed, allowing readers to understand any concerns surrounding a given effect estimate.

What do the findings of this review mean?

Qualitative findings show that gender and social norms acted as significant barriers to programme uptake across multiple intervention types, including cash transfers and safe spaces. Policy makers must ensure they recognise embedded power dynamics and specifically address patriarchal norms, as without these components, interventions are unlikely to succeed in their stated aims. Without addressing these barriers, secondary outcomes that are not specifically targeted by the interventions are unlikely to be affected. For example, only one intervention group, self-help groups, resulted in a significant reduction of violence at the household level and no interventions reduced intimate partner violence.

There is a lack of evidence from certain regions and intervention types. Very few studies came from the Arabian Peninsula, although it is a highly relevant region considering the focus of our review. We also observe gaps of evidence related to gender-specific and transformative interventions focusing either on recovery and relief or on the active role of women in peacebuilding.

The body of evidence covered in our review suggests that interventions explicitly addressing gender norms are well-positioned to strengthen women's empowerment in fragile and conflict-affected settings. Targeting the right beneficiaries, designing interventions to be gender-transformative, and implementing with the context in mind are key to realising the goals of empowering women, promoting gender equity and building more peaceful and inclusive societies.

Although only a few interventions focused on the direct role of women as agents for peacebuilding, a large body of evidence focused on the indirect impact of gender equality towards inclusive and peaceful societies. Addressing women's empowerment and gender equality is then a major contribution to the reduction of violence and the building of sustainable peace.

Executive Summary

Background

In 2020, about 23 per cent of the world's population lived in fragile contexts, including 76.5 per cent of all those living in extreme poverty (OECD 2020). The World Bank reports that over the last ten years, the number of fragile and conflict-affected states and situations (FCAS) has increased from 36 to 39 countries or states, with 17 classified as being in a 'situation of medium- to high-intensity conflict' (World Bank 2020). The Institute for Economics and Peace reports a similar deterioration of peace, finding a 0.34 percentage point drop last year in their Global Peace Index, which measures societal safety and security, ongoing domestic and international conflict and levels of militarisation (Institute for Economics and Peace 2020).

The year 2020 also marked the 20th anniversary of the United Nations Security Council Resolution 1325 on Women, Peace and Security, a resolution that recognises the inordinate impact of violent conflict and war on women and girls and their role in sustaining peace. Despite evidence demonstrating the importance of promoting gender inclusion, equality and empowerment in fragile and conflict-affected states and situations (FCAS), evidence gaps remain with respect to the effectiveness of such interventions (OECD 2013; Cornwall 2016). A synthesis on the effects of interventions targeting women's empowerment and gender equality is needed to inform decisions about how to optimise the use of limited funding to achieve the ambitious goal of achieving peace in inclusive societies. This comprehensive systematic review synthesises the findings of 104 studies (from 136 impact evaluations) evaluating the effects of a range of different programmes in 29 FCAS.

Objectives

The primary objective of this review is to identify, assess and synthesise evidence on the effect of gender-specific and gender-transformative interventions that fall within the United Nations Security Council Resolution 1325 four pillars on women's empowerment and gender equality in FCAS settings (Participation, Prevention, Protection, Recovery and Relief). The systematic review aims to facilitate the use of evidence in informing policy and practice decisions within the field of transition aid, particularly related to gender focused programming. Our second objective is to assess how these interventions contribute to inclusive and sustainable peace in conflict affected situations.

To achieve these objectives, we answered the following questions:

1. What are the impacts of gender-transformative and specific interventions on women's empowerment and gender equality in FCAS?
2. What are the effects of these interventions on sustainable peace?
3. To what extent do effects vary by population group, ecological level and intervention type?
4. What are contextual barriers to and facilitators of intervention effectiveness?

While the evidence base for this review was extensive and broad, stretching across 14 intervention groups, the stark lack of evidence in the recovery and relief pillar represents an open-ended route for future research.

The qualitative evidence base for questions 3 and 4 provided findings on many design features as well as implementation and contextual considerations. However, this evidence base was not spread equally across all intervention groups, with six groups not covered by corresponding qualitative evidence at all. We further identified a relative lack of data and evidence to conduct moderator analysis informed by the qualitative review findings.

Study selection criteria

We included studies examining gender-specific or gender-transformative interventions in fragile and conflict-affected states and situations (FCAS) that were published between 2000 and 2021. Our review included studies that use experimental or quasi-experimental study designs and measure women's empowerment or gender equality for peaceful and inclusive societies as outcomes. We included studies targeting all ages and genders in the eligible countries based on the Gender Inequality Index, the Fragile States Index, the World Bank FCAS list, the OECD FCAS list and the Women Peace and Security Index. Within these four categories, we analysed studies focusing on 14 intervention types: all-women police stations, asset transfers, cash transfers, community dialogues and reconciliations, community-based services, discussion groups, inclusive community-driven development, institutional provision of loans and savings, quotas, safe spaces, self-help groups and village savings and loan associations, sensitisation campaigns, technical and vocational education and training.

We also included qualitative studies, descriptive quantitative studies, process evaluations and project documents linked to the interventions that were evaluated in the included experimental and quasi-experimental studies.

Search

We searched a broad range of electronic academic databases, internet search engines, websites and collections. We screened titles and abstracts of approximately 100,000 papers, the majority of which were excluded either due to their gender inequality or FCAS index, or because they were not gender-specific or gender-transformative. We screened the full text of 1,330 papers for inclusion. After a final screening by at least two reviewers, we included 136 papers, corresponding to 104 unique studies covering 55 different programmes.

Included studies

The included studies cover programmes across 29 low- and middle-income countries. This includes 78 studies from Sub-Saharan Africa, 29 from South Asia, four from the Middle East and North Africa, two from East Asia and the Pacific, and two from Latin America and the Caribbean. All included studies were published between 2009 and 2021, with most from peer reviewed journal articles published between 2016 and 2020. Based on reported sample sizes, we estimate that the studies included data from over 290,494 beneficiaries.

Among the 30 quasi-experimental design studies, only 3% were appraised as low risk of bias, while 45% were high risk of bias and 52% were appraised as having some risk of bias concerns. Among the 91 included studies using an experimental design, 28% of the studies had low risk of bias and 28% had high risk of bias.

Data collection and analysis

We critically appraised included studies using standard appraisal tools and collected data on programme design, process, implementation and contextual factors. We extracted effect size data on relevant outcomes from included studies, calculating standardised mean difference, (SMD) standard errors and 95% confidence intervals (CIs) for all studies. Two independent reviewers coded all studies with a quantitative expert providing quality assurance.

Studies were then synthesised using random-effects meta-analysis, estimating average effects of different gender-specific and transformative interventions and associated heterogeneity. We conducted sensitivity analyses and explored potential sources of variation among the studies. For the synthesis of qualitative evidence, we used thematic synthesis, organising analytical themes according to the intervention categories.

Results

On average, all the identified interventions have significant and positive effects on women's empowerment outcomes closely related to the purpose of the intervention. For example:

- Asset transfers (SMD 0.34 [95% CI: 0.22 to 0.46]) and cash transfers (SMD 0.22 [95% CI: 0.12 to 0.31]) realise some of the largest impacts on women's access to and ownership of assets, credit and income (an estimated 13.3% change in the intervention group compared to the control group for asset transfers, and an estimated 8.7% change in the intervention group compared to the control group for cash transfers);
- Village savings and loan associations (SMD 0.26 [95% CI: 0.00 to 0.52]) and institutional provisions of savings and loans (SMD 0.34 [95% CI: 0.05 to 0.63]) have their largest impacts on women's capacity to use and understand financial, banking and business services effectively (an estimated 10.3% change in the intervention group compared to the control group for village savings and loan associations, and an estimated 13.3% change in the intervention group compared to the control group for institutional provisions of savings and loans);
- Life skills and capacity building programmes realise their largest impacts on women having better life skills (SMD 0.16 [95% CI: 0.12 to 0.21], an estimated 6.4% change in the intervention group compared to the control group); and
- Inclusive community-driven development has the largest impact on increased representation of women in local and sub-national civil and political processes (SMD 0.27 [95% CI: 0.17 to 0.37], an estimated 10.6% change in the intervention group compared to the control group).

Average effects are generally smaller in magnitude, or negligible, for outcomes less aligned with the main purpose or mechanism of the intervention. For example, we do not find an effect on intimate partner violence from any intervention. Through our qualitative synthesis, we explored potential explanations for the lack of trickle-down effects on behavioural outcomes.

Some interventions have rarely been studied—there is insufficient evidence to conduct a synthesis for all-women police stations and community-based services. Many

intervention-outcome combinations were not studied at all. In some cases, this indicates an important evidence gap, while in other cases the intervention would not be expected to impact the outcome. No intervention has a significant negative effect on any outcome.

We conducted thematic synthesis for eight intervention groups from 90 linked qualitative studies. The other six groups do not have sufficient studies to conduct a qualitative synthesis. These 90 papers cover 55 unique programmes across our evidence base and provide a rich complement to our quantitative findings by generating 149 descriptive themes across all studies and 37 analytical themes. Findings emphasise that context is a critical aspect of implementation; ranging from asset transfers to safe spaces interventions, programme participants cited gender and social norms as a significant barrier or opening factor for uptake. Overall, the qualitative findings give a look into how positive and significant effect sizes (such as technical and vocational education and training) in the quantitative synthesis are buttressed in their impact on behavioural outcomes by factors related to programme implementation.

Qualitative findings suggest that economic interventions, including asset and cash transfers, should be tailored to both geographical and societal settings, to account for semi-arid climates or entrenched patriarchal norms that often prohibit women from owning assets such as livestock. Gendered household and community roles often restrict women beneficiaries' ability to manage or decide how to spend money. Furthermore, implementation and community leadership are both critical pillars of influence in sensitising communities and households to participate in an intervention. In other types of economic interventions, such as self-help groups, mentoring and peer exchange is a key factor to help women participants actively challenge prevailing gender norms by collectively working towards economic empowerment.

Based on qualitative evidence, for inclusive community-led development, representation within committees in terms of gender and background diversity as well as support and integration with local government structures, emerge as key contributors to intervention success. Recognising embedded power dynamics also enables women to better participate in local and sub-national political processes. However, entrenched patriarchal norms necessitate an increased buy-in by these stakeholders such as local, cultural or religious leaders for appropriate uptake at the community level. Open dialogue forums may support this. In contexts where forced marriage is common, safe spaces and sensitisation campaigns are promising approaches to support adolescent girls' contexts by providing a physical safe location for sharing experiences and learning new skills. Regular and sustained interaction contributed to maintained uptake and long-lasting engagement at the community level.

Finally, the qualitative findings suggest that the success of life skills and capacity building programming is often determined by contextual and societal factors that can limit programme implementation. This means that embedding programming within these existing norms can enhance a community's acceptance of the offered programme. In contrast, the more embedded a programme is, the more difficult it may be to enact radical change challenging existing norms. Like our findings in technical and vocational education and training interventions, group mentorship and peer exchange offer a route to collective action and socialisation that can encourage group uptake of a life skills and capacity building programme.

Intervention-level results

Asset transfer

Asset transfer interventions yield significant and positive effects on outcomes in two of the three dimensions of the resources-agency-achievement framework. Although asset transfers interventions target outcomes across all three dimensions of women's empowerment, they only produce positive and significant results on access to assets, credit and income and women's participation in decision making. We find no evidence that asset transfer interventions improve achievement-related outcomes, such as improved self-image and confidence, the reduction of intimate partner violence and the improved quality of relationships with community and households.

Qualitative findings suggest that the provision of assets or tools for cultivation or livestock farming can generally support asset endowment, creditworthiness and income. If assets are utilised productively, it opens pathways to credit use options and income-generating activities. This was especially evident in multi-component asset transfer programmes, such as the Income-Generating Vulnerable Group Development (IGVGD) programme in Bangladesh, which added a microfinance facility which allowed women to use loans for a multitude of purposes (Ahmed et al. 2007). Additionally, agency is bolstered when women are able to own and manage animals and/or crops. Women can then experience improved decision-making and social status within households and around communities.

Cash transfer

We find that cash transfers have positive effects on outcomes across the all three dimensions of the resources-agency-achievement framework. We observe improvements in access to assets, credit and income, access to livelihood support services, engagement in enterprises, reduction of acceptability of violence, and increased political representation. However, we do not find any effects of cash transfer programmes on women's control over their bodies, freedom of movement and association, participation in decision making, better life skills, community support for women's needs, safer communities and reduced frequency of violence.

Many cash transfer programmes included additional components (e.g., a training component), which we hypothesised would be related to larger effects. However, our moderator analyses revealed no differences between the cash transfers alone and cash transfers plus an additional component on any of the women's empowerment outcomes. Similarly, we generally find no difference in the size of effects between conditional and unconditional cash transfers. The exception to this is women's participation in decision making, where we actually observe a significantly *smaller* effect of conditional transfers as compared to unconditional transfers.

Qualitative findings suggest that economic interventions should be tailored to both geographical and societal settings, including entrenched patriarchal norms. This is because gendered household and community roles often restrict women beneficiaries' ability to manage or decide how to spend money. Furthermore, implementation and community leadership are both critical pillars of influence in sensitising communities and households to participate in an intervention.

Community-based services

The effects of community-based services on women's empowerment index outcomes are unknown. The one study that is included under this intervention group only examined outcomes related to access to and ownership of assets, credit and income and access to decent work to very small, positive effects. None of the effects of this single primary study reached statistical significance. Further, all of the evidence from the community-based services intervention is of low certainty, indicating that future studies of these interventions may produce different results.

Inclusive community-driven development

Inclusive community-driven development (ICDD) only yields significant and beneficial effects in the achievement dimension of the Resources-Agency-Achievement framework. There are only positive and significant effects on the increased political representation of women. The other outcomes targeted by ICDD programmes, such as freedom of movement and association, positive attitudes towards claiming rights, reduction of justification of violence, support from community on women's needs and the pursuit of safer communities, are not affected by the community-driven development interventions. This echoes existing synthesis that found ICDD programmes use existing social cohesion rather than building it (White et al, 2018).

Based on qualitative evidence, representation within committees in terms of gender and background diversity as well as support and integration with local government structures emerge as key contributors to intervention success. Recognising embedded power dynamics also enables women to better participate in local and sub-national political processes. However, entrenched patriarchal norms necessitate an increased buy-in by these stakeholders such as local cultural or religious leaders for appropriate uptake at the community level. Regular and sustained interaction contributed to maintained uptake and long-lasting effects at the community level.

Institutional provision of loans and savings

We find institutional provision of loans and savings only have positive effects on resources, as measured by access to or ownership of assets, credit and income, but fail to have significant and positive effects on the increased participation of women in decision-making, the improved capacity of women entrepreneurs, and on women's empowerment index.

Quotas

We are unable to provide strong conclusions regarding quotas due to the low number of included studies using quota interventions. Women's improved attitudes towards taking action to claim their rights was the only outcome with two studies contributing effects, but the synthesis found a null effect. This must be interpreted with caution given the very small number of included studies. Further, all of the evidence from the quota interventions is of low certainty, indicating that future studies of quota interventions may produce different results.

Self-help groups and village savings and loans associations

We find that SHGs and VSLAs have positive effects on outcomes across the resources-agency-achievement framework, including capacity to use and understand financial and banking services, access to decent work, increased freedom of movement and

association, rights awareness, positive attitudes towards acting to claim rights, better life skills, representation in political processes, prevention of violence, decreased violence at the household level and women's empowerment index. This type of intervention, however, fails to affect women's access to and ownership of assets, credit and income, reduction of justification of violence, participation in decision-making, participation in the community, support from the community, more positive attitudes towards women, improved self-image and confidence and better relations with community and household members.

The qualitative evidence base supports this; social capital gains and skill acquisition seems to directly augment economic-focused outcomes rather than association and social-focused outcomes. SHGs and VSLAs are an attractive and convenient option for women who have been systematically excluded from formal economic participation. However, similar to quantitative findings, other factors such as market access and conditions minimise the collective gains for other outcomes.

Technical and vocational education and training

We find that TVET interventions have positive effects on outcomes across the three empowerment dimensions. This includes access to assets, credit and income, improved capacities of women entrepreneurs, reduction of forced and child marriage, political representation of women, and improved image and self-confidence. However, TVET fails to have significant and positive effects on women's control over their body, increased freedom of movement and association, reduced justification of violence, improved life skills, participation in decision making, community support, reduction of violence, quality of relationships, and women's empowerment index. Though we hypothesized that TVET would affect women's access to decent work, this is not supported by the evidence, which found a null effect of TVET programmes in fragile and gender unequal states on women's access to work. This is reflected in qualitative data as an effect of social norms since, although we observe a positive effect on capacities, this does not necessarily transform into employment for women. This is often due to conflating social norms that TVET interventions need to consider in varying contexts. Unless programmes are designed to address underlying norms that hinder women's participation in such a programme, there is likely little progress to be made regarding household power relations and women's empowerment within communities. For example, providing childcare and transportation stipends to women participants can offset the perception that they are abandoning their domestic work and offset the opportunity cost of attendance.

The qualitative evidence base indicates a need to design additional programme components within TVET in order to target outcomes beyond individual skill development and improved attitudes. Group mentorship and peer exchange offer a route to collective action and socialisation that can encourage group uptake of a life skills and capacity building programme. The most successful programmes in terms of uptake and outcomes were ones that enhanced the social aspect as well as the livelihoods aspect.

All-women police stations

We were unable to draw conclusions for all-women police stations due to the low number of included all-women police stations studies. The only included study found a significant negative effect on employment of women in law enforcement (as measured by the daily

proportion of cases assigned to women, and a null effect on women having improved and equitable access to justice. We find evidence from all-women police stations to be of low or very low certainty, indicating that future studies of all-women police stations may produce different results.

Sensitisation campaigns

Sensitisation campaigns have a beneficial effect on outcomes from two of the three dimensions of the women's empowerment framework. There is little data on resource outcomes from this type of intervention. Beneficial effects of sensitisation campaigns were found on women having more positive attitudes towards claiming their rights, increased awareness of issues affecting women and women's empowerment index. However, sensitisation campaigns fail to affect reduction of agreement with justification and occurrence of violence, improved self-image, support from households and communities, reduction of physical and sexual IPV and improved quality of relationships with household and community members.

Based on the qualitative findings, sensitisation campaigns can be limited and restricted by social factors such as gender norms and roles. Programmes that, for example, worked with constructive religious and social messaging, often through community leaders in religious institutions, were able to act as a useful channel for disseminating information. Religious leaders and institutions acted as a conduit for participants to better understand gender imbalances and how to address them. Changes that are related to altering human behaviour through learning and communication often take time to realise.

Additionally, qualitative findings reveal that training that specifically targets men and boys is useful in building awareness on IPV and conflict resolution to advocate for men directly challenging gender norms. Training men to better understand women's roles is an important exercise in encouraging men to understand women's reality in the household and community.

Community dialogues and reconciliations

We are unable to make strong conclusions related to community dialogue interventions given that only two studies were included. While the synthesis of the two studies found a positive and significant effect on an achievement-related outcome: improved attitudes and support for women's economic, social and human rights by households and communities, this result should be interpreted with caution. We have a low level of certainty of this body of evidence, suggesting that future studies of community dialogue interventions may produce different results.

Discussion groups

We are unable to make strong conclusions related to discussion group interventions given that only two studies were included. While the synthesis of the two studies found a positive and significant effect on an achievement-related outcome: improved attitudes and support for women's economic, social, and human rights by households and communities, this result should be interpreted with caution as we have low certainty in the body of evidence related to this effect, suggesting that future studies may produce different results.

Safe spaces

Safe spaces only have a beneficial effect on the resource dimension of the empowerment framework through increased access to or ownership of assets, credit and income. However, safe spaces fail to affect agency and achievement related outcomes in our sample of studies, including increased capacity of women to understand and use financial, banking, and business services effectively, better control over their bodies and sexual health, freedom of movement and association, support for or occurrence of child and forced marriage, participation in decision making by women at the household or community level, improved attitudes, self-image and confidence, improved quality of relationships between women and their household and community members.

Although we would expect it to increase safe communities improve skills and capacity building and reduce violence, there was a null effect of safe spaces on these outcomes in fragile and gender unequal states. Contextual findings from qualitative synthesis show that uptake and participation are buttressed by the maintenance of physical space and planned activities to ensure safety of participants and implementing partners. Including mentorship as a key aspect provides participants and their family members with confidence in dismantling harmful social norms related to the rights of women and girls.

Additionally, targeting women or girls in particular age brackets proves to be successful. For example, a programme that targeted adolescent girls of marriageable age showed success in preventing early and forced marriage in the target community by combating harmful social and religious beliefs that pressured or forced these girls into marriage.

Life, social, and livelihood skills and capacity building

Life, social, and livelihood skills and capacity building (LSCB) interventions result in positive and significant effects on outcomes within two of the three dimensions of the resources-agency-achievement framework. We find beneficial effects of LSCB programmes on both resources and agency outcomes. Specifically, understanding and use of financial and business services, better control over their body and sexual health, freedom of movement and association, awareness of rights, reduced agreement with justification of violence, equipment with life skills, participation in decision-making. However, LSCB interventions fail to effect access to and ownership of assets, credit, and income, reduction of child and forced marriage, participation in and support from the community, reduction of occurrence of violence, quality of women's relationships and women's self-image and confidence.

The qualitative findings suggest that the success of life skills and capacity building programming is often determined by contextual and societal factors that can make or break the success of a programme. This means that embedding programming within these existing norms can enhance a community's acceptance of the offered programme. In contrast, the more embedded a programme is, the more difficult it may be to enact radical change challenging existing norms. Similar to our findings in technical and vocational education and training interventions, group mentorship and peer exchange offer a route to collective action and socialisation that can encourage group uptake of a life skills and capacity building programme.

Implications for research

Based on the findings of our review, we note the following implications for policymakers and programme implementers:

- **Examine interventions targeting gender components of an intervention and its associated outcomes.** Our review and search process highlights that there is a large number of studies assessing effects on gender outcomes, but a much smaller literature of evaluations of gender-specific or gender-transformative interventions. For example, even though all our 14 intervention groups focused on gender equality and women's empowerment in fragile and conflict-affected states and situations (FCAS), a minority of the included studies actually has a specific focus on women's direct involvement in peacebuilding.
- Most of the studies focusing on gender-specific and gender-transformative interventions are concentrated in specific regions, mainly Sub-Saharan Africa and Southeast Asia. **More research is especially needed in the Middle East and North Africa**, given the few studies and large number of FCAS in the region.
- To provide programme designers and evaluators with a detailed view of local context, both **methodological approaches and data generation are needed**. The methodological approaches will allow to better conceptualise and measure complex issues related to women's empowerment. These approaches will then be complemented by high quality gender-specific and transformative data at community and household levels that will inform interventions designs and implementations.
- There is a **need for indices measuring gender-related outcomes at the local or community level**. Measurement of fragility should incorporate these complex differences at the local level.
- To make meaningful knowledge contributions and to produce reliable evidence, evaluation studies **need to follow rigorous methodological designs, and report their research in such a way that allows for transparency and assessment of study quality**. Good impact evaluations will register a pre-analysis plan, establish an appropriate counterfactual (either through random assignment or through an appropriate quasi-experimental design) and collect baseline data from all participants. They will collect outcome data from all participants at intervals that are equivalent across the treatment and control conditions. In reporting their findings authors will account for multiple-hypothesis testing in their analysis, and report on all outcomes, regardless of significance. When reporting on these evaluations, authors should also report information related to the assignment mechanism, baseline balance among treatment and control groups, attrition and differential attrition, deviations from the intended intervention, potential spill-over effects, and potential performance bias. Ethical clearances and external validity should also be explicitly discussed.
- To support the full integration and use of qualitative data within impact evaluation research, **more detailed reporting on the methods used to generate qualitative data within impact evaluations reports is required**. While we found a healthy body of impact evaluations reporting qualitative data in the study reports, the methods used to generate and analyse such data were not reported at a similar rate. Methods sections should equally allude to the process of the qualitative research to allow for an objective assessment of the qualitative research.

- A coordinated effort on the part of donors and implementers is needed to ensure that **researchers use validated outcome measurement tools for impact evaluations**. In addition to consistency within sectors, this also includes investing in instrument validation.
- Finally, to assess the long-term impact of interventions and sustainability, where resources allow, future evaluations should try to **collect outcomes after longer follow-up periods**.

Implications for policy and practice

Based on the findings of our review, we make the following general recommendations for policymakers and programme implementers:

- **Design and implement specific programme components that target empowerment outcomes.** This requires context analyses before, during and after programming that have an explicit focus on gender to provide an understanding of the barriers to empowerment in each context. This would include updated context analyses before, during and after programming that have an explicit focus on gender.
- **Consider multiple empowerment approaches and pathways tailored to local norms and practices.** Each implementation component should carefully apply a gender lens to assess whether implementation choices are gender-sensitive or transformative, which is essential for generating a conducive environment in which empowerment-related programme components can be established and succeed.
- **Programme designers should emphasise on the potential of women peer-to-peer exchanges and collective-building programme components such as mentoring and women's groups which create social capital and networks among women.** These design features allow women to support each other in finding tailored solutions, e.g., sharing innovation in how to operate in male-dominated industries, as well as providing a social safety net that can provide business support and access to assets.
- **Try to reach participants through local agents and structures.** The gains of embedding programmes in existing structures can enhance legitimacy and improve access to participants. However, programmes should be careful to not further entrench long-standing gender-restrictive norms and practices that can undermine empowerment outcomes.
- **Target the right beneficiaries.** While specific focus should be on women and girls in the most vulnerable communities or households, the targeting process needs to be geographically inclusive. Targeting procedures should be transparent, tailored to minimise beneficiary selection errors and maximise trust in administrative processes, programme participation, and the achievement of aimed outcomes.
- **Fund additional gender interventions with rigorous impact evaluations.** Many studies focus exclusively on peacebuilding interventions in fragile contexts, and do not have a gender component inclusive to our scope. This might be due to both the challenge of running impact evaluations in fragile contexts and measuring gender-related outcomes.

Tables 1 - 4 below provides recommendations for specific types of interventions.

Participation

Table 1: Participation intervention recommendations

Intervention	Intervention works on	Intervention does not work on
Asset transfers	<p>Resources</p> <ul style="list-style-type: none"> Women have increased access to and ownership of assets, credit and income. <p>Agency</p> <ul style="list-style-type: none"> Increased participation in decision making by women at the household or community level <p>Achievement</p> <ul style="list-style-type: none"> Effect unknown - insufficient evidence 	<p>Resources</p> <ul style="list-style-type: none"> Women can access decent work (formal and informal employment) <p>Agency</p> <ul style="list-style-type: none"> Effect unknown - insufficient evidence <p>Achievement</p> <ul style="list-style-type: none"> Women have improved attitudes, self-image and confidence Reduced frequency and distribution of types of violence by an intimate partner Improved quality of relationships between women and their household and community members Women's empowerment/equality index
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> Identify and secure multiple convenient locations for training activities (e.g., agricultural education workshops) and for asset collection to improve participation and productive use of assets. Design quality (participatory and 'hands-on'), comprehensive and inclusive training for recipients to promote understanding and successful adoption of relevant skills to carry out activities related to asset transfer such as agricultural production. Conduct quality checks for asset transfers and provide necessary supporting inputs (e.g., fences and irrigation equipment). Ensure that adequate household level social and labour support systems for women beneficiaries are available. For instance, through sensitisation or training of family members, particularly husbands. Explore multicomponent interventions that include legal support, microcredit, cash transfers, and capacity building more holistically to support the productive and sustained use of assets 	

	Contextual considerations to keep in mind <ul style="list-style-type: none"> • Traditions and norms, such as gendered divisions of labour and women's rights to land hinder the type of activities women can participate in and eventual empowerment outcomes Programmes have attempted to address this through gender sensitisation campaigns targeted at partners. • Programme activities (including compulsory training) may disrupt usual women's day-to-day work obligations and affect time use preferences 	
	Good practice examples: <ul style="list-style-type: none"> • Burkina Faso's Enhanced-Homestead Food Production (E-HFP) emphasises the need for additional input support to complement asset transfers meant for agricultural purposes. Specifically, whilst there was an indication of improved ability for women to use land for agricultural purposes, primary obstacles included insufficient water access, pest and animal attacks. Hence, the provision of additional inputs in the form of irrigation equipment, fences and pest control chemicals could have mitigated these challenges and enhanced the benefits realised from the assets transferred. • Bangladesh's Targeting the Ultra Poor (TUP) illustrates the importance of ensuring adequate social and labour support systems for women beneficiaries. The livestock and poultry that women received raised women's workloads at home and in some cases prevented women from being employed outside their homes. As an alleviating factor, the women benefited from children and husband assistance with selected tasks related to the care of their livestock. • The Bangladesh's Income-Generating Vulnerable Group Development (IGVGD) asset transfer programme included a microfinance facility which enhanced economic involvement as it allowed women to use the borrowed money for various purposes including paying for rental on land they used to start and carry out agricultural activities related to the agricultural assets they had received. 	
Intervention	Intervention works on	Intervention does not work on
Cash transfers	Resources <ul style="list-style-type: none"> • Women have increased access to and ownership of assets, credit and income. • Women and girls have equitable access to livelihood support services. • More women engaged in other micro, small, and medium-sized enterprises. 	Resources <ul style="list-style-type: none"> • Women can access decent work (formal and informal employment) Agency <ul style="list-style-type: none"> • Women have more and better control over their bodies and sexual health • Women have increased freedom of movement and association • Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them

	<p>Agency</p> <ul style="list-style-type: none"> • Reduced percentage of women agreeing with certain reasons that justify violence against women and girls <p>Achievement</p> <ul style="list-style-type: none"> • Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post-conflict restoration 	<ul style="list-style-type: none"> • Reduced support for or occurrence of child and forced marriage • Increased participation in decision making by women at the household or community level • Women participate more in their community <p>Achievement</p> <ul style="list-style-type: none"> • Women have improved attitudes, self-image and confidence • Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members • Safer and more secure household, communities and areas/territories for women, girls, men, and boys • Reduced frequency and distribution of types of violence by an intimate partner • Improved quality of relationships between women and their household and community members • Women's empowerment/equality index
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> • Incorporating couples' relationships, social and livelihood skills, mentoring and capacity building components can support household relationships and conflict resolution skills • Targeting the most vulnerable households and women who struggle to meet their basic needs and those who do not have access to credit or capital to start businesses • Targeting family units (compared to individuals) and carrying out clear awareness-raising campaigns supports participation • When determining transfer size, considerations should be made for inflation and household size, but relatively high cash transfers may discourage participation in wage labour activities. • Mentorship, training and support from family members promotes adherence and compliance to programme requirements, especially in conditional cash transfer programmes. • We find no difference in the effects of conditional transfers compared to unconditional transfers. This should be carefully considered given the higher administrative costs of conditional transfers. 	

	Contextual considerations to keep in mind <ul style="list-style-type: none"> • Social norms, existing gender roles and rigid restrictive beliefs can preserve men’s dominance (regardless of women’s income contribution) in household decision making, including participation in these programmes. 	
	Good practice example: <ul style="list-style-type: none"> • After noticing slow uptake and lack of awareness in communities, the implementation team from the Zomba Cash Transfer Programme (ZCTP) in Malawi adopted a sensitisation campaign to help prospective participants understand the objectives of the intervention, which improved participation rates. • In the Takaful and Karama Programme in Egypt, severely poor people used a large part of the cash transfer to take care of their basic needs including food, clothing and medical expenses following the receipt of the cash transfer. Notably, a significant number of the beneficiaries were also highly in debt. Hence, the cash transfer alleviated short term needs for vulnerable beneficiaries in order to focus on more sustainable, long-term economic goals. • The mentorship and coaching components in Kenya’s Rural Entrepreneur Access Project (REAP) implemented by the BOMA Project helped ensure that participants adhered to the programme, enhanced school attendance, financial literacy and beneficiaries’ knowledge. 	
Intervention	Intervention works on	Intervention does not work on
Community-based services	Resources <ul style="list-style-type: none"> • Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> • Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> • Effect unknown - insufficient evidence 	Resources <ul style="list-style-type: none"> • Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> • Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> • Effect unknown - insufficient evidence
	Intervention design suggestions to maximise benefits <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
	Contextual considerations to keep in mind <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
	Good practice example: <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	

Intervention	Intervention works on	Intervention does not work on
Inclusive community-driven development	<p>Resources</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Agency</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Achievement</p> <ul style="list-style-type: none"> • Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post-conflict restoration 	<p>Resources</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Agency</p> <ul style="list-style-type: none"> • Women have increased freedom of movement and association • Women have more positive attitude towards taking action to claim their rights • Reduced percentage of women agreeing with certain reasons that justify violence against women and girls <p>Achievement</p> <ul style="list-style-type: none"> • Increased attention and focus on the needs and priorities of women and girls, and other vulnerable groups during relief and recovery in conflict and post-conflict settings • Increased community support for women's and children's human, economic and legal rights • Women have improved attitudes, self-image and confidence • Safer and more secure household, communities and areas/territories for women, girls, men, and boys
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> • Composition of committees and representatives in decision-makers structures mattered largely for the legitimacies of these bodies within the communities targeted by the intervention (e.g., mandated parity) • Include capacity building initiatives for established governance bodies • Encourage collaborations between institutions to mitigate potential conflicts between multiple governance systems or bodies through regular meetings and combined working groups and balance politics and power relations between existing and new bodies. 	
	<p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> • Governance structures introduced may coexist with traditional governance authorities (e.g., local chiefs) as well as local government structures. Navigating this co-existence requires a careful balancing of politics and power relations, which limited the space for project implementation and reach in some instances. 	

	<ul style="list-style-type: none"> Powerful elites with a vested interest may undermine project implementation when the legitimacy and capacity of governance bodies is a challenge. 	
	<p>Good practice example:</p> <ul style="list-style-type: none"> The Democratic Republic of the Congo's Tuungane programme featured mandated gender parity in the composition of project committees, which tended to facilitate more women to participate in the programme. In Afghanistan's National Solidarity Programme (NSP), collaboration between entities or institutions such as cultural or religious leaders, public and private agencies was crucial. Bottom-up initiatives inclusive of diverse groups of community members helped legitimise the intervention's committees as local institutions. Additionally, in Afghanistan, project success hinged on timely disbursement of funds, regular meetings and follow-up on community initiatives. The timing of project implementation centred around it not coinciding with annual harvesting seasons, and the time use of community members, particularly women. 	
Intervention	Intervention works on	Intervention does not work on
Institutional provision of loans and savings	<p>Resources</p> <ul style="list-style-type: none"> Women have increased access to and ownership of assets, credit and income. <p>Agency</p> <ul style="list-style-type: none"> Effect unknown - insufficient evidence <p>Achievement</p> <ul style="list-style-type: none"> Empowerment index 	<p>Resources</p> <ul style="list-style-type: none"> Improved capacity of women entrepreneurs <p>Agency</p> <ul style="list-style-type: none"> Increased participation in decision making by women at the household or community level <p>Achievement</p> <ul style="list-style-type: none"> Women's empowerment/equality index
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	<p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	<p>Good practice example:</p> <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
Intervention	Intervention works on	Intervention does not work on
Quotas	<p>Resources</p> <ul style="list-style-type: none"> Effect unknown - insufficient evidence 	<p>Resources</p> <ul style="list-style-type: none"> Effect unknown - insufficient evidence

	Agency <ul style="list-style-type: none"> Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> Effect unknown - insufficient evidence 	Agency <ul style="list-style-type: none"> Women have more positive attitude towards taking action to claim their rights Achievement <ul style="list-style-type: none"> Effect unknown - insufficient evidence
	Intervention design suggestions to maximise benefits <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	Contextual considerations to keep in mind <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	Good practice example: <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
Intervention	Intervention works on	Intervention does not work on
Self-help groups and village savings and loans associations	Resources <ul style="list-style-type: none"> Increased capacity of women to understand and use financial, banking, and business services effectively Women have increased access to and ownership of assets, credit and income Women can access decent work (formal and informal employment). Agency <ul style="list-style-type: none"> Women are more aware of their rights and the roles and responsibilities of duty bearers. Women have more positive attitude towards taking action to claim their rights. Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them. 	Resources <ul style="list-style-type: none"> Women have increased access to and ownership to assets, credit and income Agency <ul style="list-style-type: none"> Reduced percentage of women agreeing with certain reasons that justify violence against women and girls Increased participation in decision making by women at the household or community level Women participate more in their community Achievement <ul style="list-style-type: none"> Communities have a more positive attitude towards women/marginalised groups Women have improved attitudes, self-image and confidence Improved quality of relationships between women and their household and community members

	<p>Achievement</p> <ul style="list-style-type: none"> • Effective prevention strategies supported to end violence against women and girls • Decreased violence/discrimination at the household level • Women's empowerment/equality index 	
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> • Include training components (e.g., financial literacy and business skills) for participants and group leaders in self-help groups and village savings and loan associations programmes to enhance confidence in leadership by participants. • To enhance participation, encourage existing groups to use traditional knowledge sharing platforms such as listening in meetings by prospective participants and motivational workshops. • In programme design, ensure that it is mandatory for groups to establish clear rules of participation and communication by groups to promote adherence and willingness to participate. 	
	<p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> • Group compliance obligations and domination by elites may foster intergroup conflict, whereas restrictive legislative frameworks may stifle programme expansion, which could hinder programme impact. • Women may not receive support from their partners due to pervasive gender norms and roles. • Regularity of meetings ensures reliable participation and group success. • Meeting spaces should be in centralised locations to reduce transport burden on participants. 	
	<p>Good practice example:</p> <ul style="list-style-type: none"> • During the Saving for Change (SfC) programme in Mali, some participants accepted the programme because replicating agents were considered legitimate and this legitimacy was derived from a formal training they underwent and the certificate they received. Trainers gained status and recognition by both clients and the broader community owing to this training. • In the same programme, group support was also seen to promote participation by trust building in the system of self-help groups and village savings and loan associations first. For instance, there were groups that refused to join the association due to lock-in costs. To garner interest and encourage participation, an agent brought women from different groups to visit and observe association meetings and report back to their respective groups. 	

Intervention	Intervention works on	Intervention does not work on
Technical and vocational education and training	<p>Resources</p> <ul style="list-style-type: none"> • Women have increased access to and ownership of assets, credit and income • Improved capacity of women entrepreneurs <p>Agency</p> <ul style="list-style-type: none"> • Reduced support for or occurrence of child and forced marriage <p>Achievement</p> <ul style="list-style-type: none"> • Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration • Women have improved attitudes, self-image and confidence. 	<p>Resources</p> <ul style="list-style-type: none"> • Women can access decent work (formal and informal employment) <p>Agency</p> <ul style="list-style-type: none"> • Women have more and better control over their bodies and sexual health • Women have increased freedom of movement and association • Reduced percentage of women agreeing with certain reasons that justify violence against women and girls • Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them • Increased participation in decision making by women at the household or community level <p>Achievement</p> <ul style="list-style-type: none"> • Increased community support for women's and children's human, economic and legal rights • Reduced frequency and distribution of types of violence by an intimate partner • Improved quality of relationships between women and their household and community members • Women's empowerment/equality index
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> • Design and implement a tailored and sensitive awareness-raising and outreach campaign to attract participants into the programme. • Identify and select a central and safe location that is easily accessible (e.g., via public transport, within walking distance of participants). • Provide allowances for transport or childcare to encourage participation and reduce barriers of access. • Target single sex participants (women) to allow space for open engagement and enhance feelings of safety. 	

	<ul style="list-style-type: none"> • An integration with Life, social, and livelihood skills and capacity building (LSCB) capacity building programmes are promising as assessments of stand-alone technical and vocational education and training (TVET) programmes and TVET components within LSCB programmes emphasise the interconnectedness of both types of skill sets. • A multicomponent or multi-faceted design approach to support beyond the skill building component including mentoring, peer-to-peer exchange and ‘collective-building’ formats such as women’s self-help groups and clubs. <p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> • Training curricula and materials require careful adaptation to local context and translation into local languages. • The skilfulness and approach of programme trainers is a key implementation element and driver of programme satisfaction. • Tailoring programming to context is crucial, particularly in contexts where women or girls are barred from certain occupations. <p>Good practice example:</p> <ul style="list-style-type: none"> • The Gender and Entrepreneurship Together (GET) Ahead for Women in Enterprise programme in Kenya was designed specifically with a gender lens throughout and aimed to address both the practical and strategic needs of low-income women in enterprise by strengthening their basic business and people management skills as well as supporting networking and relationship building. Women beneficiaries, in addition to the freedom to express themselves without the censure of men, valued the ‘camaraderie’, support and mentorship they received from one another. Mentoring was seen to enhance business skill transfer and development by contextualising skills in the experience of an established entrepreneur. • The Empowerment and Livelihood for Adolescents (ELA) programme in Sierra Leone referred to its multifaceted approach to simultaneously tackle multiple disadvantages young women face, related to having agency over their bodies and barriers to accumulating human capital. This approach combines the provision of life skills, vocational skills and microfinance, delivered within a social club and safe space for women. • Liberia’s Economic Empowerment of Adolescent Girls and Young Women (EPAG) programme deliberately designed a pairing, small group (3-4 girls) approach between girls with the aim of improving girls’ skill development as well as to promote the creation of social capital, such as friendships, mutual trust and support networks among the girls. The girl participants appreciated this design element and reported that they helped each other during and outside of the classroom activities.
--	--

Prevention

Table 2: Prevention intervention recommendations

Intervention	Intervention works on	Intervention does not work on
Community dialogues and reconciliations	Resources <ul style="list-style-type: none"> Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members 	Resources <ul style="list-style-type: none"> Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> Effect unknown - insufficient evidence
	Intervention design suggestions to maximise benefits <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	Contextual considerations to keep in mind <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	Good practice example: <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	

Protection

Table 3: Protection interventions recommendations

Intervention	Intervention works on	Intervention does not work on
All-women police stations	Resources <ul style="list-style-type: none"> Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> N Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> Effect unknown - insufficient evidence 	Resources <ul style="list-style-type: none"> Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> Effect unknown - insufficient evidence

	Intervention design suggestions to maximise benefits	
	<ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
	Contextual considerations to keep in mind <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
	Good practice example: <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
Intervention	Intervention works on	Intervention does not work on
Sensitisation campaigns	Resources <ul style="list-style-type: none"> • Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> • Women have more positive attitude towards taking action to claim their rights. Achievement <ul style="list-style-type: none"> • Increased awareness in communities of the issues affecting women • Women's empowerment/equality index 	Resources <ul style="list-style-type: none"> • Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> • Reduced percentage of women agreeing with certain reasons that justify violence against women and girls Achievement <ul style="list-style-type: none"> • Effective prevention strategies supported to end violence against women and girls • Women have improved attitudes, self-image and confidence • Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members • Reduced frequency of physical IPV • Reduced frequency of sexual IPV • Improved quality of relationships between women and their household and community members
	Intervention design suggestions to maximise benefits <ul style="list-style-type: none"> • Design more frequent programme activities such as training and counselling over an extended period of time. • Adopt in-person platforms and intimate sessions that largely focus on couples, and men as perpetrators. • Engage and utilise respected opinion leaders (e.g., religious leaders) to counsel, mentor and sensitise couples to IPV with focus on the tenets of healthy relationships and conflict resolution skills and knowledge. 	

	<ul style="list-style-type: none"> ● Tailor the types of community engagement to context, such as parent’s evenings, dialogue groups and exchanges. ● Draw on religious messages to support change regarding gender equality. ● As seen in other intervention groups, sensitisation campaigns can be a key “add-on” to enhance the uptake and implementation of an intervention. As a result, sensitisation may be an “easy win” to increase the opportunity of trickle-down effects across programmes. <p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> ● Gender norms, roles and culturally developed patriarchal attitudes are often the most entrenched and difficult barrier in successful sensitisation but engaging men, particularly local opinion leaders, may counter traditional gender roles. ● Issues of personal image, knowledge on rights and empowerment can be addressed gradually as they depend on the nature of close relationships and personal confidence from successful programme impact. ● Using communal and public platforms can work to reach more participants at the same time compared to intimate settings. <p>Good practice example:</p> <ul style="list-style-type: none"> ● Training men to understand the value of women’s roles better was an important exercise for Men’s Action for Stopping Violence Against Women (MASVAW) participants in India. Particularly, activities that allowed men to imagine “the average day in a woman’s life” were instrumental in provoking reflection around gender roles and norms. Opinion leaders incorporated in India’s MASVAW programme expressed excitement to engage actively with other participants and supported programme staff in identifying and embedding knowledge on gender inequalities in their communities. Their involvement encouraged them to prioritise intimate partner violence prevention and wider engagement on gender issues with the community. ● In Rwanda, the Indashyikirwa programme’s personalised settings were more conducive to engagement and listening through participatory methods of community engagement. These included community forums, to encourage openness, particularly regarding violence against women and gendered issues. Participatory training and community-based activism in the Indashyikirwa programme encouraged a culture of community engagement to activate learning and third-party intervention. In this way, participants were able to protect those that suffered from intimate partner violence. ● A religious opinion leader in Rwanda Indashyikirwa’s programme noted the importance of drawing on religious messages and principles to support change regarding gender equality as it allowed participants to understand and internalise that the genders were all equal.
--	--

Multi-pillar

Table 4: Multi-pillar interventions recommendations

Intervention	Intervention works on	Intervention does not work on
Discussion groups	Resources <ul style="list-style-type: none"> Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members 	Resources <ul style="list-style-type: none"> Women have increased access to and ownership to assets, credit and income Agency <ul style="list-style-type: none"> Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> Effect unknown - insufficient evidence
	Intervention design suggestions to maximise benefits <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	Contextual considerations to keep in mind <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	Good practice example: <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
Intervention	Intervention works on	Intervention does not work on
Life, social, and livelihood skills and capacity building	Resources <ul style="list-style-type: none"> Increased capacity of women to understand and use financial, banking, and business services effectively Agency <ul style="list-style-type: none"> Women have increased freedom of movement and association. Women are more aware of their rights and the roles and responsibilities of duty bearers. Reduced percentage of women agreeing with certain reasons that justify violence against women and girls 	Resources <ul style="list-style-type: none"> Women have increased access to and ownership to assets, credit and income Agency <ul style="list-style-type: none"> Reduced support for or occurrence of child and forced marriage Women participate more in their community Achievement <ul style="list-style-type: none"> Communities have a more positive attitude towards women/marginalised groups

	<ul style="list-style-type: none"> • Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them. • Increased participation in decision making by women at the household or community level <p>Achievement</p> <ul style="list-style-type: none"> • No significant effect identified in the list of studies 	<ul style="list-style-type: none"> • Women have improved attitudes, self-image and confidence • Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members • Decreased violence/discrimination at the household level • Safer and more secure household, communities and areas/territories for women, girls, men, and boys • Reduced frequency of physical IPV • Reduced frequency of sexual IPV • Improved quality of relationships between women and their household and community members • Women's empowerment/equality index
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> • Design and implementation of a sensitive and targeted awareness-raising and outreach campaign is key to attract participants into the programmes. • Provide participants with a central and safe location that is easily accessible and/ or provides childcare facilities promotes acceptability and programme participation. • Life, social, and livelihood skills and capacity building design benefits from the inclusion of and coupling with TVET programmes due to the interconnectedness of both types of skill sets, expressed by a demand and need for both types of skills to enhance participants life capabilities and opportunities. • Incorporate elements that address contexts and norms issues directly such as women's peer-to-peer exchange including mentoring elements, women's self-help groups and clubs • Trainer skills and approaches are an important consideration and factor for participant satisfaction • Continuous sourcing of participant feedback and consultation to allow rapid programme iteration (e.g., programme locations) 	

	Contextual considerations to keep in mind <ul style="list-style-type: none"> • Social norms and restrictive contexts determine the space in which programmes can operate and emphasise the need to strike balance between the depth of embedding programmes in these contexts to operate and reach women and the adjustment of programme empowerment outcomes to avoid strong conflict with existing norms and practices. • Failure to adapt the curriculum to each context may affect uptake and relevance. • Lack of men's engagement may be a key barrier to empowerment outcomes that may impede skill development to translate into social change. 	
	Good practice example: <ul style="list-style-type: none"> • The Creating Opportunities through Mentoring, Parental Involvement and Safe Spaces (COMPASS) programme faced opposition from parents and community leaders in Congo, Ethiopia and Pakistan with participants and stakeholders questioning the appropriateness of the programme activities. Only through extensive sensitive, targeted awareness-raising and outreach campaigns that included extensive discussions, house-to-house visits and negotiations with community leaders, local authorities, and parents, was consensus achieved and implementation proceeded. Through trust and legitimacy, the programme was able to collaborate with leaders of 'secret' societies associated with female genital mutilation enhancing its reach to participants it would have not been able to access. • In Afghanistan's participants What Works to Prevent Violence Programme participants had access to TVET-linked programme components such as cash crop farming, animal husbandry, and embroidery within a wider empowerment programme and women peers for mentoring including fellow programme participants, trainers and formal mentors provided by the programmes. Access to start-up capital and learning from women in similar contexts based on shared experiences and realities was reported as a key benefit of the arrangement. Women formed social safety nets for each other by staying in touch post the intervention and supporting each other's business ventures. These arrangements resulted in enhanced social capital, a crucial factor in targeting structural barriers to empowerment. 	
Intervention	Intervention works on	Intervention does not work on
Safe spaces	Resources <ul style="list-style-type: none"> • Women have increased access to and ownership of assets, credit and income 	Resources <ul style="list-style-type: none"> • Increased capacity of women to understand and use financial, banking, and business services effectively

	<p>Agency</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Achievement</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence 	<p>Agency</p> <ul style="list-style-type: none"> • Women have more and better control over their bodies and sexual health • Women have increased freedom of movement and association • Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them • Reduced support for or occurrence of child and forced marriage • Increased participation in decision making by women at the household or community level <p>Achievement</p> <ul style="list-style-type: none"> • Women have improved attitudes, self-image and confidence • Improved quality of relationships between women and their household and community members
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> • Target adolescent girls before they are culturally considered ready to marry to combat trends of forced marriage or teenage pregnancy. • Promote mentorship to encourage openness, collaboration, interaction and learning issues beyond what is learnt through formal education. • Secure safe, convenient meeting locations and timing of the meetings. • Provide incentives and rewards, which can encourage wide participation of the targeted population in programmes, delivered timely and without too stringent conditions. • Incorporate appropriate awareness and communication strategies within households, especially when recruiting adolescent girls to tackle negative perceptions about programme intentions and debates around children's rights. 	

	<p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> • Changes in long-standing social and cultural norms and attitudes towards women or girls may need time to reflect. • Mentors' motivation level and capacity development in advocacy, outreach, communication and networking skills are key to winning the confidence of parents to encourage girls to participate. • Baseline analyses and participatory research before implementation support appropriate and effective safe spaces. <p>Good practice example:</p> <ul style="list-style-type: none"> • The Ishraq programme in Egypt targeted girls 13-15 years of age, which was a factor that may contribute to programme impacts where girls 15 and older were considered ready for marriage. Awareness building among participants as well as their household members and the wider community at an early stage of the intervention played a critical role as many of such outcomes were compounded by social, cultural norms and traditions. • The Adolescent Girls Empowerment Programme (AGEP) programme in Zambia highlighted the importance and centrality of mentorship as it provided a platform for girls to open up to matters that they may not be able to share with their families and this built up their self-esteem.
--	--

Synthesising across the summary tables, we highlight the following themes:

- The interventions in the participation pillar show positive results across the Resources-Agency-Achievement framework, identifying promising interventions for policy-makers interested in impacting gender equality and women's empowerment.
- The low number of studied interventions under the protection or the prevention pillar precludes an assessment of these interventions, warranting additional research in this area.
- Findings from the multi-pillar interventions indicate promising results on the resources-focused outcomes, but there is a relatively low number of interventions. More research on the impact of multi-pillar interventions on agency and achievement is needed.
- Overall, gender-specific and transformative interventions are effective at addressing the three dimensions of the gender equality and women's empowerment framework.

Table 5: Overview of interventions effectiveness

Empowerment Dimensions	Resources			Agency			Achievement		
Outcome categories	Access to justice and legal services	Economic and livelihood related resources	Access to employment	Individual agency	Community level agency	Institutions supporting agency	Improved systems	Norms and behaviour change	Index
Participation		5	3	3	1		4	2	1
Asset transfer									
Cash transfer									
Community-based services									
Inclusive community-driven development									
Institutional provision of loans and savings									
Quotas									
Self-help groups and Village savings associations									
Technical and vocational education and training									
Prevention							1		
Community dialogues and reconciliations									
Protection				1			1		1
All women police stations									
Sensitisation campaigns									
Multi-pillar (participation and protection)		2		1	1		1		
Discussion groups									
Life, social, and livelihood skills and capacity building									
Safe spaces									
General	0	7	3	5	2	0	4	5	2

Legend

	Aggregation of effect at the general level and number of interventions observing a significantly positive effect on an outcome from this category at intervention level
	Aggregation of effect at the UNSCR 1325 level and number of interventions observing a significantly positive effect on an outcome from this category at intervention level
	Significantly positive effect observed on an outcome from this category at intervention level

Content

Acknowledgements	i
Plain language summary	ii
Executive Summary	iv
List of figures and tables	xxxv
List of abbreviations	xlili
1. Background	1
1.1 The problem, condition or issue	1
1.2 The intervention	6
1.3 How the intervention might work	12
1.4 Why it is important to do this review	13
2. Objectives	18
3. Methods	19
3.1 Criteria for considering studies for this review	19
3.2 Search methods for identification of studies	30
3.3 Data collection and analysis	34
4. Results	47
4.1 Description of studies: search results and characteristics of the evidence base ...	47
4.2 Effects of interventions	69
4.3 Effects of interventions under the UNSCR participation pillar	69
4.4 Effects of interventions under the UNSCR protection pillar	248
4.5 Effects of interventions under the UNSCR prevention pillar	280
4.6 Effects of interventions under multiple UNSCR pillars: participation and protection	290
5. Summary of findings and conclusions	361
5.1 Summary of findings	361
5.2 Overall completeness of the evidence	374
5.3 Quality of the evidence	375
5.4 Agreement and disagreement with other studies	376
5.5 Discussion and conclusions	378
5.6 Limitations, potential biases and deviations from the protocol	400
5.7 Deviations from the protocol	402
6. Acknowledgements	404
6.1 Contributions of authors	404
6.2 Declarations of interest	405
6.3 Preliminary timeframe	405
6.4 Plans for updating this review	405
Online appendixes	406
References	408

List of figures and tables

Figure 1: Building peaceful and inclusive societies	3
Figure 2: Our systematic review's theory of change	13
Figure 3: Overview of the review process	19
Figure 4: Map of includable countries for the systematic review	25
Figure 5: PRISMA flow diagram of systematic search and screening results	48
Figure 6: Annual (panel A, top) and cumulative (panel B, bottom) number of included publications over the review period from 2000 to 202	50
Figure 7: Types of included publications	50
Figure 8: Map of countries of included studies	51
Figure 9: Studies per country	51
Figure 10: Study FSI scores	52
Figure 11: Study GII scores	52
Figure 12: Distribution of intervention types	57
Figure 13: Intervention length (months)	57
Figure 14: Reporting on uptake and adherence per intervention type	60
Figure 15: Distribution of qualitative studies, by country	65
Figure 16: RoB Assessment of QED included studies	66
Figure 17: RoB assessment of RCT included studies	67
Figure 18: Summary of overall critical appraisal ratings across studies included in the qualitative synthesis	67
Figure 19: Critical appraisal category ratings across studies included in the qualitative synthesis	68
Figure 20: Theory of change of included studies	69
Figure 21: Theory of change of included studies under the participation pillar	71
Figure 22: Asset transfers for gender equality, women's empowerment and peace: theory of change	73
Figure 23: Asset transfer QED RoB assessment	78
Figure 24: Asset transfer RCT RoB assessment	78
Figure 25: AssetAB2: Forest plot showing the observed outcomes and the estimate of the random-effects model	80
Figure 26: AssetAB2b: Funnel plot	80
Figure 27: AssetAC2: Forest plot showing the observed outcomes and the estimate of the random-effects model	81
Figure 28: AssetBB1. Forest plot showing the observed outcomes and the estimate of the random-effects model	83
Figure 29: AssetCB3. Forest plot showing the observed outcomes and the estimate of the random-effects model	84
Figure 30: AssetCB7: Forest plot showing the observed outcomes and the estimate of the random-effects model	85
Figure 31: AssetCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model	86
Figure 32: AssetCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model	87
Figure 33: Cash transfers for gender equality, women's empowerment and peace: theory of change	97
Figure 34: Cash transfer studies QED RoB assessment	103

Figure 35: Cash transfer studies RCT RoB Assessment.....	104
Figure 36: CTAB2a: Forest plot showing the observed outcomes and the estimate of the random-effects model	105
Figure 37: CTAB2b: Funnel plot.....	106
Figure 38: CTAB3: Forest plot showing the observed outcomes and the estimate of the random-effects model	107
Figure 39: CTAC2: Forest plot showing the observed outcomes and the estimate of the random-effects model	108
Figure 40: CTBA2: Forest plot showing the observed outcomes and the estimate of the random-effects model	109
Figure 41: CTBA3: Forest plot showing the observed outcomes and the estimate of the random-effects model	110
Figure 42: CTBA6: Forest plot showing the observed outcomes and the estimate of the random-effects model	111
Figure 43: CTBA7: Forest plot showing the observed outcomes and the estimate of the random-effects model	112
Figure 44: CTBA8: Forest plot showing the observed outcomes and the estimate of the random-effects model	112
Figure 45: CTBB1: Forest plot showing the observed outcomes and the estimate of the random-effects model	113
Figure 46: CTBB2: Forest plot showing the observed outcomes and the estimate of the random-effects model	114
Figure 47: CTCB4: Forest plot showing the observed outcomes and the estimate of the random-effects model	115
Figure 48: CTCB6: Forest plot showing the observed outcomes and the estimate of the random-effects model	117
Figure 49: CTCB7: Forest plot showing the observed outcomes and the estimate of the random-effects model	117
Figure 50: CTCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model	118
Figure 51: CTCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model	119
Figure 52: Community-based services for gender equality, women's empowerment and peace: theory of change	131
Figure 53: RCT Community based services RoB assessment	133
Figure 54: Inclusive community-driven development for gender equality, women's empowerment and peace: theory of change.....	137
Figure 55: Inclusive CDD studies RCT RoB assessment	141
Figure 56: CDDBA3: Forest plot showing the observed outcomes and the estimate of the random-effects model	142
Figure 57: CDDBA5: Forest plot showing the observed outcomes and the estimate of the random-effects model	143
Figure 58: CDDBA6: Forest plot showing the observed outcomes and the estimate of the random-effects model	144
Figure 59: CDDCA2: Forest plot showing the observed outcomes and the estimate of the random-effects model	145
Figure 60: CDDCA8: Forest plot showing the observed outcomes and the estimate of the random-effects model	147

Figure 61: CDDCB3: Forest plot showing the observed outcomes and the estimate of the random-effects model	148
Figure 62: CDDCB6: Forest plot showing the observed outcomes and the estimate of the random-effects model	149
Figure 63: Institutional provision of loans and savings for gender equality, women's empowerment and peace: theory of change.....	162
Figure 64: Institutional loans and savings studies QED RoB assessment.....	165
Figure 65: Institutional loans and savings RCT RoB assessment	166
Figure 66: LSAB2: Forest plot showing the observed outcomes and the estimate of the random-effects model	167
Figure 67: LSAC3: Forest plot showing the observed outcomes and the estimate of the random-effects model	168
Figure 68: LSBB1: Forest plot showing the observed outcomes and the estimate of the random-effects model	169
Figure 69: Quotas for gender equality, women's empowerment and peace: theory of change	175
Figure 70: QED quotas RoB assessment	178
Figure 71: RCT quotas RoB assessment.....	179
Figure 72: QuoAttitudes. Forest plot showing the observed outcomes and the estimate of the random-effects model	180
Figure 73: Self-help groups and VSLAs for gender equality, women's empowerment and peace: theory of change	184
Figure 74: SHGs and VSLA studies QED RoB assessment.....	189
Figure 75: SHGs and VSLA studies RCT RoB assessment	189
Figure 76: SHGAB1: Forest plot showing the observed outcomes and the estimate of the random-effects model	190
Figure 77: SHGAB2: Forest plot showing the observed outcomes and the estimate of the random-effects model	191
Figure 78: SHGAC2: Forest plot showing the observed outcomes and the estimate of the random-effects model	192
Figure 79: SHGBA3: Forest plot showing the observed outcomes and the estimate of the random-effects model	193
Figure 80: SHGBA4: Forest plot showing the observed outcomes and the estimate of the random-effects model	194
Figure 81: SHGBA5: Forest plot showing the observed outcomes and the estimate of the random-effects model	195
Figure 82: SHGBA6: Forest plot showing the observed outcomes and the estimate of the random-effects model	196
Figure 83: SHGBA7: Forest plot showing the observed outcomes and the estimate of the random-effects model	197
Figure 84: SHGBB1: Forest plot showing the observed outcomes and the estimate of the random-effects model	198
Figure 85: SHGBB2: Forest plot showing the observed outcomes and the estimate of the random-effects model	199
Figure 86: SHGCA2: Forest plot showing the observed outcomes and the estimate of the random-effects model	200
Figure 87: SHGCA6: Forest plot showing the observed outcomes and the estimate of the random-effects model	201

Figure 88: SHGCB2: Forest plot showing the observed outcomes and the estimate of the random-effects model	202
Figure 89: SHGCB3: Forest plot showing the observed outcomes and the estimate of the random-effects model	202
Figure 90: SHGCB5: Forest plot showing the observed outcomes and the estimate of the random-effects model	203
Figure 91: SHGCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model	204
Figure 92: SHGCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model	205
Figure 93: TVET for gender equality, women's empowerment and peace: theory of change	218
Figure 94: TVET studies QED RoB assessment	222
Figure 95: TVET studies RCT RoB assessment.....	223
Figure 96: TVETAB2: Forest plot showing the observed outcomes and the estimate of the random-effects model	224
Figure 97: TVETAB2b: Funnel plot	225
Figure 98: TVETAC2: Forest plot showing the observed outcomes and the estimate of the random-effects model	226
Figure 99: TVETAC3: Forest plot showing the observed outcomes and the estimate of the random-effects model	227
Figure 100: TVETBA2: Forest plot showing the observed outcomes and the estimate of the random-effects model	228
Figure 101: TVETBA3: Forest plot showing the observed outcomes and the estimate of the random-effects model	229
Figure 102: TVETBA6: Forest plot showing the observed outcomes and the estimate of the random-effects model	230
Figure 103: TVETBA7: Forest plot showing the observed outcomes and the estimate of the random-effects model	231
Figure 104: TVETBA8: Forest plot showing the observed outcomes and the estimate of the random-effects model	232
Figure 105: TVETBB1: Forest plot showing the observed outcomes and the estimate of the random-effects model	232
Figure 106: TVETCA2. Effect of TVET on women having improved attitudes, self-image and confidence.....	234
Figure 107: TVETCB3: Forest plot showing the observed outcomes and the estimate of the random-effects model	234
Figure 108: TVETCB7: Forest plot showing the observed outcomes and the estimate of the random-effects model	236
Figure 109: TVETCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model	236
Figure 110: TVETCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model	237
Figure 111: Theory of change for included studies under the protection pillar	249
Figure 112: All Women Police Stations for gender equality, women's empowerment and peace: theory of change	251
Figure 113: QED AWPS RoB assessment	253

Figure 114: Sensitisation campaigns for gender equality, women's empowerment and peace: theory of change	257
Figure 115: Sensitisation campaign studies QED RoB assessment.....	260
Figure 116: Sensitisation campaigns RCT RoB assessment.....	261
Figure 117: SCBA5: Forest plot showing the observed outcomes and the estimate of the random-effects model	263
Figure 118: SCBA6: Forest plot showing the observed outcomes and the estimate of the random-effects model	264
Figure 119: SCCA6: Forest plot showing the observed outcomes and the estimate of the random-effects model	265
Figure 120: SCCB1: Forest plot showing the observed outcomes and the estimate of the random-effects model	265
Figure 121: SCCB3: Forest plot showing the observed outcomes and the estimate of the random-effects model	266
Figure 122: SCCB4: Forest plot showing the observed outcomes and the estimate of the random-effects model	267
Figure 123: SCCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model	268
Figure 124: SCCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model	269
Figure 125: Theory of change of the included studies under the prevention pillar.....	281
Figure 126: Community dialogues and reconciliations for gender equality, women's empowerment and peace: theory of chang.....	283
Figure 127: CDCB4. Forest plot showing the observed outcomes and the estimate of the random-effects model	287
Figure 128: Theory of change of the included studies under multiple pillars	291
Figure 129: Discussion groups for gender equality, women's empowerment and peace: theory of change	293
Figure 130: Discussion Groups studies RCT RoB assessment.....	296
Figure 131: DGAB2. Forest plot showing the observed outcomes and the estimate of the random-effects model	297
Figure 132: DGCB4. Forest plot showing the observed outcomes and the estimate of the random-effects model	299
Figure 133: Safe spaces for gender equality, women's empowerment and peace: theory of change	302
Figure 134: Safe spaces studies QED RoB assessment.....	306
Figure 135: Safe spaces studies QED RoB assessment.....	306
Figure 136: SSAB1: Forest plot showing the observed outcomes and the estimate of the random-effects model	307
Figure 137: SSAB2: Forest plot showing the observed outcomes and the estimate of the random-effects model	308
Figure 138: SSBA2: Forest plot showing the observed outcomes and the estimate of the random-effects model	309
Figure 139: SSBA3: Forest plot showing the observed outcomes and the estimate of the random-effects model	309
Figure 140: SSBA7: Forest plot showing the observed outcomes and the estimate of the random-effects model	310

Figure 141: SSBA8: Forest plot showing the observed outcomes and the estimate of the random-effects model	311
Figure 142: SSBB1: Forest plot showing the observed outcomes and the estimate of the random-effects model	312
Figure 143: SSCB3: Forest plot showing the observed outcomes and the estimate of the random-effects model	313
Figure 144: SSCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model	315
Figure 145: Life, social and livelihood skills and capacity building for gender equality, women's empowerment and peace: theory of change	325
Figure 146: Life, social and livelihood skills and capacity building studies QED RoB assessment.....	331
Figure 147: Life, social and livelihood skills and capacity building studies RCT RoB assessment.....	332
Figure 148: CBAB1: Forest plot showing the observed outcomes and the estimate of the random-effects model	333
Figure 149: CBAB2: Forest plot showing the observed outcomes and the estimate of the random-effects model	334
Figure 150: CBBA3: Forest plot showing the observed outcomes and the estimate of the random-effects model	335
Figure 151: CBBA4: Forest plot showing the observed outcomes and the estimate of the random-effects model	336
Figure 152: CBBA6: Forest plot showing the observed outcomes and the estimate of the random-effects model	337
Figure 153: CBBA7: Forest plot showing the observed outcomes and the estimate of the random-effects model	338
Figure 154: CBBA8: Forest plot showing the observed outcomes and the estimate of the random-effects model	339
Figure 155: CBBB1: Forest plot showing the observed outcomes and the estimate of the random-effects model	340
Figure 156: CBBB2: Forest plot showing the observed outcomes and the estimate of the random-effects model	341
Figure 157: CBCB2: Forest plot showing the observed outcomes and the estimate of the random-effects model	342
Figure 158: CBCB3: Forest plot showing the observed outcomes and the estimate of the random-effects model	343
Figure 159: CBCB4: Forest plot showing the observed outcomes and the estimate of the random-effects model	344
Figure 160: CBCB5: Forest plot showing the observed outcomes and the estimate of the random-effects model	344
Figure 161: CBCB6: Forest plot showing the observed outcomes and the estimate of the random-effects model	345
Figure 162a: CBCB7pv: Forest plot showing the observed outcomes and the estimate of the random-effects model	346
Figure 162b: CBCB7sv: Forest plot showing the observed outcomes and the estimate of the random-effects mod.....	347
Figure 163: CBCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model	348

Figure 164: CBCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model	349
Figure 165: Systematic Review Theory of Change	361

Table 1: Participation intervention recommendations	xv
Table 2: Prevention intervention recommendations	xxv
Table 3: Protection interventions recommendations	xxv
Table 4: Multi-pillar interventions recommendations	xxviii
Table 5: Overview of interventions effectiveness	xxxiii
Table 6: Conceptual scope of the systematic review	7
Table 7: Participation pillar, included studies	8
Table 8: Prevention pillar, included studies	9
Table 9: Protection pillar, included studies	10
Table 10: Recovery and relief pillar, included studies	12
Table 11: List of included countries for the systematic review	23
Table 12: List of included interventions for the systematic review	27
Table 13: Summary of included studies per intervention type	53
Table 14: List of outcomes and number of studies covering them	61
Table 15: List of outcomes and number of studies covering them	64
Table 16: Asset transfer activities: features of included studies	75
Table 17: Asset transfer summary of secondary and immediate outcomes	77
Table 18: GRADE summary of findings and certainty of evidence on asset transfers....	93
Table 19: Features of included cash transfer studies	98
Table 20: Summary of secondary and immediate outcomes for cash transfers	102
Table 21: GRADE summary of findings and certainty of evidence on cash transfers...	126
Table 22: Community-based services features of included studies	132
Table 23: Summary of secondary and immediate outcomes for community-based services.....	133
Table 24: GRADE summary of findings and certainty of evidence on community-based services.....	135
Table 25: Inclusive community-driven development activities design features of included studies.....	139
Table 26: Inclusive community-driven development summary of secondary and immediate outcomes.....	140
Table 27: GRADE summary of findings and certainty of evidence on inclusive community driven development.....	157
Table 28: Institutional ision of loans and savings activities features of included studies	163
Table 29: Institutional provision of loans: secondary and immediate outcomes	164
Table 30: GRADE summary of findings and certainty of evidence on institutional provision of loans and savings.....	172
Table 31: Quota studies design features	176
Table 32: Summary of secondary and immediate outcomes for quota studies	177
Table 33: GRADE summary of findings and certainty of evidence on quotas	182
Table 34: Self-help groups and VSLAs design features of included studies.....	186
Table 35: Self-help groups and VSLAs summary of secondary and immediate outcomes	188

Table 36: GRADE summary of findings and certainty of evidence on SHGs and VSLAs	213
Table 37: Activities and design features of included TVET studies	219
Table 38: TVET summary of secondary and immediate outcomes	221
Table 39: GRADE summary of findings and certainty of evidence on TVET	243
Table 40: AWPS features of included studies	252
Table 41: Summary of Secondary and Immediate outcome	253
Table 42: GRADE summary of findings and certainty of evidence on all women police stations.....	255
Table 43: Activities and design features of included sensitisation studies	258
Table 44: Summary of secondary and immediate outcomes	260
Table 45: GRADE summary of findings and certainty of evidence on sensitisation campaigns.....	277
Table 46: Community dialogues and reconciliation activities design features	285
Table 47: Summary of secondary and immediate outcomes for community dialogues and reconciliations	286
Table 48: GRADE summary of findings and certainty of evidence on community dialogues and reconciliations	289
Table 49: Discussion group activities design features of included studies	294
Table 50: Summary of secondary and immediate outcomes for discussion groups	295
Table 51: GRADE summary of findings and certainty of evidence on discussion groups	300
Table 52: Safe spaces activities and design features of included studies	303
Table 53: Summary of secondary and immediate outcomes	305
Table 54: GRADE summary of findings and certainty of evidence on safe spaces	320
Table 55: Life skills and capacity building activities design features of included studies	327
Table 56: Summary of secondary and immediate outcomes for LSCB and capacity building.....	330
Table 57: GRADE summary of findings and certainty of evidence on life, social and livelihood skills and capacity building.....	357
Table 58: Quantitative findings.....	364
Table 59: Overview of interventions effectiveness.....	380
Table 60: Participation interventions implications	383
Table 61: Prevention interventions implications.....	392
Table 62: Protection interventions implications.....	393
Table 63: Protection interventions implications.....	395
Table 64: Adaptation of our list of interventions	402

List of abbreviations

3ie	International Initiative for Impact Evaluation
AAWAZ	Voice and Accountability Programme
AFD	French Agency for Development
AG	Advisory Group
AGEP	Adolescent Girls Empowerment Programme
AGI	Adolescent Girls Initiative
AU	African Union
AWPS	All-Women Police Stations
BCC	Behaviour Change Communication
BISP	Benazir Income Support Programme
BMZ	German Ministry for Economic Cooperation and Development
BRAC	Bangladesh Rural Advancement Committee
BWC	Business Women Connect
CAs	Community Activists
CB	Capacity Building
CBT	Cognitive Behavioural Therapy
CCT	Conditional Cash Transfer
CDC	Community Development Councils
CDD	Community Driven Development
CEDAW	Convention on the Elimination of Discrimination Against Women
CETA	Common Elements Treatment Approach
CFU	Conservation Farming Unit
COMPASS	Creating Opportunities through Mentoring, Parental Involvement and Safe Spaces
CPSL	Centre for Promoting Sustainable Livelihoods
CSO	Civil Society Organisation
DDR	Disarmament, Demobilisation and Reintegration
DFID	Department for International Development
DRC	Democratic Republic of Congo
EDs	Electoral Divisions
EGM	Evidence Gap Map
E-HFP	Enhanced-Homestead Food Production
ELA	Empowerment and Livelihood for Adolescents
EMAP	Engaging Men through Accountable Practices
EPAG	Economic Empowerment of Adolescent Girls and Young Women
EU	European Union
FCAs	Farmer Community Associations
FCAS	Fragile and Conflict-Affected States and Situations
FCDO	Foreign Commonwealth and Development Office
FCFA	Central African CFA Franc
FSVGD	Food-Security Vulnerable Group Development
FT	Fambul Tok
FTS	Full Text Screening
GALS	Gender Action Learning System
GCRF	Global Challenge Research Fund

GE	Girl Empower
GHA	Gendered Household Approach
GLS	Grey Literature Search
GPI	Global Peace Index
GSR	Gender Systematic Review
HKI	Helen Keller International
IDP	Internally Displaced People
IFS	Integrated Food Security
IGVGD	Income-Generating Vulnerable Group Development
ILO	International Labour Organisation
INGO	International Non-Governmental Organisation
IPV	Intimate Partner Violence
IRC	International Rescue Committee
JLOS	Justice and Rule of Law Sector
KES	Kenyan Shilling
L&MICs	Low- and Middle-Income Countries
MASVAW	Men's Action to Stop Violence Against Women
MDGs	Millennium Development Goals
MENA	Middle East and North Africa
MoFA	Ministry of Foreign Affairs
NAP	National Action Plan
NGO	Non-Governmental Organisation
NSP	National Solidarity Programme
OECD	Organisation of Economic Co-operation and Development
PFP	Pigs for Peace
PKR	Pakistani Rupee
PSM	Propensity Score Matching
PSNP	Productive Safety Net Programme
QED	Quasi-Experimental Design
RAIN	Realigning Agriculture for Improved Nutrition
RCT	Randomised Controlled Trial
RDD	Regression Discontinuity Design
REDD+	Reducing Emissions from Deforestation and Degradation
RISHTA	Research and Intervention in Sexual Health: Theory to Action
RoB	Risk of Bias
RSS	Regional Stabilisation Strategy
SATYA	Social Awakening Through Youth Action
SDGs	Sustainable Development Goals
SEA	Sexual Abuse and Exploitation
SEWA	Self-Employed Women's Association
SfC	Saving for Change
SGBV	Sexual and Gender-Based Violence
SGSY	Swarnajayanti Gram Swarozgar Yojana
SHGs	Self-help Groups
SR	Systematic Review
SRHR	Sexual and Reproductive Health Rights

TAS	Title and Abstract Screening
TEVAW	Together to End Violence
TMRI	Transfer Modality Research Initiative
ToC	Theory of Change
TT	Tchova Tchova
TTHV	Tchova Tchova Histórias de Vida: Diálogos Comunitários
TUP	Targeting the Ultra-Poor Programme
TVET	Technical and Vocational Education and Training
UBL	Unite for a Better Life
UCT	Unconditional Cash Transfer
UN	United Nations
UNSC(R)	United Nations Security Council (Resolution)
USD	United States Dollar
VDC	Village Development Committee
VFLs	Village Farm Leaders
VGD	Vulnerable Group Development Programme
VMFs	Village Model Farms
VSLAs	Village Savings and Loan Associations
WfWI	Women for Women International
WINGS	Women's Income-Generating Support
WPS(A)	Women Peace and Security (Agenda)
ZCTP	Zomba Cash Transfer Programme

1. Background

1.1 The problem, condition or issue

Despite gains made over the last decades towards equity, inclusion, and empowerment, significant gaps remain for women accessing resources, earning livelihoods, achieving legislative and political representation, and participating in important decision-making processes (Cornwall 2016; United Nations 2010). A growing body of literature indicates that these challenges are exacerbated in contexts of conflict and fragility (Bouta et al. 2004; Buvinic et al. 2013; Caprioli 2000).

This is particularly alarming when presented alongside the reality that the prevalence of fragility and conflict-affected states is on the rise. The World Bank reports that over the last ten years, the number of fragile and conflict-affected states and situations (FCAS) has increased from 36 to 39 countries or states, with 17 classified as a 'situation of medium- to high-intensity conflict.' (World Bank n.d.) The Institute for Economics and Peace reports a similar deterioration of peacefulness, finding a 0.34 percentage point drop last year in their Global Peace Index (GPI), which measures societal safety and security, ongoing domestic and international conflict, and degree of militarisation (Institute for Economics and Peace 2020). According to the Organisation for Economic Co-operation and Development (OECD), a staggering 23 per cent of the world's population is living in a fragile context, including 76.5 per cent of all those living in extreme poverty (OECD 2020).

1.1.1 Definition of key terms

Fragile contexts

Defining fragility is a complex and sometimes contentious challenge due to the conceptual ambiguities that characterise it (Faust et al. 2013). We adopt the same definition for situations of fragility as used in the International Initiative for Impact Evaluation's (3ie) *Building peaceful societies: An evidence gap map* (Sonnenfeld et al. 2020). Situations of fragility can be understood as "...the combination of exposure to risk and insufficient coping capacity of the state, system and/or communities to manage, absorb or mitigate those risks. Fragility can lead to negative outcomes, including violence, the breakdown of institutions, displacement, humanitarian crises or other emergencies" (OECD 2016a). The focus on situations of fragility recognises that exposure to risks and vulnerabilities is not constant, neither over time nor within a state. For example, within a state that is not considered fragile, there may be communities that have the characteristics of fragility. Note that we operationalise a metrics-based definition of fragility for our inclusion criteria which can be found in Methodology of Selection of Population.

Gender equality and women's empowerment

Our systematic review (SR) will use Naila Kabeer's definition of women's empowerment: "...a process by which women who have been denied the ability to make strategic life choices acquire such an ability" (Kabeer 1999). This ability to exercise choice relies on three interrelated dimensions:

- **Resources:** material, human and social resources which serve to enhance the ability to exercise choice;
- **Agency:** ability to define one's goals and act upon them and operationalised decision-making; and

- **Achievement:** ways of being and doing which can be realised by different individuals.

The different dimensions emphasise that *empowerment* encompasses different categories of daily life and as such, empowerment-focused interventions for women and girls may take the form of many different programmes, as further specified in the *interventions* section of this protocol (Lwamba et al., 2021). Overall, empowerment contributes to the overall equality between men and women, improving one's ability to make choices and live a safe and fulfilling life (Cornwall, 2000).

Peacebuilding

The former United Nations (UN) Secretary-General Boutros Boutros-Ghali provided one of the first definitions of the concept of *peacebuilding* as “an action to identify and support structures which will tend to strengthen and solidify peace in order to avoid relapse into conflict” (Barnett et al. 2007). Like the *Building peaceful societies* evidence gap map (EGM), our SR adopts the definition of peacebuilding developed by the UN Secretary General's Policy Committee in 2007, which defines peacebuilding as “a range of measures aimed at reducing the risk of lapsing or relapsing into [violent] conflict, by strengthening national capacities for conflict management and laying the foundations for sustainable peace. It is a complex, long-term process aimed at creating the necessary conditions for positive and sustainable peace by addressing the deep-rooted structural causes of violent conflict in a comprehensive manner” (United Nations Peacebuilding Support Office 2010).

There are three types of peacebuilding processes:

- Track I describes activities that bring parties to a conflict into direct negotiation to achieve an agreement or a resolution through official discussions between high level governmental and military leaders (includes ceasefires, peace talks, treaties etc.);
- Track II describes activities of unofficial dialogue and problem-solving aimed at building relationships between civil society leaders and influential individuals that can impact Track I; and
- Track III describes activities of people-to-people interactions at the grassroots level to encourage interaction and understanding between communities (includes meetings, media exposure, political and legal advocacy etc.) (Dudouet, 2017).

Sustainable peace relies on numerous stakeholders, at all levels and beyond the lifetime of an active conflict or crisis. Our SR analyses the role of these stakeholders: private sector actors, the public, civil society organisations, civil servants and service providers, and individual households.

Human security

A broad approach to fostering peaceful societies is one that examines and understands the dangers of fragility to not only be about physical security, but human security as well. The concept of human security originates from UNOCHA's Commission on Human Security to develop a world “free from want” and “free from fear” (UNOCHA 2009). The Berghof Foundation defines human security as “...a comprehensive, people-centred and prevention-oriented concept that includes protection from threats in the area of economic, food, health, environmental, personal, community and political security” (Berghof Foundation 2019). Human security expands the concept of security beyond a

state-centric framework to focus more on the macro-level of households and individuals. For women, there are vast differences and obstacles in attaining the same level of human security as men.

Peaceful and inclusive societies

As presented in the Building Peaceful Societies EGM, peaceful and inclusive societies cover a wider spectrum than the absence of violence and the resolution of conflict. It is then important to understand the concept of peaceful and inclusive societies as the conjunction of addressed fragility, human security, positive peace but also sustainable and inclusive governance. This target of peaceful societies has been made a priority by a series of international agencies such as the Sustainable Development Goals (Keuleers 2016) and the UK Foreign and Commonwealth Office who identified four objectives for peaceful states and societies (Department for International Development 2010):

1. Address the causes and effects of conflict and fragility, and build conflict resolution mechanisms;
2. Support inclusive political settlements and processes;
3. Develop core state functions; and
4. Respond to public expectations.

Building on the approach initiated in the Building Peaceful Societies EGM, our review adopts a definition of *peaceful and inclusive societies* based on the association of the following concepts:

- Addressing the causes and drivers of fragility by building sustainable economic foundations and livelihoods
- Addressing the roots causes and drivers of conflicts, strengthening social cohesion, supporting peace processes to build sustainable peace
- Supporting human security through the prevention of violence and protection of human rights and security from any form of violence
- Supporting good governance through the development of responsive and sustainable institutions and governance practices
- Developing inclusivity by addressing the specific needs of marginalised and vulnerable groups (including women and girls)

Figure 1: Building peaceful and inclusive societies



1.1.2 The issue: women's empowerment in FCAS

Women and girls are economically disempowered through restricted access to livelihoods and resources (Sweetman and Rowlands 2016). The causes of this disenfranchisement extend from a lack of economic or social protection, to social norms and traditions, to legal and political barriers (Chen et al. 2006) (Doss 2013) (Strickland 2004). Examples of the kinds of systemic barriers that discourage women's economic autonomy are early and forced marriage, bureaucratic hurdles to accessing capital for entrepreneurship, sexual and gender-based violence, and discrimination and harassment (OECD 2016b) (Perrin et al. 2019).

In addition to economic disempowerment, women are also often excluded from having a seat at the policy or decision-making table. Women's movements are localised and often restricted to the grassroots, and when they have broader national reach, are often co-opted by NGOs or government bodies (Jong and Kimm 2017). Women and girls have insurmountable hurdles if they seek to enact change at all levels, from their communities to the national government (Goetz 2008). Furthermore, at the legislative and political levels, equal rights that may be enjoyed by women and girls in national legislation are often not codified at the local or sub-national level due to contravening customary law and social norms (UN Women 2015).

In FCAS, these kinds of disadvantages and inequalities are greatly amplified. (Speake 2013). Violence and fragility affect women and girls in many ways, some of which differ from their impact on men and boys. For example, sexual assault and exploitation of women are often used as "tools of war," and threats, such as child marriage and human trafficking of women, are often exacerbated during conflicts (USAID 2007). There are many drivers of the differences in effects of conflict on women and girls from men and boys, such as the implications of the different social roles assigned to men and women, which vary by context. Understood in this way, the effects of conflict on women do not stem from any intrinsic weakness in women, but are rather a consequence of their position in society (Pankhurst 2000).

In addition to the immediate threats to women, conflicts can have a destabilising effect on gender norms in the long term. Gender roles can change during conflicts by increasing the burden and responsibilities of women while men are away fighting. Such shifts in gender and social norms can manifest as new sources of conflict when women are expected to return to pre-conflict status (United States Institute for Peace 2012).

At the root of this issue is the exclusion of women from the processes that foster peaceful and inclusive societies. A study of 31 major peace processes between 1992 and 2011 by UN Women found that only 2.4% of chief mediators, 4% of peace signatories and 9% of negotiating delegations were women (O'Reilly 2013). Between 2000 and 2010, less than a third of peace agreements signed contained a gender reference (Hedström et al. 2015). The exclusion of women's voices and contributions to peace processes not only leave women disenfranchised, underrepresented and vulnerable, but studies have shown that limiting women's participation is associated with greater recidivism and return to conflict (Hedström et al. 2015) (Bigio and Vogelsteing 2016) (Council on Foreign Relations n.d.) (O'Driscoll 2017).

In summary, there cannot be positive peace, or the creation of social systems and foundations for durable, peaceful, and inclusive societies, without an empowered female population that participates in the restoration and/or construction of relationships (Galtung 1996).

1.1.3 A solution: women's empowerment and gender equality in FCAS

Processes that develop peaceful societies are not only opportunities to strengthen, reconstruct or promote the resilience of social cohesion but also to transform the roles and definition of social relations towards inclusivity, participation and peace (Aall and Crocker 2019). Rebuilding social cohesion and inclusive societies may be buttressed by women's empowerment and by the promotion of gender equality for peace and through peace (Björkdahl and Höglund 2013). Protecting women and girls' human rights, safety, physical and mental health and security, promotion of their socio-economic recovery and increased participation in decision-making processes and responses related to conflict or fragility are key processes that lead to overall progress towards gender equality and women's empowerment (Buvinic et al. 2013).

Empowering women across all three dimensions of gender equality and women's empowerment, according to Kabeer, often requires structural and systematic changes (Kabeer 1999). Implementing transitions based on peace, nonviolence and inclusion can be driven by repealing old laws and instituting new structures, but must also transform harmful social norms that exclude women from publicly and meaningfully participating in society (McWilliams and Kilmurray 2015). This includes ensuring that women have equal rights, are treated equally and have their voices heard and needs met. Promoting the role of women and facilitating exchanges between them help change those mindsets and promote greater gender equality, reduce conflict, and discourage extremism. In that sense, women's empowerment can both be a driver and a consequence of peace: empowerment *through* and *from* peace (Cheldelin and Mutisi 2016).

1.1.4 UNSCR 1325 as the starting point of change towards inclusive peace

When the United Nations Security Council Resolution (UNSCR) 1325 (United Nations Security Council 2000) was adopted on October 31, 2000, it represented a major milestone in acknowledging the disproportionate impact of armed conflict on women and girls in continuation of previous work started in the 20th century. The resolution, which has been used extensively as a policy tool to implement gender-sensitive conflict-related policies, affirms three major points:

- I. The recognition of the inordinate impact of violent conflicts and war on women and girls;
- II. The recognition of the crucial role that women should play in conflict prevention, resolutions, building peaceful and inclusive societies; and
- III. The necessary adoption of a gender perspective in conflict prevention, resolutions, and building peaceful and inclusive societies.

UNSCR 1325 has widely been interpreted as asserting those objectives by establishing four dimensions of women's roles in conflict: participation, prevention, protection and relief and recovery through gender equality and for gender equality.

Although UNSCR 1325 prioritised women's empowerment for peace and initiated the Women Peace and Security Agenda (WPS), some criticisms have been raised

(Senarathna, 2015). First, women are not a homogeneous group that can be represented as an entity in peace processes. Gender is both multidimensional and intersectional and thus a *one-size-fits-all* approach may not be appropriate. Second, the formulation of *women as agents of peace* has been criticised as being *instrumentalist*, focusing on what women can bring to peace and not what peace can bring to women. Finally, women are not only exposed to violence in periods of conflict but may also be highly exposed to violence in times of peace and both should be considered in the WPS. Despite these critiques, the UNSCR 1325 has been the starting point of 20 years of change toward inclusive peace (Desmidt and Davis 2019). See Online Appendix I, Table 2 for a more complete chronology of the impact of UNSCR 1325.

UNSCR 1325 has initiated a global movement towards consensus on the importance of the role of women in building peaceful and inclusive societies, but concrete implementation of the recommendations remains scattered. In 2018, out of 52 agreements across a range of issues for peace, only four contained gender-related provisions (UN Women 2020). In 2019, only 41 per cent of the UN Member States had adopted a national action plan (NAP) for the UNSCR 1325 (Desmidt and Davis 2019). Though it is still early to evaluate its impact, the novel coronavirus (COVID-19) has and will continue to aggravate existing intersectional risks, strain the coping capacities of those most vulnerable, increase marginalisation and it is quite likely that these will be variables to consider in current or ongoing peace processes. This systematic review operationalises UNSCR 1325 by incorporating its four pillars into our framework to understand and classify intervention groups.

1.2 The intervention

1.2.1 Conceptual scope of the review

A wide breadth of different interventions is currently being implemented to help empower women in FCAS contexts. As per Table 6 below, to structure our review of this broad range of interventions and to specify the types of intervention we include, we have developed a guiding framework to:

1. **Include only those intervention groups that are either *gender-specific* or *gender-transformative***, as guided by the World Health Organization's *Gender Responsiveness Assessment Scale* (World Health Organisation n.d.). Gender-specific interventions are those that consider gender norms, roles and relations for women and men and how they affect access to and control over resources. They intentionally target and benefit a specific group of women or men to achieve certain policy or programme goals or meet certain needs. Gender-transformative interventions are those that address the causes of gender-based inequities and include ways to transform harmful gender norms, roles and relations to foster progressive changes in power relationships between women and men. Given that our review aims to identify the effectiveness of interventions whose specific and explicit aim is to empower women in FCAS, we have used this device as a filter to exclude studies which may include a disaggregation for sex in their study design, but for which the intervention was more broadly targeted. As a result, we considered intervention groups of *all* genders and ages, with an explicit focus on gender. This includes any focus on Lesbian, Gay, Bisexual, Transgender, Queer and/or Questioning, Intersex and Asexual (LGBTQIA+) groups to represent sexual and gender diversity.

2. Include only intervention groups that operate within the micro- (individual) and meso- (household and community) spheres of the ecological framework not those which operate at the macro (or society) level. As is outlined by Bronfenbrenner, we recognised that an individual's immediate setting "...is affected by the larger context in which the settings are embedded" (Bronfenbrenner 1979). That said, our decision to focus on the lower levels of the ecological framework was driven by this review's focus on those interventions which work directly with individuals and their communities.
3. **Classify interventions based on the four pillars of UNSCR 1325.** While some interventions work across multiple pillars, we have decided to operationalise this widely used framework to help organise our findings. In doing so, our review is in line with national, international, and multilateral development priorities and the evidence that we present is easily navigable and applicable for policy makers and programme managers.

Table 6: Conceptual scope of the systematic review

Outcome	Gender approach	Ecological framework	UNSCR 1325 pillars
Women's empowerment and gender equality in FCAS	Gender unequal	Society	Participation
	Gender unaware		
	Gender sensitive	Community and institutional	Protection
	Gender specific	Household and interpersonal	Prevention
	Gender transformative	Individual	Recovery and relief
Included			
Excluded			

These UNSCR 1325 pillars are larger than the possible scope of this review, and not all interventions within them are included. Further detail follows in the summary of included interventions.

1.2.2 Summary of included interventions

The scope of efforts which aim to empower women in FCAS is vast, and likewise, the body of impact evaluations significant. In Table 6, we operationalise our conceptual scope as a tool to help identify specific categories of interventions to include in this systematic review. As an extension of that, the following are descriptions of the categories of interventions we intended to include within each of the four UNSCR pillars.

Participation

For our review, we define participation interventions as those that create opportunities for, build acceptance of, or strengthen capacities for the equal participation and full involvement of women and girls in political, economic and social institutions and decision-making processes.

Table 7: Participation pillar, included studies

Intervention	Description
Civil society, associations and networks	Interventions that aim to strengthen the technical capacity and institutionalisation of civil society, associations and networks, governments and market driven associations working for gender equality or women's empowerment. This also includes the establishment and support of women's economic associations, such as cooperatives and activity groups.
Economic rights and entitlements	Interventions that protect or support women's access to land rights and entitlements, such as unemployment benefits.
Formal education	System, school, and teacher-level interventions that promote gender equality, girls' access to and the gender sensitivity of formal education, and/or develop human capital amongst girls and women through school-based education. This can also include education financing and planning interventions such as vouchers for girls to attend low-fee private schools.
Non-formal education	Access to education and educational resources for women and girls that occurs outside of traditional schools including community-based education, camp-based education, etc.
Technical and vocational education and training	Education for women and girls that provides knowledge and skills useful for employment, jobs placement in the formal sector.
Micro, small and medium enterprises	Support for women-run micro, small and medium enterprises. This may include incubation-style interventions, loans, technical support, entrepreneurship training, mentorship or start-up kits.
Cash-based approaches to support women's access to and participation in education and/or the economy	Cash-based interventions, which aim to develop human capital (e.g., enable girls' access to education) or empower women to participate in the economy.

Intervention	Description
Financial inclusion	Interventions that provide access to credit, savings, insurance and other financial products, either on an individual or group-based approach. This includes self-help groups, village savings and loan associations, savings groups, etc.
Community and leisure activities	Interventions that create opportunities or encourage support for women's participation in community and leisure activities, such as sports, art, theatre, etc.
Civic education and youth leadership	Interventions that build capacity and create opportunities for women and girls to understand and participate in politics and political processes. This also includes support to girls and young women to learn skills for and access mentorship and opportunities to become active members of their communities.
Voice and participation in local and subnational governance and development bodies	Interventions that aim to increase the voice and participation of women and girls in local government through quotas, outreach and encouragement campaigns, etc. This also includes efforts to strengthen the responsiveness of local and subnational governance and development bodies to women's needs and priorities. Finally, it includes interventions that provide training or opportunities enabling women and girls to hold public services and government accountable for service provision and protection of their rights.

Prevention

For our review, we define prevention pillar interventions as those that build capacities and systems to support the gender-responsiveness and inclusivity of violence prevention and conflict transformation processes. This also includes efforts to hold perpetrators of violence accountable through formal or informal means.

Table 8: Prevention pillar, included studies

Intervention	Description
Conflict early warning systems	Interventions that facilitate women's involvement in, or the gender-inclusiveness or gender-responsiveness of community-based or online conflict early warning systems.
Dialogue groups	Interventions that either facilitate women's involvement in or set up women-specific community dialogue groups. These can be forums that draw participants across the community to exchange ideas in face-to-face moderated sessions, share personal stories and experiences, express perspectives, clarify viewpoints, and develop solutions to problems. These may include community conflict prevention fora, dialogues to resolve or transform

Intervention	Description
	specific conflicts, or dialogues to build social cohesion.
Community consultations	This includes interventions that create specific consultations for women at community and subnational levels as part of a formal peace process.
Capacity building for conflict transformation	Interventions that build women's capacity to participate in or the gender-sensitivity of community and subnational conflict transformation processes. These may also be referred to as building skills for mediation, negotiation, conflict resolution or conflict prevention.
Peace education	This includes interventions that aim to teach people about the importance of including women in peace processes.

Protection

For our review, we define the protection pillar interventions as those that create, facilitate access to, or build awareness of and support for legal or social protections for women's and girls' rights. This includes behavioural, legal and environmental interventions that aim to reduce women and girls' risk of experiencing sexual and gender-based violence (SGBV).

Table 9: Protection pillar, included studies

Intervention	Description
Legal rights education	This includes interventions that disseminate information and build understanding of women's rights, among men and women, and capacity-building on ways to demand their rights.
Behaviour change communication around support for women's rights and preventing SGBV	Community-based and subnational-level interventions that aim to change behaviours, attitudes and beliefs around gender equality and the role of women in society, as well as efforts to reduce the prevalence of SGBV. This could include the targeting of gender norms within communities through social institutions, such as churches, community groups or cooperatives. It could also include efforts to sensitise families on the negative consequences of SGBV, particularly for children. These interventions may target local community and civil society leaders, local and subnational government, the private sector, etc. This would also include media interventions such as radio dramas or participatory theatre aiming to encourage discourse around women's rights and empowerment.

Intervention	Description
Capacity building and technical support to subnational government officials to strengthen service provision for women and gender equality	These interventions build the capacity of subnational and local government officials to provide services for women, meet women's needs, and improve understanding of women's rights and the importance of gender equality. This may target gender equality-specific officials such as officials within a Department of Women's Affairs or aim to strengthen the capacity and service provision for women within sector-specific services such as education officials.
Preventative protection measures	These interventions comprise non-police or security force-based efforts to reduce incidences of violence, especially SGBV. They include making the physical environment less conducive to such acts and minimising the exposure of vulnerable groups to risky situations. This can include crime prevention through environmental design intervention, installing lighting in public spaces, women-friendly transport systems and public facilities, removing obstacles so there is a better line of sight and reclaiming spaces for positive community activities. They may also include interventions to reduce vulnerable groups' risk of exposure to violence (e.g., through providing firewood or alternative fuel sources to women in refugee camps).
Gender-sensitive policing	Interventions that strengthen the approaches and capacity of police forces to prevent SGBV, improve police support to and treatment of women and girls, and support victims in a respectful way that does no harm. This includes, for example, training and capacity building interventions targeting police forces that strengthen their awareness of the importance of gender equality and women's legal rights, train them on best practices for protecting victims, and interventions that create women-specific police officers to facilitate access to reporting crimes.
Informal judicial system	Interventions supporting women's access to or participation in informal or semi-formal justice processes, such as alternative dispute resolution mechanisms or village courts

Relief and recovery

For our review, we define relief and recovery interventions as those that build capacities and systems to support the gender-responsiveness and inclusivity of relief and recovery processes related to conflict, displacement and natural disasters.

Table 10: Recovery and relief pillar, included studies

Intervention	Description
(Re)integration of forcibly displaced populations	This includes gender-specific or transformative interventions that target forcibly displaced populations, with the aim of supporting (re)integration into host or home communities. This may include economic support such as cash-based approaches (e.g., cash transfers and vouchers); social reintegration such as through the organisation of social activities or sports leagues; and community development inclusion such as strengthening the inclusion of displaced women in community driven development interventions.
Disaster risk reduction	Interventions to support women's participation in or the gender responsiveness of community-based disaster risk reduction activities (such as hazard vulnerability and risk assessments)
Disarmament, demobilisation and reintegration (DDR)	These interventions support forcibly displaced populations to (re)integrate into host or home communities. This may include economic support such as cash-based approaches (e.g., cash transfers and vouchers); social reintegration such as through the organisation of social activities or sports leagues; and community development inclusion such as strengthening the inclusion of displaced women in community driven development interventions.

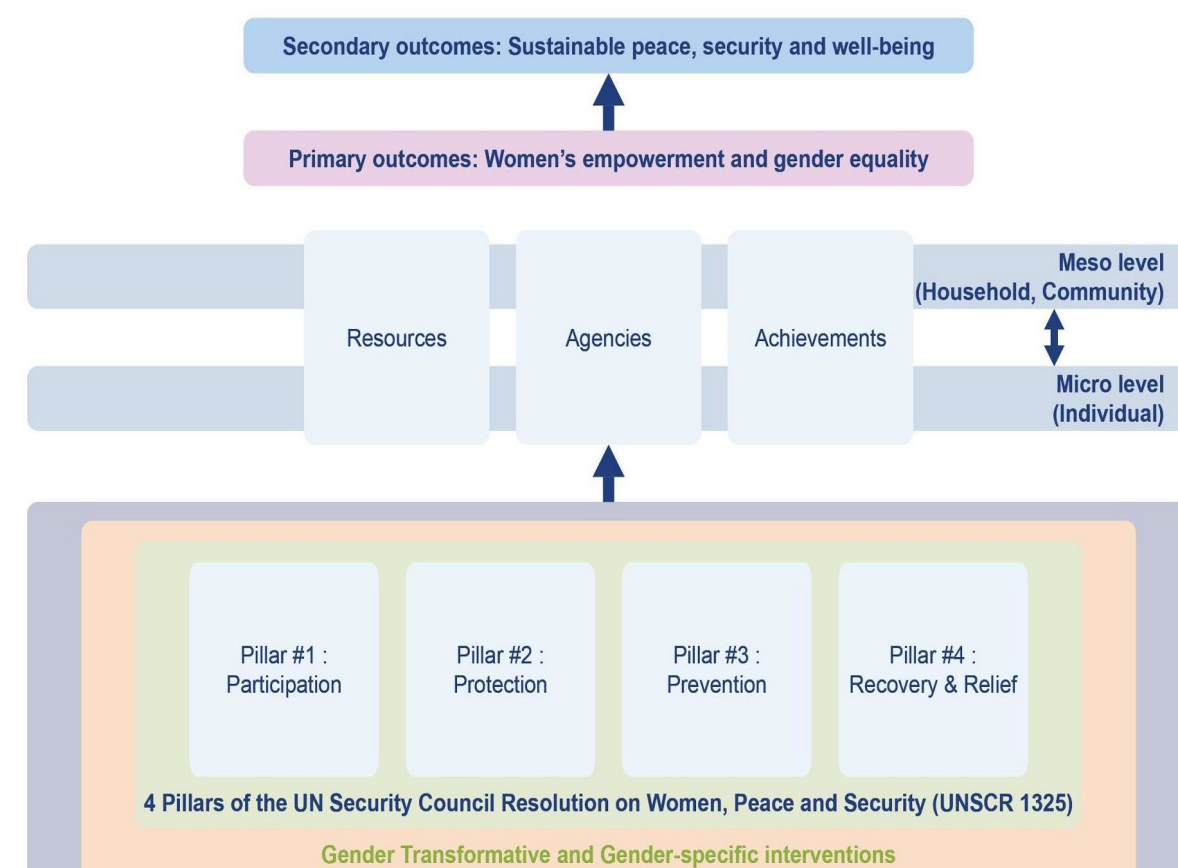
1.3 How the intervention might work

Our framework acknowledges both the wider and more specific contexts within which gender-specific and transformative interventions are implemented in FCAS. Interventions operate within a wide context of international standards and practices, including UNSCR 1325, the 2030 SDGs or the WPS Agenda, and this global context overlays the unique local context of any given intervention where those general standards are applied. The framework captures the reality that while broader principles often guide intervention design, the characteristics of fragility and conflict depend on the location, motivations, interactions and local specificities of the intervention fields.

As per the conceptual scope, the below Theory of Change (Figure 2) includes interventions operating in the micro- and meso-levels of the ecological framework and that are gender-specific or transformative and those that work to develop the resources, agency and achievement levers of women's empowerment and gender equality.

Importantly, the UNSCR Pillars are not necessarily mutually exclusive. Interventions can work across multiple pillars, levels and levers with the common objective to impact women's empowerment and gender equality. Lastly, the primary outcome of the review is women's empowerment and gender equality that will then contribute to achieving the secondary outcome of peaceful and inclusive societies.

Figure 2: Our systematic review's theory of change



1.4 Why it is important to do this review

1.4.1 Review of existing literature

The existing literature

3ie's recent EGM, *Building peaceful societies*, demonstrated that a considerable body of literature exists regarding the impact of interventions that focus on women and girls in FCAS. There is a wide scope of experimental and quasi-experimental primary research focusing on peaceful and inclusive societies and women's empowerment. As is outlined above, much of this evidence suggests a link between the integration of gender-sensitive approaches in FCAS context and lasting achievements in building, ensuring human rights, and promoting inclusive societies. That said, only a handful of systematic reviews focused on the role of women in peace processes and there is a stronger emphasis on empowerment of women in fragile contexts for sustainable peace.

After completing a comprehensive scoping of existing SRs and impact evaluations (IE) through the 3ie Evidence Portal, we identified seven IEs and two EGMs focusing on the topics of building peaceful societies and fragility. 3ie notably contributed to the production of the two evidence gaps maps focusing on *Building peaceful societies* (Sonnenfeld et al. 2020) as an update and expansion of the 2015 EGM on *Evidence for peacebuilding* (Cameron et al. 2015). The *Evidence for peacebuilding* EGM identified over 100 IE, four SR and two protocols. The *Building peaceful societies* EGM identified about 250 IEs, 34 SRs and four protocols.

Women's empowerment in fragile contexts has been substantially reviewed. A search of 3ie's Development Evidence Portal identified 39 impact evaluations, 16 systematic reviews, and one EGM focusing on topics of women and girls' empowerment, including contexts in FCAS.¹ As identified through the focus of the EGM (Picon et al. 2017), most of the reviews with an explicit and specific focus on women in FCAS focus on the prevention of and recovery from intimate partner violence (IPV).

There is also a substantial literature of SRs and IEs integrated within the second pillar of our framework (prevention) with a high prevalence of IEs focusing on the incidence or reaction to IPV with 62 out of 153 identified IEs. Sexual and gender-based violence (SGBV) and IPV are also the focus of five SRs and two protocols (Delkhosh et al. 2017; Signorelli et al. 2018).

The Spangaro-led review (Spangaro et al. 2013) notably found that although guidelines and training courses have been developed at the policy level, few initiatives have been implemented and evidence is lacking to judge the effectiveness of interventions designed to address or prevent sexual violence in conflict- or crisis-affected settings. IPV has also been the focus of a *systematic review of structural interventions for IPV in L&MICs* (Bourey et al. 2015) and *in humanitarian crises* (Warren et al. 2015). Some reviews also focused on the recovery and health aspects of IPV (Rivas et al. 2015; Tol et al. 2013).

Our literature scoping also identified a cluster of studies focusing on women's economic empowerment. Reviews of economic empowerment interventions include community-based programmes such as economic self-help groups where some reviews identified positive effects on women's political and economic empowerment (Brody et al. 2017). Others focus on the role of economic resources transfer for women's empowerment (Kabeer and Waddington 2015). A third type of economic empowerment focuses on the impacts of business-oriented programmes such as business and vocational training to improve women's labour market outcomes that tends to have more impact when they are combined with cash transfer or life skills training (Balarin et al. 2017).

Several reviews include a specific focus on livelihoods (Gibbs et al. 2012) at the community level to complement reviews integrating a health lens at the individual level (Kraft et al. 2014). Such characteristics of the literature landscape confirms the relevance of analysing women's empowerment through an ecological framework.

The political dimension of empowerment is also highly reliant on the community level of the ecological framework through the mobilisation of civil society groups to serve empowerment purposes notably in the sub-category of reproductive health for marginalised groups (Handanagic et al. 2016; Moore et al. 2014). Although we were able to find numerous references to gender empowerment in systematic reviews, none of them have a specific focus on gender empowerment for and through building peaceful and inclusive societies processes in FCAS.

Additionally, though women's empowerment in fragile contexts is at the core of our review, it is important not to forget the wider spectrum of governance and social

¹ The team used the following keywords for the literature search: "gender", "women empowerment", "adolescent girls' empowerment", "gender-based violence", "intimate partner violence".

protection themes integrated in this review. The 3ie Evidence Portal gathers over 170 impact evaluations, 11 systematic reviews and two evidence gap maps, including one in-progress, covering issues of governance. Themes such as accountability, transparency, management of resources, public administration are all relevant aspects for the analysis of fragile and conflict-affected situations as they are both sources of conflict and opportunities for peace. The two EGMs notably focused on the State-Society relations (Coffey et al. 2017) (18 systematic reviews, two protocols and 365 impact evaluations) with a major focus on public institutions and services and on the effect of rule of law interventions on justice outcomes (Doherty et al. 2020).

Our contribution to the existing literature

The above summary of findings from existing reviews reveals that although both fragility and women's empowerment are key areas of focus of systematic research, only a handful of studies focus on women's empowerment in FCAS and/or gendered approaches to address fragility. Even when they do, there usually is only particular emphasis on the impact of fragility on women rather than women's roles in shaping peace.

This initial summary confirms the relevance of the conceptual scope of our review as we saw that several reviews focused on some of its aspects, including a specific attention to gender-transformative and gender-specific IE in the systematic reviews covering women empowerment that matches our approach. Many of the systematic reviews focus through the lens of one of the levels of the ecological model with many of the reviews focusing on the community level. Finally, gender- and peacebuilding-focused SRs overlapped on one or several of the UNSCR 1325 pillars (participation, prevention, protection and recovery and relief).

These gaps identified in the systematic literature should not be interpreted as a lack of interest for women's empowerment for peace and inclusivity. This literature review allowed us to identify several impact evaluations already covering relevant aspects for our study. The gaps of literature are mainly due to the challenging analytical aspects that this topic raises. The high amount of non-systematic literature covered through the scoping stages of our research is a good illustration of the interest of these topics for researchers. This review contributes to the systematic review literature through the development of understanding of the causal impacts of gender sensitive and gender transformative interventions in FCAS.

1.4.2 Relevance to policy and practice

Relevance to international and national development goals

An objective of our study is to feed into some of the international goals and approaches supporting women empowerment and gender equality in fragile contexts. Internationally, the interest in gender equality and its relationship to peaceful societies can be measured in the number of investments being made in both the SDGs relevant to gender equality and fragility, and national and regional-level initiatives. An estimated \$1 trillion USD would be necessary to meet SDG #5 on gender equality and #16 on peaceful and inclusive societies (Peace Women 2016). There is growing recognition that investment in military spending is not the only solution for peace, but that solution can also come through an increased investment and recognition of women's rights and gender equality worldwide. Some international mechanisms have been created, such as the Addis

Ababa Action Plan on Transformative Financing for gender equality and women's empowerment (Peace Women n.d.), the Compact on WPS-HA of the Generation Equality Forum, and the Global Acceleration Instrument (GAI) for Women, Peace and Security and Humanitarian Action, that represent pooled, transnational sources of funding for gender equality and women's empowerment.

At the national level, donors and countries have taken some measures to target women's empowerment and gender equality in FCAS since 1325 was introduced but more generally in the support of achievement of international development and growth. The German international development agency, GiZ, published its vision for promoting gender equality as one of the key values of its approaches abroad and as a prerequisite and driver of sustainable development (Langenkamp n.d.). Global Affairs Canada (GAC) updated its international assistance policy in 2020 to include a feminist perspective. The updated policy recognises the importance of supporting gender equality and the empowerment of women and girls as it is the best way to build a more peaceful, more inclusive and more prosperous world (Global Affairs Canada 2017). The United Kingdom's former Department for International Development (DfID), which is now the Foreign, Commonwealth, and Development Office (FCDO), published a Strategic Vision for gender equality in 2018. The document is presented as a call to action, recognising that action is needed if gender equality is to become a lasting reality (Department for International Development n.d.-c). The Ministry of Foreign Affairs of the Kingdom of the Netherlands publishes reports every four years to focus its impact on UNSCR 1325 and the Women, Peace and Security Agenda (Ministry of Foreign Affairs n.d.).

Relevance for donors and intervention design

Donors and implementing agencies have led interesting and innovative interventions in FCAS to increase the role of women in fostering peaceful societies.

The German Ministry for Economic Cooperation and Development (BMZ) implements a project in conjunction with the African Union (AU), entitled 'Support to the African Peace and Security Architecture' (APSA), to work with regional organisations and other international stakeholders to improve the participation of women in activities in the peace and security sector. BMZ has also supported the operationalisation of the AU-supported Network of African Women in Conflict Prevention and Mediation ([FemWise-Africa Network](#)), which collaborates with civil society to build the mediation skills of young women and bridge intergenerational gaps. The project seeks to promote the role of women in preventive diplomacy and mediation at all levels, ensure channels for their meaningful participation, initiate their action in line with the SDGs, bridging the gap between different levels of mediation and establishing local and national peace infrastructure for long-term initiatives. Germany also recently adopted its third UNSCR 1325 National Action Plan (NAP) for the period 2021-2024 following its two previous plans for 2017-2020 and 2013-2016.

BMZ also supports compliance training addressing fundamental human rights and zero tolerance for sexual abuse and exploitation (SEA). More recently, BMZ has offered support to the implementation of the regional stabilisation strategy (RSS) for the Lake Chad Basin region, mental health support in Boko Haram-affected regions and hosted regional workshops on SGBV.

Since 2011, the *Red Nacional de Mujeres* publishes an annual report on monitoring the progress of the implementation of UNSCR 1325 (Red nacional de mujeres n.d.), which has been in conjunction with the country's NAP on 1325. *Ruta Pacífica de las Mujeres*, an organisation that was key in advocating for the inclusion of women's voices in the 2016 Colombian Peace Treaty, includes a list of indicators across all levels of the ecological framework in their strategic plan to monitor progress on women's empowerment in the Colombian context (Ruta pacífica de las mujeres 2016).

In Pakistan, the FCDO funds the Voice and Accountability programme (AAWAZ), which brings together three critical components focusing on the interconnection of gender, conflict resolution and citizen engagement (Department for International Development n.d.-a). This strives to focus on the enhanced political participation of women and in larger public life without fear of sexual or gender-based violence, while striving to attain social harmony. This is set against the backdrop of promoting civil society, by having active and informed participation of citizens and their organised groups. More broadly, the FCDO also funds the Global Challenge Research Fund (GCRF) Gender, Justice and Security Hub (Department for International Development n.d.-b). On a higher level, the Hub functions to deliver ambitious and impactful research facilitating gender justice and inclusive security in FCAS.

The Dutch Ministry of Foreign Affairs funds, through its SDG5 Fund, a variety of modalities, including the Women, Peace and Security fund, the Power of Women fund and grants for protecting women and girls sexual and reproductive health rights (SRHR). Consequently, the overall objective of the strategic WPS-partnerships is to contribute to an enabling environment for women's participation and empowerment in conflict and post-conflict environments, so they can meaningfully participate in conflict prevention, resolution, fostering peaceful societies, protection, relief and recovery.

Relevance for policy design and decision making

Although it is still an ongoing process, we can already see some positive impacts of the WPS agenda. Data notably shows an increase in bilateral aid in support of gender equality and women's empowerment in FCAS (UN Women n.d.) from some of the key states of funding of international development programmes. According to UN Women, all top bilateral donors increased their proportion of spending in support of gender equality and women's empowerment in FCAS over the last decade alone. For example, the United States gave approximately \$2 billion USD in 2010, yet increased that number still to nearly \$3 billion in 2017. However, we see that the proportion allocated by these same countries to bilateral aid dedicated to gender equality and women's empowerment as a principal objective in FCAS is lower than 6 per cent of aid and has decreased over the last decade.

The adoption of National Action Plans (NAP) by individual countries to adopt the WPS Agenda as outlined in UNSCR 1325 has some encouraging results in Central and West Africa and Latin America but is still an ongoing process in Southeast Asia and in the MENA region; we see a relatively low rate of adoption of NAP in medium to high level of violence conflict regions (UN Women n.d.). Some examples of NAP are briefly presented below (London School of Economics 2019). See Online Appendix I Table 1 for some examples of different National Plans for Action on UNSCR 1325.

Since the first NAP was pioneered in 2005 by Denmark, over 80 countries have piloted their own NAPs focusing on the WPS agenda (Peace Women, 2014). However, despite there being nearly 84 NAPs (as of December 2019), very few (33%) contain an allocated implementation budget. Additionally, only 31% have references to disarmament. Lastly, while the majority of NAPs mention the role of grassroots leaders and civil society (75%), most allocate them to an advisory role. There are 11 additional regional action plans (RAPs), such as for regional bodies like the AU and EU. Regional coordination efforts also include the Asia-Pacific Regional Symposium on National Action Plans on Women, Peace and Security where the Member States, alongside civil society representatives, share their lessons learned and best practices in the implementation of UNSCR 1325.

Twenty years after the introduction of UNSCR 1325, the purpose of NAPs and their impact has yet to be fully explored. The dynamic nature of conflict and the intersections of gender often mean that perspectives are constantly shifted and not always accurately captured in plans or data (Hamilton et al. 2020). Additionally, without implementation plans or budgets, NAPs ultimately end up being solely documents that do not move past rhetoric towards commitment. These trends illustrate the challenges raised to increase the role of women in building peaceful and inclusive societies. Our review then helps inform decisions about how to spend the resources available by providing a comprehensive review of the evidence on the impact of women as agents of change to build peaceful and inclusive societies in FCAS.

2. Objectives

This review builds on 3ie's EGM of the impact evaluation and systematic review evidence base of interventions aiming to promote peaceful and inclusive societies in fragile contexts (Sonnenfeld et al. 2020). The EGM identified a cluster of studies evaluating gender equality-focused behaviour change communication programmes and raised interest in investigating the evidence base for understanding the role of women more broadly as agents of change in developing peaceful and inclusive societies.

Building on the cluster of evidence identified in the EGM, our review increased generalisability of findings from single studies and focuses on interventions across a broad range of geographical locations, settings and populations, types of implementations and outcomes. We also address (when possible) the identified gaps in literature regarding meta-analysis in conflict-affected contexts.

As such, we propose the following objectives:

1. The primary objective of this review is to identify, assess and synthesise evidence on the effect of gender specific and gender transformative interventions within the context of the four pillars of UNSCR 1325 on women's empowerment and gender equality in FCAS. The SR facilitates the use of evidence in informing policy and practice decisions within the field of transition aid, particularly as it relates to gender focused programming.
2. Our second objective is to assess how these interventions contribute to inclusive and sustainable peace in conflict affected situations. We compare the effectiveness of these different types of interventions through the lenses of their ecological level, types of impact on women's empowerment, local context of gender inequality and conflict.

To achieve these objectives, we aimed to answer the following questions:

1. What are the impacts of gender transformative and specific interventions on women's empowerment and gender equality in FCAS?
2. What are the effects of these interventions on sustainable peace?
3. To what extent do effects vary by population group, ecological level and types of interventions?
4. What are contextual barriers to and facilitators of intervention effectiveness?

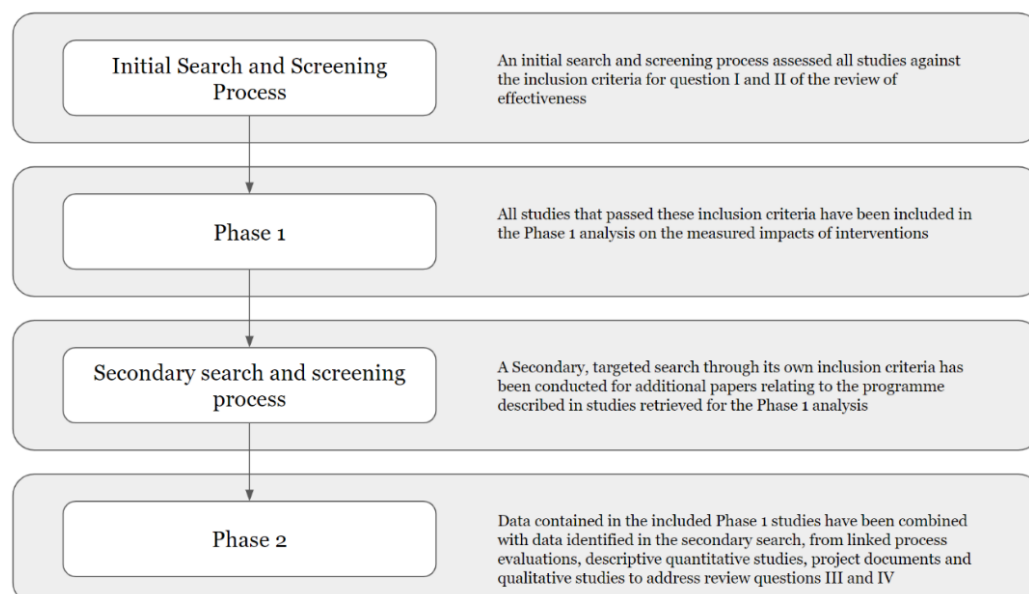
3. Methods

The review followed the Campbell and Cochrane Collaborations approaches to systematic reviewing (Cochrane Collaboration n.d.; Hammerstrøm et al. 2010; Shemilt et al. 2013). The review also drew on the concepts of theory-based impact evaluation (White 2009) and theory-based systematic reviews (Snijlsteit 2012) to provide a mixed-methods systematic review and analysis along the causal chain, reviewing evidence on context, process and implementation to identify barriers and facilitators and analyse the effectiveness of interventions aiming to improve women's empowerment and gender equality outcomes.

The review systematically collected and synthesised quantitative evidence from impact evaluations meeting our inclusion criteria to answer our review questions. If sufficient data was available, outcomes were synthesised in sub-groups of populations, types of interventions, and local contexts. For the review to be more useful for policymakers and practitioners, we also collected qualitative evidence from the included studies to assess factors that determine or hinder the effectiveness of interventions using a combination of qualitative synthesis and meta-regression analysis.

The review included studies in two phases as per figure 3 below. To address questions I and II, we included studies meeting the impact evaluation study design inclusion criteria. To address the questions III and IV, IE studies that pass these criteria were then used as the basis for a second phase to identify and include qualitative studies, project documents, and process evaluations.

Figure 3: Overview of the review process



3.1 Criteria for considering studies for this review

3.1.1 Types of study design

We included studies in two stages, in a similar approach to Snilstveit and colleagues (2015; 2018). In the first stage, we included studies that assessed the effects of interventions using experimental designs or quasi-experimental designs with non-random assignment that allow for causal inference (to address primary research objectives 1 and 2). Specifically, we included the following:

- Randomised controlled trials (RCTs), with assignment at individual, household, community or other cluster level, and quasi-RCTs using prospective methods of assignment such as alternation.
- Non-randomised studies with selection on unobservables:
 - Regression discontinuity designs, where assignment is done on a threshold measured at pre-test, and the study uses prospective or retrospective approaches of analysis to control for unobservable confounding.
 - Studies using design or methods to control for unobservable confounding, such as natural experiments with clearly defined intervention and comparison groups, which exploit natural randomness in implementation assignment by decision makers (e.g., public lottery) or random errors in implementation, and instrumental variables estimation.
- Non-randomised studies with pre-intervention and post-intervention outcomes data in intervention and comparisons groups, where data are individual level panel or pseudo-panels (repeated cross-sections), which use the following methods to control for confounding:
 - Studies controlling for time-invariant unobservable confounding, including difference-in-differences, or fixed- or random-effects models with an interaction term between time and intervention for pre-intervention and post-intervention observations; and
 - Studies assessing changes in trends in outcomes over a series of time points (interrupted time series, ITS), with or without contemporaneous comparison (controlled ITS), with sufficient observations to establish a trend and control for effects on outcomes due to factors other than the intervention (e.g., seasonality).
- Non-randomised studies with control for observable confounding, including non-parametric approaches (e.g., statistical matching, covariate matching, coarsened-exact matching, propensity score matching) and parametric approaches (e.g., propensity-weighted multiple regression analysis).

In a second stage, to address questions 3 and 4 on factors related to intervention design, implementation, and context, we extracted descriptive and qualitative data from the included experimental and quasi-experimental studies. In addition, we conducted a targeted search for additional papers on the interventions covered by the included impact evaluations to provide additional detail on these factors. We have included paper related to the intervention in the included impact evaluations and corresponding to one or more of the following types of studies (Snilstveit 2012):

- A *qualitative study* collecting primary data using mixed- methods or quantitative methods of data collection and analysis and reporting some information on all of the following: the research question, procedures for collecting data, procedures

for analysing data, and information on sampling and recruitment, including at least two sample characteristics.

- A *descriptive quantitative study* collecting primary data using quantitative methods of data collection and descriptive quantitative analysis and report some information on all of the following: the research question, procedures for collecting data, procedures for analysing data, and information on sampling and recruitment, including at least two sample characteristics.
- A *process evaluation* assessing whether an intervention is being implemented as intended and what is felt to be working well, and why. Process evaluations may include the collection of qualitative and quantitative data from different stakeholders to cover subjective issues, such as perceptions of intervention success or more objective issues, such as how an intervention was operationalised. They might also be used to collect organisational information.
- A *project document* providing information about planned, ongoing or completed interventions. Such documents may describe the background and design of an intervention, or the resources available for a project. As such, these documents do not typically include much analysis of primary evidence, but they provide information about interventions. The purpose of including them in our review was to ensure we had sufficient information about the context and interventions in included studies.

3.1.2 Types of participants

We included participants from fragile communities in low-and middle-income countries (L&MICs), including participants from the general population from all ages and genders and those from specific population sub-groups, such as displaced populations, refugees, women, youth, etc. This wide definition of the population then allowed us to run subgroup analyses in our review such as rural and urban populations, socioeconomic groups or classes and castes, age, sex etc.

Challenges for the population definition

The definition of the population for this review raised a series of challenges to be considered:

- **The combination of women's empowerment and fragility criteria:** there is a gap in data available on both women's empowerment *and* fragile contexts. The team had run a scoping of indicators, indexes, and data available on women's empowerment in FCAS. The only index that covers both topics is the Women Peace and Security index designed by the University of Georgetown in 2017 but this index is not consensual in the research sphere as an indicator we could use on its own (Mundkur and Shepherd 2018). Although this index is relevant to our SR, its time frame was not close enough to our start date of intervention (2000) to be used as the only criteria for inclusion. This gap then required the team to build its population criteria on the combination of existing indexes on gender *or* on FCAS.
- **The analysis of conflict affected situations at state-level:** as presented in the previous sections, the study focuses on the micro- and meso- levels of the ecological framework. There are currently no indices that allowed us to draw our population selection on community and local levels indicators of fragility. Data on fragility are generally available at the State level with indexes such as the OECD and World Bank FCAS lists, the Fragile State Index (FSI) and the Global Peace Index (GPI). This gap then required the team to build its population criteria on

State-indicators although the analysis focused on lower levels. While the core team attempted to extract data from available international datasets on femicide, homicide and other SGBV-related indicators, most datasets were incomplete, resulting in an inability to add any value to the analysis.

- **The potential size of the population:** the time extension and thematic focus would have made almost all low- and middle-income countries (L&MICs) eligible for inclusion in our SR. The resources available and timeframe would then not make it do-able for the team and would have required the team to set a series of thresholds on the identified indexes to build its inclusion criteria. This then meant that some countries and regions have been prioritised for this study and that some excluded could have brought relevant studies for inclusion.

Methodology of selection of the population

We created binary variables for whether a country was considered fragile or not, or gender unequal or not depending on the index, for each year for which data were available, using the thresholds noted above. We then created a series of consolidated lists for whether a country had appeared as fragile or unequal for at least one year during the period 2000-2020. We then grouped the countries into five sets, as follows:

- Set 1: There are 51 countries that appear on each of the World Bank, FSI², OECD and GII³ lists for at least one year.
- Set 2: There are 62 countries that appear on the GII and at least one of the fragility indices for at least one year (FSI, OECD or World Bank).
- Set 3: There are 23 countries that appear on the WPS list, but do not fall within Sets 1 or 2. This includes 10 countries that appear on the gender unequal GII list but none of the fragility indices, including contexts such as Brazil, El Salvador and India that were highlighted as potentially relevant contexts during the AG call. It further includes three countries that appear on the FSI list but not the GII index. Finally, it also includes contexts such as Mexico that were identified as of interest by members of the Advisory Group (AG) due to high levels of violence.
- Set 4: There are 11 countries that have appeared on one of the fragility indices since 2000, but not the GII or WPS.
- Set 5: There is one country that appeared on the GII list for at least one year, but never on the WPS or other fragility lists (Argentina).

² The Fragile States Index (FSI) is an multifaceted indicator of state fragility managed by the Fund for Peace that captures content analysis, qualitative and quantitative data to “highlight not only the normal pressures that all states experience, but also ... identify when those pressures are outweighing a states’ capacity to manage those pressures to ... make political risk assessment and early warning of conflict accessible to policy-makers and the public at large.”

³ As per UNDP definition : “The GII is an inequality index. It measures gender inequalities in three important aspects of human development—reproductive health, measured by maternal mortality ratio and adolescent birth rates; empowerment, measured by proportion of parliamentary seats occupied by females and proportion of adult females and males aged 25 years and older with at least some secondary education; and economic status, expressed as labour market participation and measured by labour force participation rate of female and male populations aged 15 years and older. The GII is built on the same framework as the IHDI—to better expose differences in the distribution of achievements between women and men. It measures the human development costs of gender inequality. Thus the higher the GII value the more disparities between females and males and the more loss to human development.”

We developed two options for addressing the risk that including all countries within Set 3 at the title and abstract stage could create an unrealistic scope during full-text screening:

- Set 3a: restrict the threshold for the WPS index to <0.6 instead of <0.7. This would shrink Set 3 to three countries (Benin, Gabon and India). However, it would exclude the potential contexts highlighted by the AG as of interest.
- Set 3b: include only the 10 countries that show up on both the WPS and the GII (Benin, Botswana, Brazil, El Salvador, Gabon, India, Morocco, Paraguay, Peru and Saudi Arabia). This would effectively treat the WPS as a similar fragility index to the ones included within Set 2. From the above-noted countries, it would exclude Mexico, but it may be the most realistic approach that is nonetheless drawn on consistent criteria.

Inclusion criteria at Title and Abstract Screening (TAS) stage

Based on the challenges identified above and the preferred options for the sets 1,2 and 3b, we included studies based on the following population criteria at TAS stage:

- L&MICs that are listed on all our indicators: in the World Bank list of FCAS (since 2006), on the OECD list of FCAS (since 2007), on the FSI with a score higher than 90 (since 2009) and that have a score of at least 0.6 on the GII for at least one year since 2000.
- L&MICs that are listed in one of the fragility lists or had at least one year rated higher than 80 on FSI and that have a score of at least 0.6 on the GII for at least one year since 2000.
- L&MICs that do not appear on the fragility lists but are listed with a score lower than 0.7 on the Women Peace Security Index between 2016 and 2019 and that have a score of at least 0.6 on the GII for at least one year since 2000.

The above criteria allowed the team to build a list of included countries as below (Table 11 and Figure 4):

Table 11: List of included countries for the systematic review

Set	Countries	
Set 1 There are 51 countries that appear on each of the World Bank, FSI, OECD and GII lists for at least one year.	Afghanistan	Mali
	Angola	Marshall Islands
	Burkina Faso	Mauritania
	Burundi	Micronesia, Fed. Sts.
	Cambodia	Mozambique
	Cameroon	Myanmar
	Central African Republic	Nepal
	Chad	Niger
	Comoros	Nigeria
	Congo, Dem. Rep.	Papua New Guinea
	Congo, Rep.	São Tomé and Príncipe
	Côte d'Ivoire	Sierra Leone
	Djibouti	Solomon Islands
	Eritrea	Somalia
	Guinea	South Sudan

Set	Countries	
	Guinea-Bissau Haiti Iraq Kiribati Kosovo Lebanon Liberia Libya Madagascar Malawi	Sudan Syrian Arab Republic Timor-Leste Tonga Tuvalu Uzbekistan Vanuatu Venezuela, RB West Bank and Gaza Yemen, Rep. Zimbabwe
Set 2 There are 62 countries that appear on the GII and at least one of the fragility indices for at least one year (FSI, OECD or World Bank).	Algeria American Samoa Antigua and Barbuda Azerbaijan Bahrain Bangladesh Belarus Bhutan Bolivia Bosnia and Herzegovina Cabo Verde Chile Colombia Czech Republic Dominica Egypt, Arab Rep. Equatorial Guinea ESwatini Ethiopia Fiji Gambia, The Ghana Grenada Guatemala Guyana Honduras Indonesia Iran, Islamic Rep. Isle of Man Jordan Kenya	Korea, Dem. People's Rep. Kyrgyz Republic Lao PDR Lesotho Maldives Montenegro Namibia Nauru Nicaragua North Macedonia Northern Mariana Islands Oman Pakistan Palau Panama Puerto Rico Romania Rwanda Senegal Serbia Seychelles Sri Lanka St. Kitts and Nevis St. Lucia St. Vincent and the Grenadines Suriname Tanzania Togo Trinidad and Tobago Uganda Zambia

Set	Countries	
Set 3b This list includes countries that are not in the previous list and appear at least once in the WPS index also showed up on the GII for at least one year (similar to Set 2)	Benin Botswana Brazil El Salvador Gabon	India Morocco Paraguay Peru Saudi Arabia

Figure 4: Map of includable countries for the systematic review



Inclusion criteria at full text screening (FTS) stage

Given constraints in time and scope, for the full-text screening level, we developed a more nuanced operationalisation of fragility which considers the fact that fragility is not time constant, and decisions need to be made for each country within the context of the year in which the intervention took place. To do so, we applied the following inclusion criteria at full-text screening:

- Does the country score over 0.6 on GII at the start date of the intervention or under 0.7 on the WPS for studies starting between 2014 and 2020 (if intervention start date not available use publication date)?
- Does the country score over 80.0 on FSI or is listed on WB or OECD list at the start date of the intervention (if intervention start date not available use publication date)? or, if after 2016, does the country belongs to the fourth or fifth quintile of the WPS index (under 0.700)?

We are aware of some of the limitations of the international indices measuring fragility levels. An important limitation for us has been the absence of an index allowing us to measure fragility at the community and local level. To address this limitation, we made the decision that if a study was not meeting the FSI, WB or OECD criterion we would look at whether the local level of implementation of the study was meeting our definition of fragility, human security, building peaceful and inclusive societies and overall relevant characteristics of our definition of developing *peaceful and inclusive societies*. A study meeting the characteristics of our definition at the local or community level could then be included on this criterion.

Inclusion of refugee populations

Few populations are as inherently vulnerable, fragile, and exposed to conflict as refugees, including those that have been resettled as well as those that are living in camps. For that reason, this review has taken the decision to include any study whose explicit population is refugees from a country which would otherwise have been included and are living in a L&MIC that would not be otherwise included. An example of that would be the evaluation of an intervention targeting Syrian refugees in Jordan (an L&MIC country that would otherwise be excluded on the FCAS measure). Refugees living in a high-income country (HIC) have not been included in the population in our review.

3.1.3 Types of interventions

The main challenge of the definition of the intervention criteria for our SR is the wide range of interventions that fall under the four pillars of UNSCR 1325. In a similar way as for the population, including all the gender specific or transformative interventions that fall under the four pillars of UNSCR 1325 would be too broad. A set of exclusion grounds has then been agreed:

- We recognise the importance of the health sector for women's empowerment but considering the broad range of literature on this topics and existing research, health has not been included in the SR.
- We have not included studies focusing on the *treatment* of sexual and gender-based violence (SGBV) and only included studies focusing on preventing SGBV.
- We recognise the definition of the four pillars of UNSCR 1325 but have adopted a specific list of interventions under some pillars:
 - Participation: we have broadened the scope of the pillar to include interventions targeting economic, social and political participation.
 - Protection: we have narrowed inclusion to efforts to combat SGBV.
 - Recovery and relief: we have narrowed inclusion to reintegration intervention and DDR.

We used an intervention typology based on the four pillars to specify interventions groupings we included in our review. We will not repeat the descriptions of all interventions here, but Table 12 below summarises included interventions under each pillar:

Table 12: List of included interventions for the systematic review

Pillar	Intervention type
Participation	Civil society, associations and networks Land rights Formal and non-formal education Technical and vocational education and training Financial inclusion MSMEs and entrepreneurship support Cash-based approaches to support access to and participation in education and/or the economy Voice and participation in local and subnational governance and development bodies Community and leisure activities Civic education and youth leadership
Prevention	Conflict early warning systems Dialogue groups Community consultations Capacity building for conflict resolution Peace education
Protection	Legal rights education Behaviour change communication around support for women's rights and preventing SGBV Preventative protection measures Capacity building and technical support to subnational government officials to strengthen service provision for women and gender equality Gender-sensitive policing Informal judicial system
Recovery and Relief	(Re)integration of forcibly displaced populations Disaster risk reduction Disarmament, demobilisation and reintegration (DDR)

The following interventions do not meet the inclusion criteria and have been excluded from the review:

- *SGBV punishment*: we focused our work on proactive interventions for women's empowerment and gender equality and excluded reactive interventions focusing on the treatment of violence or inequality.
- *Formal peace processes*: this component would mainly focus on the higher levels of the ecological framework, and we only included the community and lower levels for inclusion.
- *Formal judicial system*: this group was excluded based on the ecological framework.
- *SGBV - medical and psychosocial treatment and services for victims*: this group was excluded on the basis that Health-focused intervention has not been included in this study and we have not focused on treatments of SGBV.

- *Family planning/reproductive rights*: this might be included if it was a component of a larger intervention but excluded if the intervention exclusively focuses on it.
- Capacity building of national-level actors on the importance of and how to support women's empowerment and gender equality: this group is excluded based on the ecological framework.
- *Women's health*: health has not been covered in this review.
- *Education in emergencies*: Education has only been covered in the aspect of peace education but not general education.
- *Monetary policy*: these interventions are rarely gender-specific and mainly are concentrated at the national/society level.
- *Cash-based approaches to meet basic needs*: these might be included if it is a component of a larger intervention but excluded if the intervention exclusively focuses on it.
- *Humanitarian aid and emergency response*: this intervention type is too broad, and we have exclusively focused on conflict-related emergencies.

Both our list of inclusion and exclusion are not exhaustive, and we therefore expect these lists to change based on the results of our search and screening process.

3.1.4 Types of comparison

We included studies that compare the effects of an intervention aiming to empower women in fragile and/or conflict-affected situations against similar situations where individuals, households and/or communities either received a “placebo” intervention or “business as usual” conditions.

3.1.5 Types of outcome measures

The review takes a rights-based approach that recognises the importance of gender equality as a priority in and of itself. Given the expected limited evidence for effects of gender equality programmes on outcomes of sustainable peace, the primary objective of the review is to understand the effects on women's empowerment and gender equality. Secondary objectives include understanding effects on outcomes of sustainable peace, human security, and wellbeing.

Given our use of Kabeer's (and Sen's) three-pronged definition of women's empowerment, we developed a taxonomy of outcomes organised by category (Agency, Resources, and Achievement), subcategory, and specific outcome. The subcategories are organised as follows:

- We define Agency as the ability to act upon one's goals, but also be able to bring meaning, motivation and purpose to an activity. We divide the outcomes into the following sub-categories:
 - Individual Agency
 - Community Level Agency
 - Institutions Supporting Agency
- We define Resources, under Kabeer's umbrella, broadly as defined to include not only access, but also future claims, to both material and human and social resources. We divide the outcomes into the following sub-categories:
 - Access to Justice and Legal Services
 - Economic and Livelihoods Related Resources
 - Access to Employment

- We define Achievements as the capabilities that underlie both agency and resources, as a means of facilitating their procurement. We divide the outcomes into the following sub-categories:
 - Improved Systems
 - Norms and Behaviour Change
 - Empowerment Index

The complete list of outcomes, their definitions and associated UNSCR 1325 pillars is found in Online Appendix H.

Duration of follow-up

We have not included or excluded studies based on duration of follow-up. If studies included multiple follow-ups, we have included the outcomes measures most like those presented in the other studies included in any single meta-analysis and report any additional follow-ups narratively.

3.1.6 Types of settings

We include interventions started between 2000 and 2020 conducted in FCAS in L&MICs only. We have noted above that fragility varies over time and within states. In order to identify interventions implemented in fragile situations, we operationalised the following approach. First, to facilitate initial screening, we developed a list of eligible L&MICs that have been classified as fragile according to one of a few state-level lists. For TAS criteria, we included studies where the population falls within at least one of the following categories:

- L&MICs that are listed on all our indicators: in the World Bank list of FCAS (since 2006), on the OECD list of FCAS (since 2007), on the FSI with a score higher than 80 (since 2009) and that have a score of at least 0.6 on the GII for at least one year since 2000
- L&MICs that are listed in one of the fragility lists or had at least one year rated higher than 80 on FSI and that have a score of at least 0.6 on the GII for at least one year since 2000
- L&MICs that do not appear on the fragility lists but are listed with a score lower than 0.6 on the Women Peace Security Index between 2016 and 2019

For FTS criteria, the call for inclusion has been made on a series of criteria depending on the characteristics of the population on the first year of implementation against the following indicators:

- L&MICs status of the country of implementation
- GII score
- List of FCAS according to the FSI, World Bank and OECD
- Nuanced definition of fragile contexts applied in previous work
- Specific local context

An additional point of interest is the emphasis on fragile and conflict-affected *situations*, rather than nation-states. As a result, countries that do not score within our inclusion criteria might still be included in the final list of studies due to a variety of reasons. As indices are often capturing national-level data, as a research team we sought to recognise that sub-nationally, the context may be varying including in situations such as refugee camps in non-L&MICs or rural, unstable regions in non-fragile states.

3.1.7 Other

We included both completed and ongoing studies, such as protocols of ongoing studies that appear to meet all other inclusion criteria or studies listed in registries of ongoing impact evaluations.

We included studies published in any language, although search terms were in English only. We included studies published in 2000 or after, including interventions implemented in 2000 or after, taking the UNSCR 1325 as the starting point of our study.

3.2 Search methods for identification of studies

A comprehensive search of the literature for a systematic review on a topic in international development should cover key bibliographic databases, those specific to international development, those specific to social sciences, and specific to the subject of the review (Kabeer and Waddington 2015). The search strategy has been developed in collaboration with an information specialist and with reference to guidance in Hammerstrøm and colleagues (2010).

In order to capture the relevant literature as comprehensively as possible, we have developed both a general set of search terms and series of sub-strategies and terms grouping around the typologies of the types of interventions included in our review, gender empowerment and gender equality, and building peaceful and inclusive societies. All searches were limited by the list of countries filters and by the year from 2000 onwards and integrated diverse range of literature: articles, impact evaluations, reports, dissertations, conference documents, etc.

3.2.1 Electronic searches

We searched the following academic databases:

- Africa-Wide (Ebsco): <https://www.ebsco.com/products/research-databases/africa-wide-information>
- CAB Global health: <https://www.cabi.org/publishing-products/global-health/>
- Communication and Mass Media: <https://www.ebsco.com/academic-libraries/subjects/communication-mass-media>
- Econlit: <https://www.aeaweb.org/econlit/>
- Embase: <https://www.embase.com/login>
- Gender Studies: <https://www.ebsco.com/products/research-databases/gender-studies-database>
- International Political Science Abstracts: <https://www.ebsco.com/products/research-databases/international-political-science-abstracts>
- Medline: <https://pubmed.ncbi.nlm.nih.gov/>
- PsycInfo: <https://www.apa.org/pubs/databases/psycinfo>
- REpec: <http://www.repec.org/>
- Web of Science (SSCI): <https://clarivate.com/webofsciencegroup/solutions/webofscience-ssci/>
- World Bank e-library: <https://elibrary.worldbank.org/>

3.2.2 Grey literature searches

We searched the following specialist organisational databases:

- CARE International: <http://www.careevaluations.org/>
- Catholic Relief Services: <https://www.crs.org/our-work-overseas/research-publications>
- Centre for Public Impact: <https://www.centreforpublicimpact.org/observatory/>
- Chemonics International: <https://www.chemonics.com/technical-areas/democracy-and-governance/>
- EGAP (Evidence in Governance and Politics): <http://egap.org/biblio>
- IRC: <https://www.rescue.org/reports-and-resources>
- LSE ICG: https://www.theigc.org/search/?select-post_type%5B%5D=publication
- Mercy Corps: <https://www.mercycorps.org/research>
- Oxfam International: <https://policy-practice.oxfam.org.uk/publications>
- Search for Common Ground: <https://www.sfcg.org/ilt/evaluations/>
- FHI360: <https://www.fhi360.org/>
- World Vision: <http://www.wvi.org/resources>

Bilateral and multilateral agencies and general repositories of impact evaluations in international development to be searched include:

- 3ie Repository of Impact Evaluations: <https://developmentevidence.3ieimpact.org/>
- 3ie RIDIE (Registry for International Development Impact Evaluations):: <http://ridie.3ieimpact.org/>
- African Development Bank (AfDB): <https://www.afdb.org/en/documents/publications/>
- AgEcon: <https://ageconsearch.umn.edu/?ln=en>
- AGRIS: <http://agris.fao.org/agris-search/index.do>
- Asian Development Bank (ADB): <https://www.adb.org/publications>
- BREAD: <http://ibread.org/bread/papers>
- Center for Effective Global Action (CEGA): <http://cega.berkeley.edu/evidence/>
- CGIAR: Consultative Group on International Agricultural Research: <https://cgspace.cgiar.org/handle/10568/83389>
- Design, Monitoring and Evaluation for Peace: www.dme4peace.org/learn/resources/
- DEval: <https://www.deval.org/en/home.html>
- DFID Research for Development (R4D): <http://r4d.dfid.gov.uk/>
- GEF (Global Environmental Facility) evaluation database: [http://www.gefio.org/evaluations/all?f\[0\]=field_iew_grouping%3A312](http://www.gefio.org/evaluations/all?f[0]=field_iew_grouping%3A312)
- Global Facility for Disaster Reduction and Recovery: <https://www.gfdr.org/en/publications>
- ICNL Research Centre: <http://www.icnl.org/research/library/ol/>
- IFPRI: <http://www.ifpri.org/publications>
- Independent Development Evaluation, AfDB: <http://idev.afdb.org/en/page/evaluations>
- Innovations for Poverty Action (IPA): <http://www.poverty-action.org/search-studies>
- Inter-American Development Bank Publications: https://publications.iadb.org/facet-view?locale-attribute=enandfield=type_view
- J-Poverty Action Lab (J-PAL): <https://www.povertyactionlab.org/evaluations>
- Locus (International Development Coalition): <https://locus.ngo/resources>

- OECD: <http://www.oecd.org/derec/?hf=5andb=0ands=score>
- UNEG Database of Evaluation Reports: <http://www.uneval.org/evaluation/reports>
- USAID: <https://dec.usaid.gov/dec/search/SearchResults.aspx?q=KERvY3VtZW50cy5CaWJ0eXBIX05hbWU6KCgiU3BIY2lhbCBFdmFsdWF0aW9uulikgT1lgKCJGaW5hbCBFdmFsdWF0aW9uIFJlcG9ydCpKSsk=>

Our methodology for the grey literature search followed a systematic approach:

1. First, a list of the most common variants of the intervention terms should be extrapolated based on any scoping work and existing studies for inclusion identified. Considering the broad aspect of our review we used a short set of terms related to gender equality and women's empowerment.
2. The strategy differs for sites with more or less than 300 resources. First, any relevant filters within the specific site should be applied (e.g., choosing only relevant document types, sectors, etc.).
 - a. For repositories with more than 300 resources:
 - i. Each search term is run alongside the methodology term “impact evaluation” and “impact assessment” for a total of 10 searches per site (women, men, girls, boys and gender).
 - ii. Screen results until no potentially relevant results are identified on two consecutive pages of 50 search results (i.e., 100 excludes in a row).
 - b. For repositories with fewer than 300 resources:
 - i. Only the methodology term “impact evaluation” is searched.
 - ii. Screen all results.
3. For all sites, where possible, display settings are set to show 50 results per page, and the first page of results gets printed to PDF. Each set of PDFs gets saved in a folder with the site name and date of search, and the search data get recorded in a master Excel file as follows:
 - a. For each website searched, the following information should be recorded:
 - i. the date of the search;
 - ii. the specific URL of the website searched (e.g., to differentiate between a search of publications vs a search of programmes, which on J-PAL's website will return different results);
 - iii. the specific search terms used in each search (e.g., 1: impact evaluation; 2: “impact evaluation”; 3... etc.);
 - iv. the number of hits for each search result (which can be recorded in a single cell with a formula, = [no. results from first search] + [no. results from second search] +...+ [no. results from nth search]);
 - v. the number of potentially relevant documents identified and uploaded to EPPI, recorded by search as in (iv)
4. For each set of results, the titles and any abstract information shown on the main page are screened for potential inclusion, and if clearly irrelevant based on intervention, methodology or population, are ignored. If the title indicates possible relevance, click through to the full-text and screen the abstract or executive summary.
5. Any result that appears of potential relevance gets downloaded and saved in the GLS shared folder for that site.
6. The documents collected from the GLS are then added to the project's EPPI environment.

- a. Create a new record in EPPI for each identified study.
- b. Key bibliographic information should be added: Title, all authors, year of publication, URL to the full text, and publisher. Abstract is not required.
- c. Once the bibliographic info is added, click “save,” which will then allow you to upload the full text.
- d. The study should be marked as “include” at TAS and coding marked as completed (with the record open in EPPI, right-click the heading for Title and abstract screening, select “Properties” then check the box next to “Coding completed”)
- e. The study should be marked as grey literature search under “origin”

When information is not directly available on the listed website, we made use of Google Search (especially Google Scholar) for the retrieval of missing and/or additional references.

Given the breadth of our SR we are going to trial web scraping tools at the beginning of the search. We tried the use of a data *miner* to extract grey literature reference that would then be screened through EPPI:

- The data scraping was only tested on platforms presenting both the Title and Abstract of the listed reports and resources
- The same approach as for the standard Grey Literature Method and save a PDF copy of the results of the search term
- Using data miner, we then extracted the data of the results pages to an excel template that would allow the upload of all the results on EPPI for Title and Abstract Screening

We piloted this web scraping approach by testing it on similar websites to the standard method and assess the interest in scaling it up.

3.2.3 Searching other resources

General approach

We screened the bibliographies of included studies and existing reviews for additional eligible studies and conducted forward citation-tracking of included studies in Google Scholar:

- Backwards citation search: Comb through the list of references in the study's full text and see if you can identify any potentially relevant studies from the titles, looking up any that look even possibly relevant to check their abstract.

Forwards citation search: We used Google Scholar to find the included paper, then under the entry, clicked the hyperlink where it says "Cited by X" and screened the subsequent list of papers citing the included study.

This citation search applied to the studies shortlisted through the Title and Abstract screening and Grey Literature search and, considering the potential size of our sample, not to the retrieved literature and citations selected for analysis through the Full Text Screening. Considering the potential size of our sample we only did one level of citation search and did not do rounds of citation searches until no resources can be found.

We also published a blog post presenting our SR and calling for the provision of includable studies from the international development and humanitarian community. We also identified and contacted key researchers and organisations working in the relevant fields of our study for suggestions and engaged our advisory group for suggestions of relevant studies. We also hand-searched journals of relevance to the review to identify

papers that have not yet been indexed, covering issues published in the last 12 months. This hand-search was also based on journals citing the included studies.

Titles and abstracts have been screened against the inclusion criteria and relevant records have been downloaded into the review management software EPPI reviewer. The initial screening of records was conducted by several reviewers screening the records from different databases. At this stage we have been over-inclusive to ensure relevant studies are not omitted because sufficient information is not reported in title or abstract. Two reviewers then independently reviewed abstracts that have been judged to be potentially relevant at the first stage in more detail to determine which papers should be retrieved and reviewed at full text. Two reviewers then independently assessed full text studies for inclusion, with any disagreements resolved by a third reviewer.

Targeted search for addressing review questions 3 and 4

Once we identified our set of included impact evaluations, we undertook targeted searching for qualitative studies, process evaluations, and project documents for those interventions evaluated in the included studies. We conducted citation tracking of included studies to identify relevant sister papers and conduct internet and database searches using the names of programmes from included studies. To identify relevant project documents, process evaluations and other qualitative studies, we conducted targeted searches of Google and Google Scholar as well as the funder and implementer websites of the identified programmes.

3.3 Data collection and analysis

3.3.1 Description of methods used in primary research

Using the inclusion criteria set out in the previous sections, we anticipated that primary studies included in this review would use experimental or quasi-experimental study designs and/or analysis methods to examine the extent to which changes in outcomes can be attributed to the intervention under study. To this end, we included randomised studies as well as non-randomised studies that were able to suitably account for selection and confounding bias (Waddington et al. 2017).

3.3.2 Criteria for determination of independent findings

Complex data structures are a common occurrence in meta-analyses of impact evaluations. There are several scenarios through which these complex structures with dependent effect sizes might occur. For instance, there could be several publications that stem from one study, or several studies based on the same data set. Some studies might have multiple treatment arms that are all compared to a single control group. Other studies may report outcome measurements from several time points or use multiple outcome measures to assess related outcome constructs. All such cases yield a set of statistically dependent effect size estimates (Borenstein et al. 2009).

The research team assessed the extent to which relationships existed across the studies included in the review. We avoided double counting of identical evidence by linking papers prior to data analysis. Where we have several publications reporting on the exact same effect, we used effect sizes from the most recent publication. We utilised information provided in studies to support these assessments, such as sample sizes, programme characteristics and key implementing and/or funding partners.

We extracted effects reported across different outcomes or subgroups within a study, and where information is collected on the same programme for different outcomes at the same or different periods of time, we extracted information on the full range of outcomes over time. Where studies report effects from multiple model specifications, we used the author's preferred model specification. If this is not stated or is unclear, we used the specification with the most controls. When studies reported multiple outcome sub-groups for the same outcome construct, we calculated a "synthetic effect size" (Borenstein et al. 2009, chapter 24). Where studies reported multiple outcomes or evidence according to sub-groups of participants, we recorded and reported data on relevant sub-groups separately. Further information on criteria for determining independent effect sizes is presented below.

We dealt with dependent effect sizes through data processing and selection techniques, utilising several criteria to select one effect estimate per study. When we had several publications reporting on the same study, we used effect sizes from the most recent publication. For studies with outcome measures at different time points, we followed De La Rue and colleagues (2013) and synthesised outcomes measured immediately after the intervention (defined as 1-6 months) and at follow-up (longer than six months) separately. If multiple time points exist within these time periods, we used the most recent measure. We anticipated many of the interventions we include in our review would be ongoing programmes and the follow-up would, therefore, reflect duration in a programme rather than time since intervention. When such studies reported outcome measures at different time points, we identified the most common follow-up period and included the follow up measures that match this most closely in the meta-analysis. When studies included multiple outcome measures to assess related outcome constructs, we followed Macdonald and colleagues (2012) and selected the outcome that appears to most accurately reflect the construct of interest without reference to the results. If studies included multiple treatment arms with only one control group and the treatments represent separate treatment constructs, we calculated the effect size for treatment A versus control and treatment B versus control and included them in separate meta-analyses according to the treatment construct. If treatments A and B represented variations of the same treatment construct, we calculated the weighted mean and standard deviation for treatment A and B before calculating the effect size for the merged group versus control group, following the procedures outlined in Borenstein and colleagues (2009, chapter 25). Where different studies reported on the same programme but use different samples (e.g., from different regions, or separately for men and women) we included both estimates, treating them as independent samples, provided effect sizes are measured relative to separate control or comparison groups.

3.3.3 Selection of studies

We imported all search results into EPPI-Reviewer 4 (Thomas et al. 2010) and removed duplicates through the EPPI-Reviewer 4 deduplication process. We double screened at title and abstract for the first 10 per cent of search results, including any studies we knew would be included, to train the machine learning (ML) algorithm. In this review, we took advantage of an innovative text-mining machine learning capability of EPPI-Reviewer 4 to reduce the initial screening workload: the inclusion/exclusion classifier (O'Mara-Eves et al. 2015; Thomas et al. 2011).

We utilised the inclusion/exclusion classifier to organise studies into groups based on their probability of inclusion in the review. We conducted piloting and verification of the machine learning functioning and based on our experience in previous reviews, we expected to be able to exclude all studies with less than 20 per cent probability of inclusion automatically from the review. We screened a random 10 per cent sample of the automatically excluded studies to double check the accuracy of the function, and if all are excludable, we would auto-exclude the rest. We would then double screen at title and abstract all records with likelihood of inclusion at 20% or greater.

Where a study's title and abstract do not include sufficient information to determine relevance, we included the study for review at full text. We double screened all studies flagged for full-text review using two independent reviewers. We resolved disagreements on inclusion or exclusion by discussion with a core review team member and the input of an additional core reviewer if necessary. We assessed the results of the study-specific key-word searches for relevance, that is, whether they cover one of the programmes included to answer our research questions and whether they provide information on the design, implementation processes, context or mechanisms at play.

We also expected to identify multiple papers related to the same study. In this case we used the Linked studies functionality of EPPI reviewer to identify the main study and other linked studies. The main study would be the study used for data extraction and the linked studies would complement the potential missing information of the main studies. To identify the main study, the priority was given to journal articles, in the case of multiple journal articles or only reports/working papers the most comprehensive was selected. In the case of equivalent quality of the paper, the most recent paper was selected.

3.3.4 Data extraction and management

We extracted the following descriptive, methodological, qualitative and quantitative data from each included study using standardised data extraction forms (see Online Appendix B):

- Descriptive data including authors, publication date and status, as well as other information to characterise the study including country, type of intervention and outcome, population, and context.
- Methodological information on study design, analysis method, and type of comparison (if relevant).
- Quantitative data for outcome measures, including outcome descriptive information, sample size in each intervention group, outcomes means and standard deviations, and test statistics (e.g., t-test, F-test, p-values, 95% confidence intervals).
- Information on intervention design, including how the intervention incorporates participation, inclusion, transparency and accountability characteristics, participant adherence, contextual factors, and programme mechanisms.

We extracted all quantitative, qualitative, descriptive, and methodological data using Excel. Descriptive and qualitative data was single coded by one reviewer and checked by a second reviewer. Two independent reviewers double coded quantitative data for outcomes analysis and risk of bias assessments, and any disagreement was resolved through discussion with a third reviewer (who was a core team member).

Once all effect sizes are calculated and converted to a standardised mean difference (as described in detail below), we examined the data for outliers. We defined outliers as any effect sizes ± 3.29 standard deviations from the mean, following the guidance of Tabachnick and Fidell (2001), and utilized an outlier analysis on the entire dataset during the data cleaning process to identify any potential errors. We examined sensitivity to outliers as discussed in the section on sensitivity analysis below.

For the qualitative analysis, we extracted detailed data on population characteristics, intervention design and implementation, and contextual variables (e.g., region, political climate) to address questions 3 and 4.

3.3.5 Assessment of risk of bias in included studies

We assessed the risk of bias in the included studies by drawing on the signalling questions in the 3ie risk of bias tool, which covers both internal validity and statistical conclusion validity of experimental and quasi-experimental impact evaluation designs (Hombrados and Waddington 2012). It includes the bias domains and extensions to Cochrane's ROBINS-I tool and RoB2.0 (Higgins et al. 2016; Sterne et al. 2016). The risk of bias assessment helps us to determine the extent to which the findings in each study are reliable. Two reviewers undertook the risk of bias assessment independently. If there are disagreements, we resolved them by discussion and the involvement of a third reviewer (who must be a member of the core team), as necessary. The risk of bias tool can be found in Online Appendix C. We did the risk of bias at the paper level.

We assessed risk of bias based on the following criteria, coding each paper as "Yes", "Probably Yes", "Probably No", "No" and "No Information" according to how they address each domain:

- Factors relating to baseline confounding and biases arising from differential selection into and out of the study (e.g., assignment mechanism).
- Factors relating to bias due to missing outcome data (e.g., assessment of attrition).
- Factors relating to biases due to deviations from intended interventions (e.g., performance bias and survey effects) and motivation bias (Hawthorne effects).
- Factors relating to biases in outcomes measurement (e.g., social desirability or courtesy bias, recall bias).
- Factors relating to biases in reporting of analysis.

We reported the results of the assessment for each of the assessed criteria for each study. In addition, we used the results of the risk of bias assessments to produce an overall rating for each study as either "High risk of bias", "Some concerns" or "Low risk of bias", drawing on the decision rules in RoB2.0 (Higgins et al. 2016), rating studies as follows:

- "High risk of bias": if any of the bias domains were assessed as "No" or "Probably No."
- "Some concerns": if one or several domains were assessed as "No Information" and none were "No" or "Probably No."
- "Low risk of bias": if all the bias domains were assessed as "Yes" or "Probably Yes."

In addition, we attempted to explore whether there are systematic differences in outcome effects between primary studies with different risk of bias. If meta-analysis was feasible, we conducted sensitivity analysis to assess the robustness of the results to the risk of bias in included studies.

3.3.6 Measures of treatment effect

An effect size expresses the magnitude (or strength) and direction of the relationship of interest (Valentine et al. 2015; Borenstein et al. 2009). We extracted data from each individual study to calculate standardised effect sizes for cross-study comparison wherever possible. For continuous outcomes comparing group means in a treatment and control group, we calculated the standardised mean difference (SMDs), or Cohen's d , its variance and standard error using formulae provided in Borenstein and colleagues (2009). A SMD is a difference in means between the treatment and control groups divided by the pooled standard deviation of the outcome measure. Cohen's d can be biased in cases where sample sizes are small. Therefore, in all cases we simply adjusted d using Hedges' method, adjusting Cohen's d to Hedges' g using the following formula (Ellis 2010):

$$g \cong d(1 - \frac{3}{4(n_1 + n_2) - 9})$$

We chose the appropriate formulae for effect size calculations in reference to, and dependent upon, the data provided in included studies. For example, for studies reporting means (X) and pooled standard deviation (SD) for treatment (T) and control or comparison (C) at follow up only:

$$d = \frac{x_{Tp+1} - x_{Cp+1}}{SD}$$

If the study does not report the pooled standard deviation, it is possible to calculate it using the following formula:

$$SD_{p+1} = \sqrt{\frac{(n_{Tp+1} - 1)SD_{Tp+1}^2 + (n_{Cp+1} - 1)SD_{Cp+1}^2}{n_{Tp+1} + n_{Cp+1} - 2}}$$

Where the intervention was expected to change the standard deviation of the outcome variable, we used the standard deviation of the control group only.

For studies reporting means (\underline{X}) and standard deviations (SD) for treatment and control or comparison groups at baseline (p) and follow up ($p+1$):

$$d = \frac{\Delta \underline{X}_{p+1} - \Delta \underline{X}_p}{SD_{p+1}}$$

For studies reporting mean differences ($\Delta \underline{X}$) between treatment and control and standard deviation (SD) at follow up ($p+1$):

$$d = \frac{\Delta \underline{X}_{p+1}}{SD_{p+1}} = \frac{\underline{X}_{Tp+1} - \underline{X}_{Cp+1}}{SD_{p+1}}$$

For studies reporting mean differences between treatment and control, standard error (SE) and sample size (n):

$$d = \frac{\Delta X_{p+1}}{SE\sqrt{n}}$$

As primary studies have become increasingly complex, it has become commonplace for authors to extract partial effect sizes (e.g. a regression coefficient adjusted for covariates) in the context of meta-analysis. For studies reporting regression results, we followed the approach suggested by (Keef and Roberts 2004) using the regression coefficient and the pooled standard deviation of the outcome. Where the pooled standard deviation of the outcome is unavailable, we used regression coefficients and standard errors or t-statistics to do the following, where sample size information is available in each group:

$$d = t \sqrt{\frac{1}{n_T} + \frac{1}{n_C}}$$

where n denotes the sample size of treatment group and control. We used the following where only the total sample size information (N) is available, as suggested in (Polanin et al. 2016):

$$d = \frac{2t}{\sqrt{N}} \quad Var_d = \frac{4}{N} + \frac{d^2}{4N}$$

We calculated the t -statistic (t) by dividing the coefficient by the standard error. If the authors only reported confidence intervals and no standard error, we calculated the standard error from the confidence intervals. If the study does not report the standard error, but report t , we extracted and used this as reported by the authors. In cases in which significance levels are reported rather than t or SE (b), then t was imputed as follows:

Prob > 0.1: $t = 0.5$

0.1 ≥ Prob > 0.05: $t = 1.8$

0.05 ≥ Prob > 0.01: $t = 2.4$

0.01 ≥ Prob: $t = 2.8$

Where outcomes are reported in proportions of individuals, we calculated the Cox-transformed log odds ratio effect size (Sánchez-Meca et al. 2003):

$$d = \frac{\ln(OR)}{1.65}$$

where OR is the odds ratio calculated from the two-by-two frequency table.

Where outcomes were reported based on proportions of events or days, we used the standardised proportion difference effect size:

$$d = \frac{p_T - p_C}{SD(p)}$$

where p_t is the proportion in the treatment group and p_c the proportion in the comparison group, and the denominator is given by:

$$SD(p) = \sqrt{p(1-p)}$$

where p is the weighted average of p_c and p_t :

$$p = \frac{n_T p_T + n_C p_C}{n_T + n_C}$$

An independent reviewer evaluated a random selection of 10 per cent of effect sizes to ensure that the correct formulae were employed in effect size calculations. In all cases after synthesis, we converted pooled effect sizes to commonly used metrics such as percentage changes and mean differences in outcome metrics typically used (e.g., weight in kg) whenever feasible.

3.3.6 Unit of analysis issues

Unit of analysis errors can arise when the unit of allocation of a treatment is different to the unit of analysis of effect size estimate, and this is not accounted for in the analysis (e.g., by clustering standard errors at the level of allocation). We assessed studies for unit of analysis errors (The Campbell Collaboration 2019), and where they exist, we corrected for them by adjusting the standard errors according to the following formula (Higgins et al. 2020; Waddington et al. 2012; Hedges 2009):

$$SE(d)' = SE(d) * \sqrt{1 + (m - 1)c}$$

where m is the average number of observations per cluster and c is the intra-cluster correlation coefficient. Where included studies used robust Huber-White standard errors to correct for clustering, we calculated the standard error of d by dividing d by the t-statistic on the coefficient of interest.

3.3.7 Dealing with missing data

In cases of relevant missing or incomplete data in studies identified for inclusion, we contacted study authors to obtain the required information. If we were unable to obtain the necessary data, we reported the characteristics of the study but state that it could not be included in the meta-analysis or reporting of effect sizes due to missing data.

3.3.8 Assessment of heterogeneity

We assessed heterogeneity by calculating the Q-statistic, I^2 , and Tau^2 to provide an estimate of the amount of variability in the distribution of the true effect sizes (Borenstein et al. 2009). We complemented this with an assessment of heterogeneity of effect sizes graphically using forest plots. Additionally, we explored heterogeneity using moderator analysis in meta-regression specifications where there were at least four studies and significant heterogeneity. While some have suggested 10 studies as a minimum for moderator analysis, as Borenstein and colleagues (2009) note, there are no hard and fast rules. However, we ensured that for categorical moderators, there were a minimum of two effects per cell.

3.3.9 Assessment of reporting biases

To reduce the possibility of publication bias, we searched for and included unpublished studies in the review. We also tested for the presence of publication bias using funnel plots and statistical tests. Specifically, the rank correlation test (Begg & Mazumdar, 1994) and the regression test (Sterne & Egger, 2005), using the standard error of the observed outcomes as predictor, were used to check for funnel plot asymmetry.

3.3.10 Data synthesis

We conducted meta-analyses of studies that we assess to be sufficiently similar. The inclusion criteria for the review were broad and we anticipated including studies that report on a diverse set of interventions, sectors and outcomes. It was therefore difficult to predict how meta-analysis would be used in the review prospectively. However, the minimum criteria were to only combine studies using meta-analysis when we identify two or more effect sizes using a similar outcome construct and where the comparison group state is judged to be similar across the two, similar to the approach taken by (Wilson et al. 2011). We combined studies in the same analysis when they evaluate the same intervention type, or the same outcome type. Moderator analyses can take into account multiple interventions as moderator variables, allowing us to also examine the impact of different intervention types by outcome. Where there were too few studies, or included studies are considered too heterogeneous in terms of interventions or outcomes, we presented a discussion of individual effect sizes along the causal chain. As heterogeneity exists in theory due to the variety of interventions and contexts included, we used inverse-variance weighted, random effects meta-analytic models (Higgins et al. 2020).

We used the metafor package (version 2.4.0; Viechtbauer 2010) in R software to conduct the meta-analyses (version 4.0.4; R Core Team 2020). The amount of heterogeneity (i.e., τ^2), was estimated using the DerSimonian-Laird estimator (DerSimonian & Laird, 1986). In addition to the estimate of τ^2 , the Q -test for heterogeneity (Cochran, 1954) and the I^2 statistic (Higgins & Thompson, 2002) are reported.

We conducted separate analyses for the major outcome categories for each intervention group: primary outcome (women's empowerment, gender equality), secondary outcome (well-being, security, peace) and other important outcomes such as changes in knowledge (of rights), attitudes and behaviours, tolerance of violence, self-confidence, self-efficacy. Based on an analysis of the interventions that we find, we attempted to further elaborate on the pathway of change that was outlined above to the extent possible. We also used sub-group analysis to explore heterogeneity by different treatment sub-groups (described in more detail in the section on subgroup analysis and investigation of heterogeneity).

We also collected qualitative information from studies about the interventions. This information may subsequently be coded quantitatively to be used in moderator analysis. It may also be used to classify intervention mechanisms in synthesis or in the further development of intervention causal chains. These characteristics may include: intervention objectives (to change processes, behaviours or both); whether interventions are strategic (complex, adaptable strategy to realise change) or tactical (tool-based); the source of intervention (local, NGO, government or researcher-led); the scale of the intervention (pilot experiment versus adoption of formal policy/law); extent to which members of both targeted groups are engaged (equally or primarily one group); and initial power differences between the groups targeted.

3.3.11 Subgroup analysis and investigation of heterogeneity

Whenever feasible, we conducted moderator analyses to investigate sources of heterogeneity. Following the PROGRESS-PLUS approach (Oliver et al. 2017), we assessed moderators falling into three broad categories of extrinsic, methodological and substantive characteristics to address inequity aspects within the gender equality context. Examples of these categories include:

- Extrinsic characteristics: funder of the study (e.g., NGO vs private sector vs government investments), publication type, publication date.
- Methodological characteristics: study design, risk of bias, study quality characteristics, evaluation period, length of follow-up.
- Substantive characteristics: participant characteristics (gender, age, socio-economic status, education, land ownership), context (geographical setting, market access), intervention type, intervention features, type of implementing agency.

We used random effects meta-regression to investigate the association between moderator variables and heterogeneity of treatment effects (Borenstein et al. 2009) and sub-group analyses to investigate heterogeneity by treatment sub-groups (e.g., men and women, poor and non-poor, and so on). If the latter strategies were not possible (that is, if we do not have enough studies or data), we discussed and explored the factors which may be driving heterogeneity of results narratively by conducting cross-case comparisons (Miles and Huberman 1994).

The qualitative analysis also included the following subgroups: women's age, education, marital status, health (currently pregnant or not), ethnically marginalised group, race and caste. For groups, heterogeneity by mixed-sex groups and women-only groups. If possible, sex of service/intervention frontline provider.

3.3.12 Sensitivity analysis

We conducted sensitivity analysis to assess whether the results of the meta-analysis were sensitive to the removal of any single study. We did this by removing studies from the meta-analysis one-by one and assessing changes in results. We also assessed sensitivity of results to inclusion of high risk of bias studies by removing these studies from the meta-analysis and comparing results to the main meta-analysis results. Studentised residuals and Cook's distances are used to examine whether studies may be outliers and/or influential in the context of the model (Viechtbauer & Cheung, 2010). Studentised residuals express the difference between the predicted effect size (based on the entire body of evidence in the analysis) and the observed effect size for any given study. These are a standard diagnostic tool for outliers in meta-analysis. Studies with a studentised residual larger than the $100 \times (1 - 0.05/(2 \times k))$ percentile of a standard normal distribution are considered potential outliers (i.e., using a Bonferroni correction with two-sided $\alpha = 0.05$ for k studies included in the meta-analysis). Studies with a Cook's distance larger than the median plus six times the interquartile range of the Cook's distances are influential.

3.3.13 Treatment of qualitative research

As a mixed-methods systematic review, we included a distinct review component to synthesise qualitative evidence on the review questions 3 and 4. While the identification of qualitative evidence is limited to studies linked to the included impact evaluations, the

process of data extraction, critical appraisal, and evidence synthesis is independent and follows methods and guidelines tailored to the conduct of qualitative evidence synthesis (Noyes et al. 2020). This approach to mixed-methods synthesis of development interventions builds on Snilstveit and colleagues 2015 and 2018.

Assessment of quality in descriptive quantitative studies, qualitative studies and process evaluations

We assessed the quality of included qualitative studies, process evaluations, and descriptive quantitative studies using a mixed-methods appraisal tool developed by Langer and colleagues (2017) and applied in Snilstveit and colleagues (2019). This tool builds on the Critical Appraisal Skills Programme checklist (CASP 2006) and Pluye and colleagues' (2011) mixed-methods appraisal tool and is provided in Online Appendix A. Our appraisal tool made judgments on the adequacy of reporting, data collection, presentation, analysis and conclusions drawn. The appraisal assessed the quality of the included qualitative studies and descriptive quantitative studies using six appraisal domains:

1. The defensibility of the applied research design to answer the research question under investigation.
2. The defensibility of the selected research sample and the process of selecting research participants.
3. The rigour of the technical research conduct, including the transparency of reporting.
4. The rigour of the applied analysis and credibility of study's claims given the nature of the presented data.
5. The consideration of the study's context (for qualitative studies only).
6. The reflexivity of the reported research (for qualitative studies only).

We filtered out studies of particularly low quality at this stage, using a fatal flaw approach following Dixon-Woods and colleagues (2005). Studies that did not meet either criterion of appraisal domains 1–4 above were excluded from the synthesis. That is, they were included in the review, and we reported on the studies' descriptive data, for example, applied intervention. However, no research findings were extracted from these studies to feed into the review's synthesis. Each appraisal domain was assessed from a scale of critical trustworthiness to low, medium and high trustworthiness. An overall critical appraisal judgement per study was allocated using a numerical threshold of the appraised quality domains (Online Appendix A).

We did not undertake a critical appraisal of included project documents. They typically provide information about planned, ongoing or completed programmes, providing information about the design or resources available for a project for instance. As such these documents did not typically include much analysis of primary evidence, but they provided information about interventions. The purpose of including them in our review was to ensure we had sufficient information about the context and interventions included. We therefore focused the appraisal on assessing the relevance of the documents against the interventions assessed in our review. Before extracting any data, we ensured that the name of the intervention, the implementing agency, context and timeline of the intervention described in the project document corresponds to the intervention assessed in the impact evaluation included in our review. Finally, collecting data from a range of sources, especially if used for triangulation, can enhance confidence in the trustworthiness of the information included. If several sources were available, we extracted data from all sources for purposes of triangulation.

Methods for qualitative evidence synthesis (applicable to review questions 3 and 4)

To address questions 3 and 4, we complemented any statistical meta-regressions with a qualitative evidence synthesis (Noyes et al. 2020). After having completed the detailed coding of all of the included studies as described above, we assessed the coding of data on factors related context, intervention design and implementation, and population characteristics for the most relevant qualitative synthesis approach. The identified approach to qualitative evidence synthesis aligns and is informed by the review's overall theory of change as outlined in Figure 2. We expected thematic synthesis (Thomas and Harden 2008) and qualitative comparative analysis (Rihoux 2006; Thomas et al. 2014) to present two relevant approaches to the qualitative synthesis in this review.⁴ Following a review of the included qualitative evidence base, we opted for thematic synthesis as the most suitable synthesis method.

Thematic synthesis

This qualitative evidence synthesis approach was applied as we had sufficient in-depth qualitative studies and empirical primary data reported across the identified evidence-base and linked to groups of interventions and outcomes along the review's theory of change. Its objective was to identify analytical themes on intervention mechanisms and contexts that mitigate or reinforce intervention effects to complement the review's statistical moderator analysis and/or meta-regression. Following Thomas and Harden's (Thomas and Harden 2008) thematic synthesis, we used inductive coding techniques to first identify common descriptive themes based on the reported findings of the primary studies. We used EPPI-Reviewer's coding software to illustrate the link between the inductive codes in the primary studies and the identified descriptive themes. In a second step, following the identification of descriptive themes, these then were configured into higher level analytical themes, which present the results of the thematic synthesis. Again, this configuration from descriptive to analytical themes was conducted in EPPI-Reviewer and we produced an overview table of both types of themes and their linkages for transparency in this final synthesis step. The process of configuring descriptive and analytical themes from the inductive coding applied the same consistency checks as the general data extraction process outlined above.

The process of generating inductive codes, descriptive themes, and final analytical themes was configured around four analytical lenses derived from the research question 3 and 4 of this review. These refer to the interplay of context, intervention design, intervention implementation, and population characteristics with programme effects, and are outlined in more detail below:

- I. **Context:** Any variable related to external factors beyond the programme's control that affect programme impact. This can refer to political factors such as types of governance, societal factors such as norms, economic factors such as a recession, and cultural factors such as beliefs.
- II. **Intervention design:** any variable that is related to the design and planning of the applied intervention. Design and planning of an intervention refers to the blueprint or schedule of the intervention and will typically outline what components the intervention consists of and in what sequence they will be

⁴ This did not rule out other synthesis approaches such as framework synthesis (Brunton et al 2020).

applied. Examples of design variables refer to: size or type of cash transfer; outreach strategy, posters; reminders; type of training.

- III. **Intervention implementation:** any variable that is related to the implementation of the intervention in practice. This refers to variables that emerge while the intervention is applied and are usually not known in advance. Examples of implementation variables refer to the lack of attendance or uptake, payment difficulties, corruption, elite capture.
- IV. **Population characteristics:** any variable related to the population targeted by the intervention or the population in which the effects are measured (in cases where these differ). This can refer to the socio-economic status of the population, its educational status, and asset ownership.
- V. **Interplay with programme effect:** extracted data and codes need to relate to the programme effect, outcome, or impact. That is, we did not extract descriptive data on the intervention design, implementation, context, and population—we were only interested in data that reports on how variables in these four categories are affecting programme effects.

3.3.14 Confidence in cumulative bodies of evidence

To summarise our confidence in the cumulative body of evidence, we used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach for confidence in quantitative bodies of evidence on intervention effects (Guyatt et al., 2008), and we explored use of the GRADE-CERQual (Confidence in the Evidence from Reviews of Qualitative research) approach for confidence in qualitative bodies of evidence on barriers, facilitators, and moderators of intervention effectiveness (Lewin et al. 2018).

GRADE approach for quantitative bodies of evidence

We used the GRADE approach to rate our confidence in the body of evidence for each intervention contrasts on all primary and secondary outcomes. We specifically rated our confidence that the true effect of an intervention (versus a comparator) on a given outcome lies on one side of the line of no effect, or “difference from the null” (Hultcrantz et al. 2017). Accordingly, our ratings referred to our confidence in the existence (or not) of intervention effects and the direction of effects; they not referred to our confidence in the magnitude of effects. There are four possible levels of confidence:

- High confidence: It is highly likely that the intervention does (not) have an effect on the outcome of interest.
- Moderate confidence: It is likely that the intervention does (not) have an effect on the outcome of interest.
- Low confidence: It is possible that the intervention does (not) have an effect on the outcome of interest.
- Very low confidence: It is not clear whether the intervention does (not) have an effect on the outcome of interest.

We followed the traditional GRADE approach, in which evidence from RCTs starts at high confidence, while evidence from all other study designs starts as low confidence. We assessed the body of evidence according to four factors that can lead to a downgrading confidence in the evidence by one level (serious concern) or two levels (very serious): limitations in individual studies (risks of bias), inconsistency of results, imprecision, and publication bias. We did not downgrade based on indirectness of the evidence, the reasons for which are addressed thoroughly in the *Limitations Related to*

the Use of GRADE section. We then assessed the body of evidence according to three factors that can lead to an upgrading confidence in the evidence: large magnitude of an effect, a dose-response gradient, and the effect of plausible residual confounding. We followed recent guidance on how to consider these factors when adopting a complexity perspective in a systematic review on intervention effects (Montgomery et al. 2019), such as using sources of intervention complexity to explain substantial heterogeneity and using information on intervention implementation to inform assessments of a dose-response gradient.

We presented and explained the results of our GRADE ratings in summary of findings tables (Santesso et al. 2016). In addition, we described our GRADE ratings narratively using language approved by the GRADE Working Group (Santesso et al. 2020):

- High confidence: The intervention reduces/increases the outcome.
- Moderate confidence: The intervention probably reduces/increases the outcome.
- Low confidence: The intervention may reduce/increase the outcome.
- Very low confidence: We are very uncertain about the effect of the intervention on the outcome.

Finally, researchers ranked the importance of different bodies of evidence by applying a framework inspired Maslow's hierarchy of needs to different outcome types as follows:

- Critical: All outcome groups that are directly related to the physical wellbeing of participants, including those related to security, IPV, protection, child-marriage, etc.
- Some importance, but not critical: All outcome groups related to basic needs, including economic security and livelihoods.
- Limited importance: All other measures of empowerment including capacity, autonomy, attitudes, political representation, etc.

Confidence in the body of evidence included in the qualitative review component

We explored the relevance and feasibility of conducting a confidence in the body of evidence assessment for the qualitative review component. The decision to conduct this assessment depended on the size and nature of the included qualitative evidence-base as well as the contribution of the qualitative review findings to the stated research questions. If relevant and feasible, we aimed to apply GRADE-CERQual (Lewin et al. 2018) to conduct the confidence of the evidence assessment. However, given the narrow contribution of the qualitative review component, which is an explanatory supplement to the quantitative findings, we opted against conducting a confidence of the evidence assessment for the qualitative component. As the qualitative review does not search for an independent body of qualitative evidence, and only focused on qualitative evidence linked to the included impact evaluation, its external validity is reduced by design limiting the contribution of the CERQual application.

4. Results

4.1 Description of studies: search results and characteristics of the evidence base

In this chapter, we report descriptive results for the review to provide an overview of the characteristics and distribution of the evidence base across all interventions covered in the review. We start by providing the results of the search and screening of the literature followed by a section providing a summary of the characteristics of included studies.

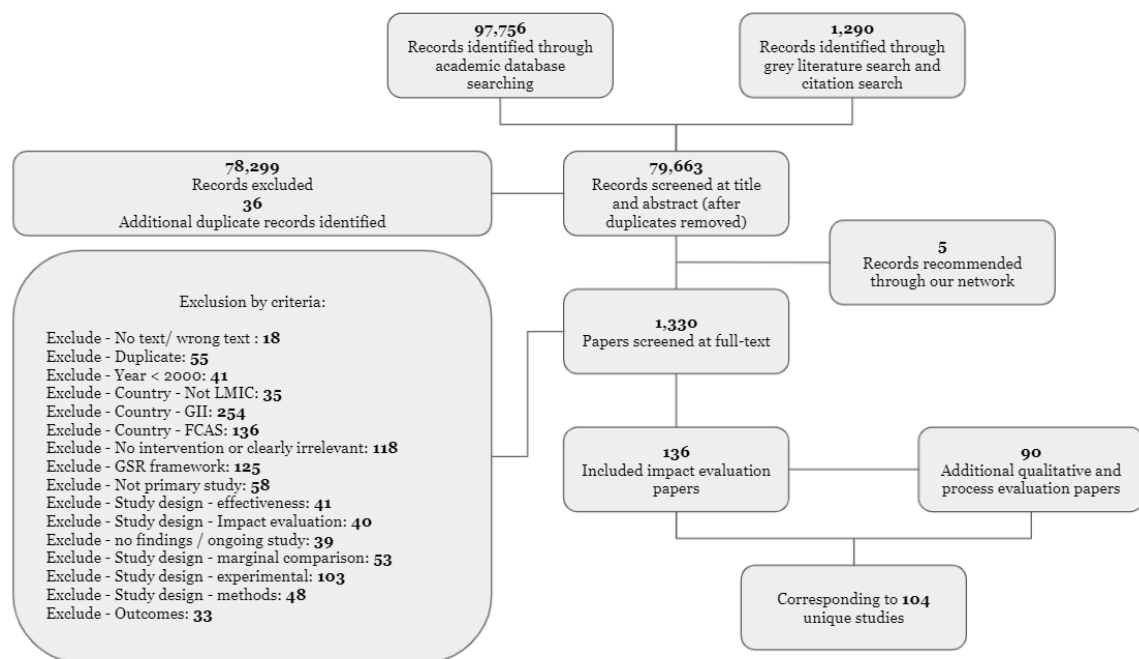
4.1.1 Results of the search

Search process

The PRISMA flow diagram below (Figure 5) details the results of the search and screening process, following guidance from reporting in systematic reviews (Page et al. 2021). The initial academic search identified 97,756 studies. Hand searches of relevant grey literature, citation searches, journal searches and additional suggestions brought 1,290 results. Following the removal of duplicates, a total of 79,663 studies were left for screening at title and abstract. Title and abstract screening identified 1,330 potentially relevant records for inclusion in the review, which were then screened at full text. The full-text screening ultimately identified 136 impact evaluation papers, corresponding to 104 studies of 55 programmes for inclusion in this review. In addition, 39 ongoing studies, or without published findings, were also identified that appear likely to meet inclusion criteria. Although not included in the review, these are listed in the Online Appendix J. There are also an additional 90 papers with complementary qualitative components that underwent qualitative extraction.

The number of included papers is larger than the number of included studies. This is because it is common for studies to be reported in more than one paper, typically one or more working papers and a journal version. As noted in our methodology, our approach was to make the most recent version of the journal article the main paper, and then include any other version(s) of the paper which contained additional information. Typically, working papers and other “unpublished” reports included more detail than journal versions, such as effects on additional outcomes, sub-groups or more details about the programme and would therefore be included.

Figure 5: PRISMA flow diagram of systematic search and screening results



Reasons for exclusion

The main reasons for exclusion at Title and Abstract (TA) stage were:

- The country of focus of the study was not included;
- The study had no intervention or was not relevant for our study: meaning that the intervention was not focusing on gender or, if focusing on gender, was not gender specific or gender transformative; or
- The study was not included through our study framework: the focus was outside of our ecological level or on a non-includable intervention type.

Our flow chart (Figure 5) provides a detailed breakdown of the reasons for exclusion at full text. The most common reason for excluding studies at full text was that they did not our criteria for country inclusion either by not scoring enough (or having a highly relevant context) on GII (n = 254) or by not being a FCAS (or having a highly relevant context) on the first year of intervention (n = 136). The exclusion on the ground of no intervention or clearly irrelevant (n = 116) or GSR framework (n = 125) was still high because some cases of exclusion from TA required more details from the full text to decide. The absence of a valid counterfactual was the main reason for exclusion by study design. In addition to full-text exclusion, 18 texts were not able to be retrieved after three rounds of full-text retrieval. Fifty-five duplicates were identified within the included and excluded studies and removed. The list of all studies excluded at full text with reasons for exclusion is available in Online Appendix J.

However, many studies that could have otherwise been included in this review were excluded based on study design. With quantitative methods, we looked for rigorous methods that included a counterfactual. Reporting on baseline was not present in many studies; without this data it was difficult to ascertain changes between baseline and the reporting of results. Overall, while this review is significantly larger than many existing systematic reviews on this topic, the breadth of interventions were concentrated in

specific interventions and outcomes, with other interventions lacking consistent or sufficient data to be used in quantitative analysis.

Inclusion of process evaluations and qualitative studies

The qualitative study identification process involved three stages of work namely searching, screening, and critical appraisal. The search and screening for qualitative studies as indicated in the methods sections yielded 420 relevant results that meet the inclusion criteria for qualitative studies. The conducted searches were related to each unique programme identified in the included impact evaluations, though multiple studies may have covered the same programme. The reasons for exclusion used for the screening process were limited to document type, e.g., excluding project PowerPoint Presentations, newspaper articles, budgetary reports and other informal reports.

The 420 identified results referred to qualitative studies, descriptive quantitative studies, process evaluations, and projects documents. Project documents did not undergo the critical appraisal process. Typically, project documents provide information about planned, ongoing or completed programmes as well as, for instance, details about the design or resources available for a project. Consequently, project documents did not contain the relevant analysis of primary evidence, but they provided information about interventions. They were retained in the review to ensure there was sufficient information about the context and interventions included and consulted as background information in the qualitative synthesis. In total 109 project documents were identified and excluded from the critical appraisal process.

This left a total of 311 studies for critically appraisal. Following the critical appraisal of these studies, a total of 90 qualitative, descriptive quantitative and process evaluation studies were included for the data extraction process (see section 4.7.2. for more detail).

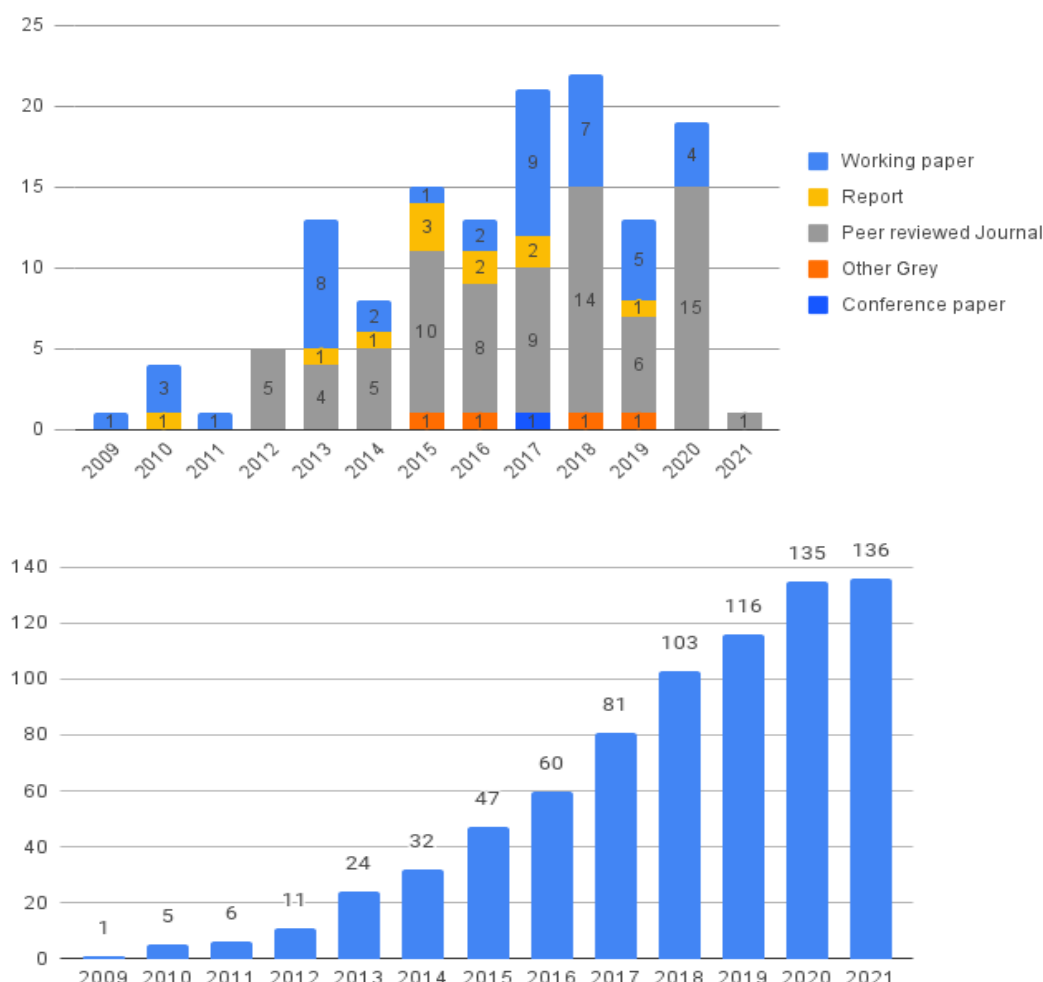
4.1.2 Characteristics of included studies

General characteristics of included studies

The search and screening process resulted in 104 unique studies. A single evaluation study was often reported in more than one report. In this case, the reviewed 104 studies were reported in 136 evaluation reports, inclusive of 32 linked publications.

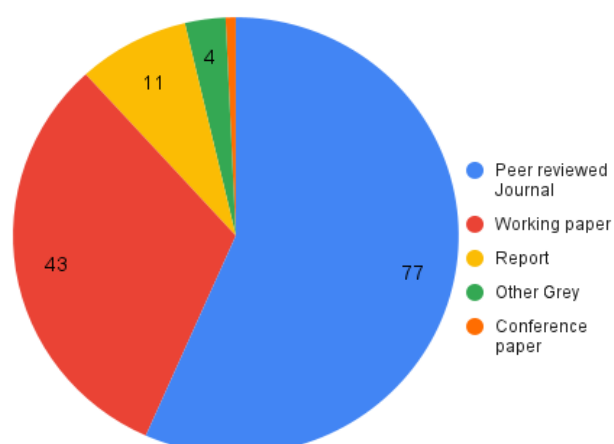
In line with this review's specific focus on interventions that commenced on or after 2000, the earliest included publication was Legovini (2005) and the most recent was Leight (2021). The annual number of publications saw a steady increase over the years with 89 (65% of the total) published between 2016 and 2021. 2018 ranked the highest, with 22 publications. Figure 6 panel A presents the number of publications, classified by types, while panel B presents the cumulative number of publications over the period under consideration.

Figure 6: Annual (panel A, top) and cumulative (panel B, bottom) number of included publications over the review period from 2000 to 202



Of the 136 included publications, 77 were published in peer-reviewed journals. Employing systematic search for grey literature and citation searches of the included publications, this review also includes 43 working papers and 11 organisational reports. Other publications included intervention implementation documents, conference papers and book or book chapters. Figure 7 presents the types of included publications.

Figure 7: Types of included publications



Geographies of included studies

The included unique studies retained after the screening process were geographically diverse (Figures 9 and 10) and representative of several fragile contexts around the globe. Most of the studies originated in Southeast Asia and Sub-Saharan Africa, with the highest number of studies coming from India. Other regions represented include East Asia and the Pacific, Latin America and Caribbean and the Middle East and North Africa.

Figure 8: Map of countries of included studies

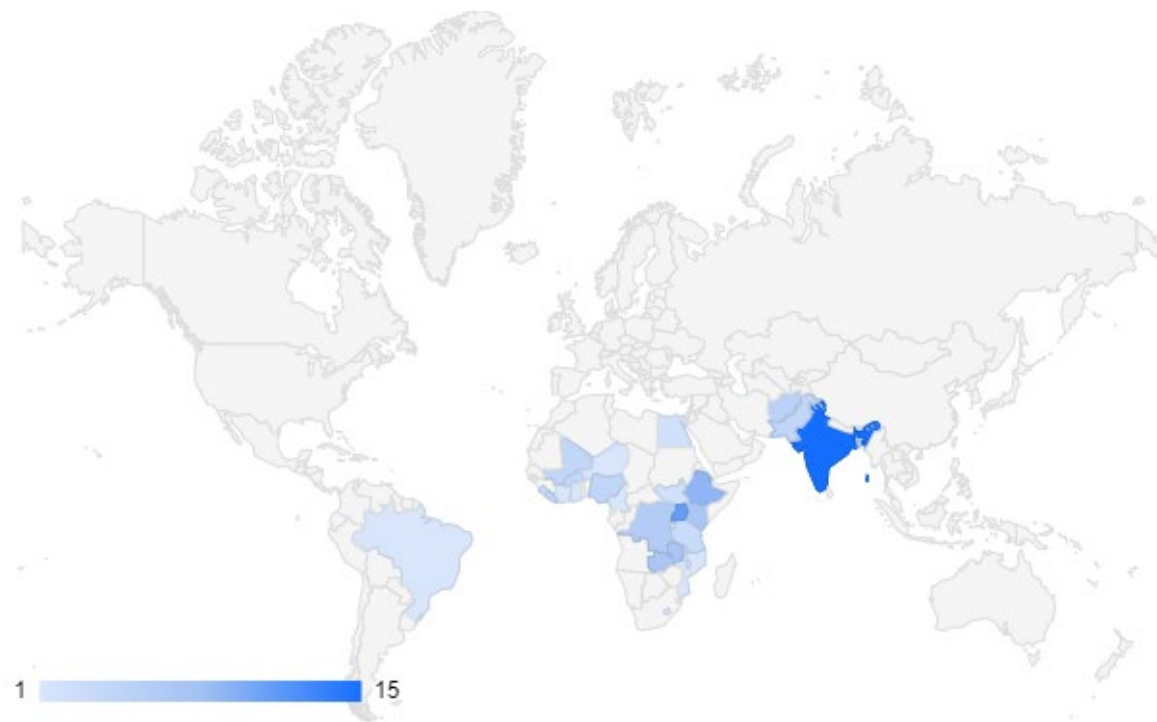
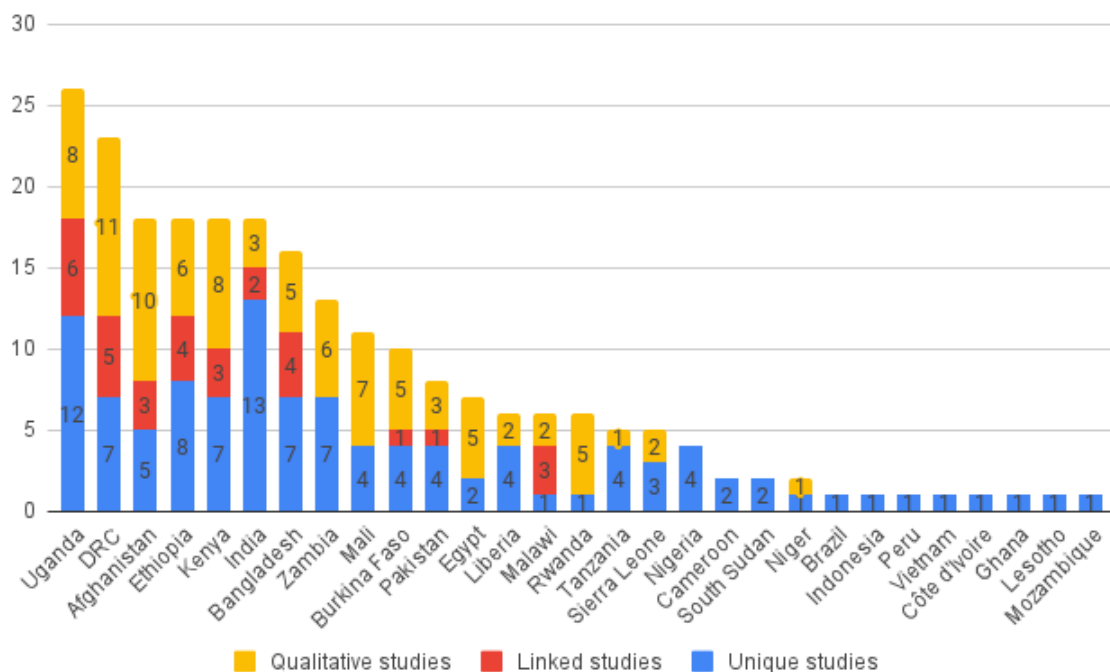


Figure 9: Studies per country



The inclusion criteria for our review specified lower and lower-middle income countries only. Given this, we saw 75 studies from low-income countries and 35 studies from lower-middle income countries, with some studies occurring across multiple countries.

Figure 10: Study FSI scores

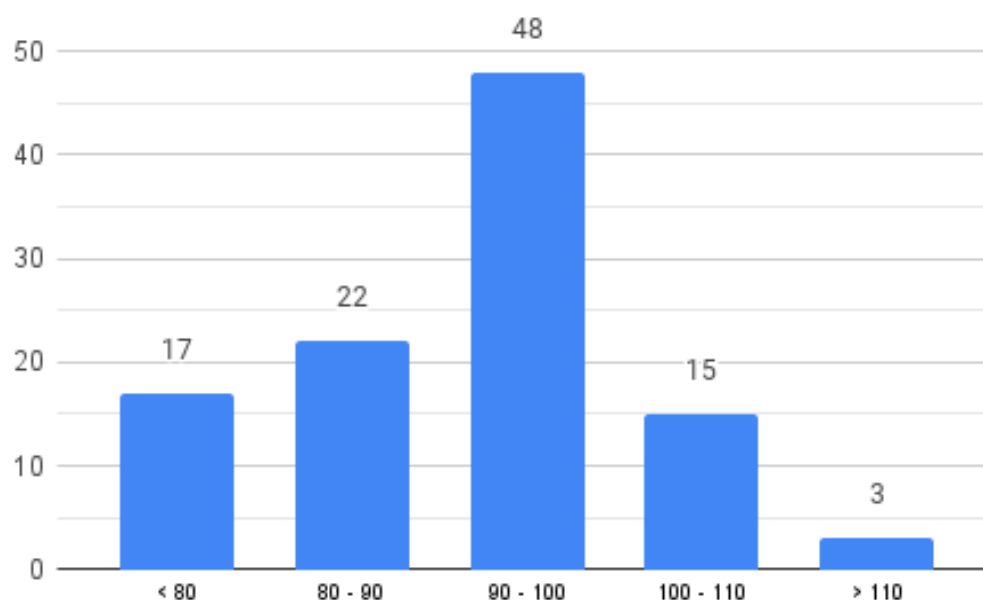
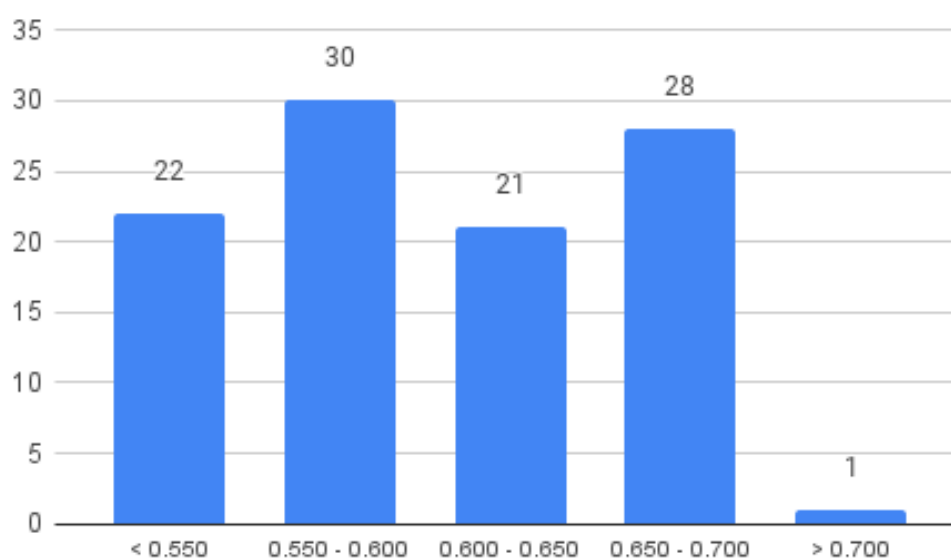


Figure 11: Study GII scores



Additionally, the studies were situated in a variety of contexts that differ in terms of gender equity, fragility and human security. We evaluated the distribution of study contexts on the included dimensions. All but five included studies scored above 0.5 on the Gender Inequality Index (GI), with 50 per cent scoring at or above 0.6 (Figure 11). On the Fragile States Index (FSI), the highest number of studies ($n = 47$) scored between 90 and 100 (Figure 10). Twenty-two studies were classified as fragile and conflict-affected states and situations (FCAS) by the World Bank while 65 studies were classified as the same by the OECD. Lastly, out of the 11 studies for which we were able to gather Women, Peace and Security (WPS) Index data, five studies scored below 0.6.

Types of interventions of included studies

In this section, we describe the interventions, including their components, duration and target populations. We also discuss the ways in which the interventions aimed to build gender equality and women's empowerment towards peaceful and inclusive societies.

While we began the extraction process with 27 intervention groups, the scope of literature quickly revealed that there was significant overlap. After several sessions of brainstorming and discussion, the team reduced it to 14 groups. Table 13 below provides an overview of the intervention categories and included studies:

Table 13: Summary of included studies per intervention type⁵

Intervention group	Included studies - country
Participation pillar	
Asset transfer (n = 14)	<ul style="list-style-type: none"> • Ahmed (2009) - Bangladesh • Bandiera et. al (2017) - Bangladesh • Bedoya (2019) - Afghanistan • Bedoya et. al (2020) - Afghanistan • Glass et. al (2017) - Democratic Republic of the Congo • Nkonya et. al (2012) - Nigeria • Olney et. al (2016) - Burkina Faso • Pradhan (2017) - Bangladesh • Van den Bold, et. al (2015) - Burkina Faso • Roy (2019) - Bangladesh • Bandiera et. al (2013) - Bangladesh • Das (2013) - Bangladesh • Roy (2015) - Bangladesh • Roy (2013) - Bangladesh
Cash transfer (n = 30)	<ul style="list-style-type: none"> • Ambler (2017) - Pakistan • Baird (2010) - Malawi • Blattman (2016) - Uganda • Bonilla (2017) - Zambia • Breisinger (2018) - Egypt • Brooks (2020) - Kenya • Clark (2019) - Kenya • Eze (2019) - Cameroon • Gelagay (2019) - Ethiopia • Gobin (2017) - Kenya • Green (2015) - Uganda • Handa (2015) - Kenya • Haushofer (2018) - Kenya • Heath (2020) - Mali • Iqbal (2020) - Pakistan

⁵ *Italics* are linked studies.

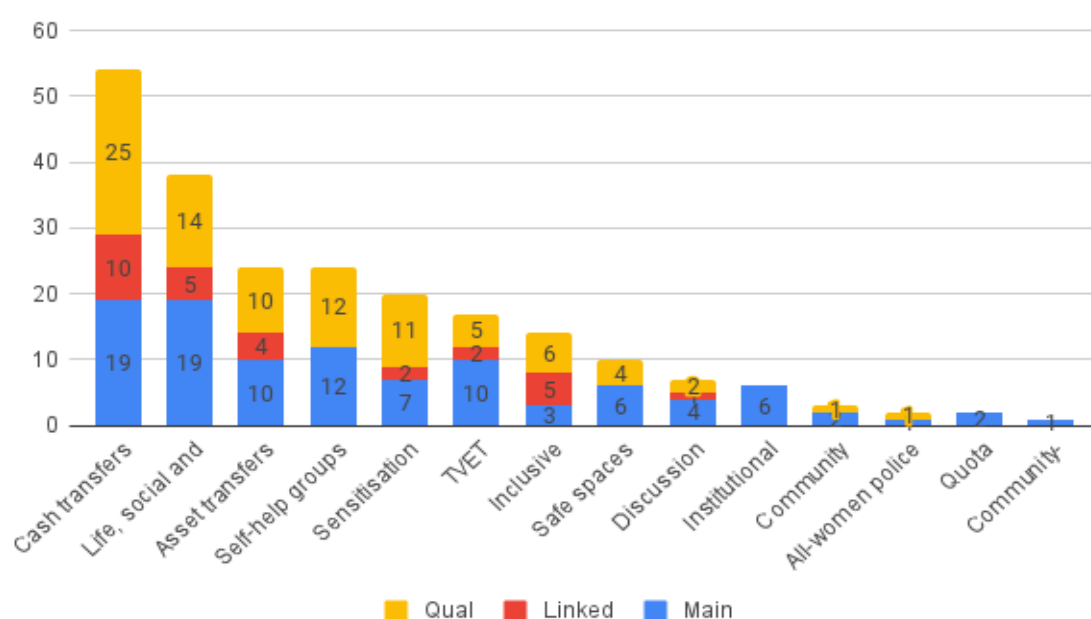
Intervention group	Included studies - country
	<ul style="list-style-type: none"> • Karimli (2021) - Burkina Faso • Malaeb (2017) - Uganda • Mekonnen (2017) - Ethiopia • Müller (2019) - South Sudan • Natali (2016) - Zambia • Baird (2019) - Malawi • Baird (2009) - Malawi • Baird (2013) - Malawi • Blattman, (2013) - Uganda • Green (2016) - Uganda • Haushofer (2019) - Kenya • Haushofer (2013) - Kenya • Haushofer (2016) - Kenya • Iqbal (2020) - Pakistan • Karimli (2019) - Burkina Faso
Community-based services (n = 1)	<ul style="list-style-type: none"> • Nandi et. al (2020) - India
Inclusive community-driven development (n = 7)	<ul style="list-style-type: none"> • Beath (2013) - Afghanistan • Beath (2015) - Afghanistan • Larson (2018) - Brazil, Cameroon, Vietnam • Laudati et. al (2018) - DRC • Beath et. al (2008) - Afghanistan • Van der Windt et. al (2018) - DRC • Mvukiyehe and van der Windt (2020) - DRC
Institutional provision of loans and savings (n = 8)	<ul style="list-style-type: none"> • Tarozzi (2013) - Ethiopia • Tarozzi (2015) - Ethiopia • Duflo (2013) - India • Banerjee (2015) - India • Ifelunini (2012) - Nigeria • Ikenwilo (2016) - Nigeria • Johnson (2018) - Liberia • Weber (2014) - Pakistan
Quotas (n = 2)	<ul style="list-style-type: none"> • Beaman et. al (2012) - India • Clayton (2014) - Lesotho
Self-help groups and VSLAs (n = 13)	<ul style="list-style-type: none"> • Alemu et. al (2018) - Ethiopia • Bass (2016) - Democratic Republic of the Congo • Beaman (2014) - Mali • De Hoop (2010) - India • De Hoop (2014) - India • Deininger (2013) - India • Desai and Joshi (2013) - India

Intervention group	Included studies - country
	<ul style="list-style-type: none"> • Ismayilova et. al (2017) - Burkina Faso • Karlan (2017) - Ghana, Malawi, Uganda • Kundu (2011) - India • Lombardini and Yoshikawa (2015) - Uganda • Vigneri and Lombardini (2016) - Mali • Yaron (2018) - India
Technical and Vocational Education and Training (TVET) (n = 12)	<ul style="list-style-type: none"> • Adoho (2014) - Liberia • Amin (2016) - Bangladesh • Bandiera (2012) - Uganda • Bandiera, et. al (2018) - Uganda • Croke (2017) - Nigeria • Gibbs (2020) - Afghanistan • Lecoutere (2018) - Uganda • Lombardini and Bowman (2015) - Pakistan • Maitra (2017) - India • Mckenzie et. al (2019) - Kenya • Bandiera (2018) - Sierra Leone • Gibbs (2018) - Afghanistan
Prevention pillar	
Community dialogues and reconciliations (n = 2)	<ul style="list-style-type: none"> • Cilliers et. al (2018) - Sierra Leone • Figueroa (2016) - Mozambique
Protection pillar	
All-women police stations (n = 1)	<ul style="list-style-type: none"> • Jassal (2020) - India
Sensitisation campaigns (n = 9)	<ul style="list-style-type: none"> • Abramsky (2014) - Uganda • Das et. al (2012) - India • Decker et. al (2020) - Malawi • Dunkle et. al (2020) - Rwanda • Green (2020) - Uganda • Quisumbing (2020) - Bangladesh • Schensul et. al (2015) - India • Green (2018) - Uganda • Watts (2015) - Uganda
Multi-pillar	
Discussion groups (n = 5)	<ul style="list-style-type: none"> • Hossain (2014) - Cote d'Ivoire • Lubega (2017) - Uganda • Mvukiyehe, (2017) - Liberia • Vaillant (2020) - DRC • Vaillant (2016) - DRC
Life, social and livelihood skills and	<ul style="list-style-type: none"> • Ashraf (2018) - Zambia

Intervention group	Included studies - country
capacity building (n = 26)	<ul style="list-style-type: none"> • Bastian (2018) - Tanzania • Buehren (2017) - South Sudan • Croke (2020) - DRC • Chinen (2016) - Uganda • Decker (2018) - Malawi • Field (2010) - India • Gottlieb (2016) - Mali • Halim (2019) - Tanzania • Lecoutere (2017) - Uganda • Leight, (2021) - Ethiopia • Murray, (2020) - Zambia • Noble (2019) - Afghanistan • Oxfam (2014) - Sierra Leone • Özler (2020) - Liberia • Scales (2013) - Bangladesh • Stark (2018) - DRC • Stark (2018) - Ethiopia • Tanner (2017) – DRC • Tanner (2017) - Ethiopia • Lecoutere, (2018) - Uganda • Sharma (2020) - Ethiopia • Stark (2018) - Ethiopia • Tanner (2017) - DRC, Ethiopia • Kumar et. al (2018) - Zambia
Safe spaces (n = 6)	<ul style="list-style-type: none"> • Buehren (2017) - Tanzania • Erulkar (2017) - Ethiopia • Sieverding and Elbadawy (2016) - Egypt • Mercy Corps (2015) - Niger • Austrian (2014) - Uganda • Austrian et. al (2020) - Zambia

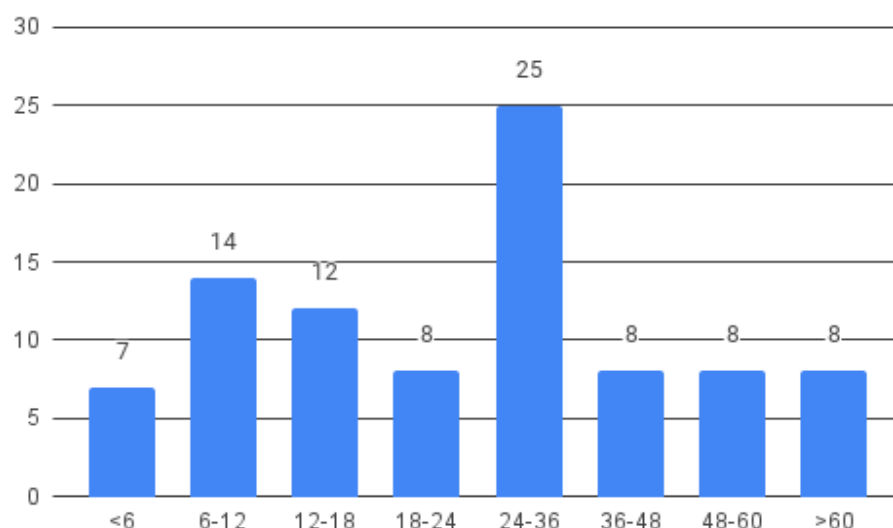
Figure 12 provides a list of all included interventions and the corresponding number of studies identified for included.

Figure 12: Distribution of intervention types



The focus of the intervention types aligned with the UNSCR Pillars covered by our studies: The Participation pillar was the most frequently occurring pillar in our sample ($n = 96$) followed by prevention ($n = 31$), protection ($n = 15$) and recovery and relief ($n = 5$). As demonstrated in Figure 13, the diversity of the types of interventions were also illustrated by the range of duration of intervention from 0.5 months to over 140 months, although most of our studies ($n = 74$) had an intervention length below 48 months.

Figure 13: Intervention length (months)



The ecological framework, or targeting level, of studies was primarily at the individual level ($n = 78$). Studies would often focus on both the individual and household level or integrate community-level ($n = 35$) interventions for broad-based interventions. The household ($n = 32$) was also a key location for interventions. Women are the target group of most of our studies ($n = 84$), but men were also a key target group ($n = 33$) for our studies, either through their involvement in cross-group interventions or through targeted interventions engaging directly with men for gender equality and women's

empowerment. This is followed by girls (n = 29), with only a minority of studies specifically focusing on boys (n = 5). Our studies also highlight a gap in literature on interventions focusing on non-binary individuals. This further corroborates the lack of studies overall in LGBTQIA+ communities in fragile contexts.

We observed a lack of reporting on the interventions funding (to be distinguished from the study funding) since approximately one-third of the included studies do not report on the source of funding of the intervention (n = 40). The main sources of intervention funding were governments (n = 35) and multilateral organisations (n = 16), followed by NGOs (n = 14), foundations (n = 11) and private sources (n = 7). The teams of the study and the intervention were often distinct since a vast majority of interventions were implemented by an implementing agency (n = 72) and only a minority were implemented by researchers (n = 9).

Equity dimensions of included studies

A key aspect of studies was equity. Equity is the absence of avoidable and unfair conditions between or among people that hinder or prevent them from attaining their full potential. It is inherently a moral and socioeconomic judgment of fairness. Often, vulnerable populations have intersecting equity concerns, such as gender, socioeconomic class, race, ethnicity and religious belief.

The focus on gender specific and transformative interventions resulted in 100 per cent (n = 104) of all included unique studies demonstrating an equity dimension. We evaluated studies on eight different measures of equity, including the disaggregation of data by sex or other sub-group analysis, equity sensitive analytical frameworks or theory of change, equity-sensitive methodologies, targeting specific vulnerable population(s), measuring effects on a gender or equity outcome, as well as a research process informed by gender or equity.

The most common equity markers were vulnerable population(s) targeted (n = 90). As mentioned earlier, women were the largest target group. Because women tend to have intersecting identities and vulnerabilities that are compounded by those identities, most studies were placed into this category. Equity sensitive analytical framework (n = 85) and equity sensitive methodology (n = 82) follow closely behind. For example, multiple studies, including Millan (2020) looked at differential impacts on social and structural inequalities, such as the stratification of inequality towards indigenous versus non-indigenous women. Other impact evaluations include equity-driven theories of change that consider gendered social relations in a community or household and how they may reinforce or change gender norms and roles. This is coupled with a focus on an equity sensitive research process (n = 66) that considers the safety and bias of potential respondents when collecting data. For example, Abramsky, et. al, 2014 ensured that women's security in treatment communities for the SASA! programme was prioritised during data collection.

Measuring effects on an inequality outcome was frequent (n = 66), as the outcomes of interest for this review (gender equality and women's empowerment, as well as human security outcomes) automatically were inclusive of the concept of inequality. When interventions targeted both men and women or boys and girls, data was disaggregated by sex (n = 48), though a smaller portion of studies had disaggregation by sub-groups,

including caste, ethnicity and age. Lastly, approximately one-third of the included studies (n = 33) reported having ethical approval from an institutional review board (IRB) or other governing body.

We also evaluated 19 different equity dimensions that go beyond the demographics of study participants but look at specific factors or dimensions of a study population. Primarily mentioned (n = 97) is sex, including the use of gender denoting biological sex, followed by place of residence (n = 80), as economic empowerment and financial inclusion studies tended to focus on rural or informally settled populations.

Additionally, studies stratified findings by age (n = 55) to look at the differential impacts on adolescents or young adults, including educational level/social capital (n = 36) and whether the treatment group is the head of household (n = 18). Conflict-affected (n = 24) situations were also common, as studies conducted in post-conflict situations or ongoing war contexts were applicable. Socioeconomic status (n = 63), including income or poverty status and social capital were most common in economic-focused studies that took participants' socioeconomic status into account.

Less mentioned dimensions or not appearing altogether included people living with HIV/AIDS (PLHIV), land size, ownership, disability, culture, refugees and religion, as well as sexual orientation and identity. The latter did not appear in any studies despite the inclusion of an additional search strategy for papers using LGBTQIA+ search terms.

Study funding and independence

Study funding and independence refers to funding of the included papers, the evaluations and the independence of study authors to the programmes being evaluated. Funding for included studies came from five sources: public institutions (n = 51), private institutions (n = 4), multilateral organisations (n = 12), foundations (n = 21) and NGOs (n = 9). Seventeen remaining studies did not provide information on their funding and two were not funded. Several studies were funded through multiple sources.

Independence of a study was defined by two variables: the relationship between the implementing agency and the study team, and whether the data collected by an independent party. For 55 studies, the funding and author team were independent of implementers/funders of the programme. In eight cases, the funding was independent of implementers/funders, but authors of the study were from the programmes' funder/implementers. Ten studies were not independent since the evaluation was funded and undertaken by funders/ implementers of the programme being evaluated. Finally, 29 cases did not have clear enough data to verify independence.

For whether the data had been collected by an independent party, 46 studies reported independence in data collection, 13 did not have independently collected reported data and 43 did not provide enough information to form a decision.

We also searched for whether there were any issues with conflict of interest. This meant searching for whether there was any declaration of interest from the study authors. For most studies (n = 63), there was not enough information to make any clear decision, and in 31 there was no conflict of interest. Eight studies had a clear conflict of interest, whether due to the fact the study was an internal evaluation, or that evaluators were associated with international organisations. All studies (n = 104) were published in English.

Implementation of study protocol of included studies

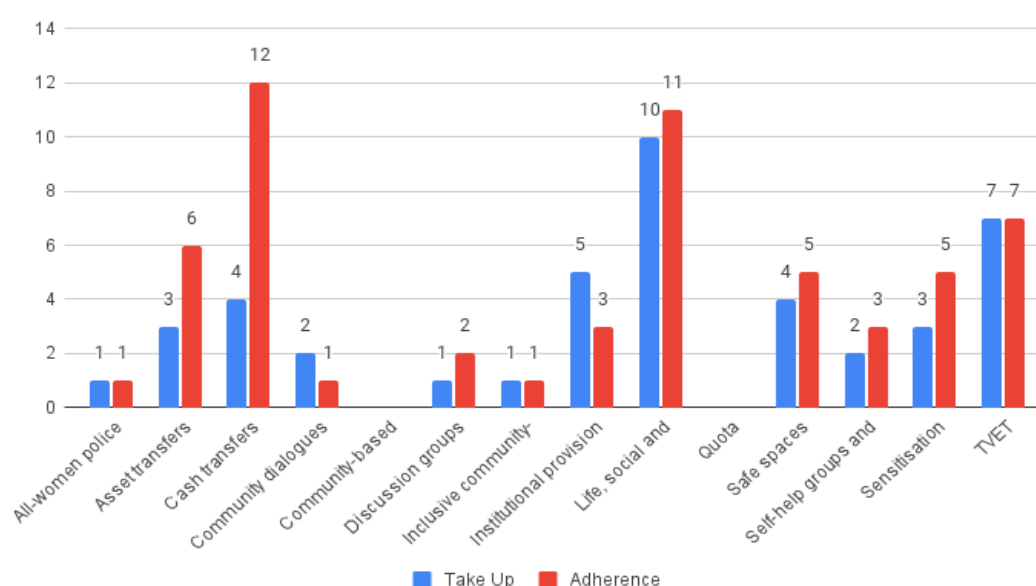
In this section, we describe the studies, including their components, duration and implementation characteristics. We also discuss the ways studies reported on generalisability of their findings and achievements of the theory of change of the programme.

Our studies were assessed through three components:

- Reporting on adherence to project and programme activities: refers to the actual participation of local population and target groups to the activities of the programme. This can be notably reviewed through the level of attrition in the programme or project: it refers to the loss of study units from a sample over the course of the programme or project. Did the beneficiaries take part in the project over the implementation period?
- Reporting on uptake of project and programme activities: refers to the level of engagement of the beneficiaries to the activities of the project. Did the beneficiaries take active part in the activities during and after the course of the project?
- Reporting on the fidelity of implementation of project and programme activities: refers to the level of implementation of the study protocol by the study and/or intervention teams. Did the teams implement the intervention and study as initially planned?

Approximately half of the studies reported on programme adherence (n = 56) but only 26 formally assessed programme adherence. Among these, the main reporting methods were commentaries (i.e., mentioned without formal analysis) by the authors (n = 40) and/or observation by implementation staff (n = 42). Only 39 studies reported uptake, and formally assessed by 15. Once again, when reported, uptake is covered through commentaries by the author (n = 24) and/or observation by the implementation staff (n = 28).

Figure 14: Reporting on uptake and adherence per intervention type



This adherence and uptake could be encouraged by the frequency of contact since most of the studies' activities were conducted on a weekly (n = 37) or monthly basis (n = 24), followed by daily contact (n = 15). A vast majority of the studies were 40 months or less in duration (n = 70) and only three were 60 months or longer. Incentives were rarely used to drive adherence and uptake; only 16 studies clearly reported using incentives and the use of incentives was not clear for 43 studies. We observed that uptake and adherence were often reported together and not always distinguished within studies.

Implementation fidelity, to the contrary, was rarely covered in our studies (n = 30) and only three studies formally assessed it. This is not the result of a lack of reporting on the programme characteristics and information since 68 studies make an explicit reference to theory of change of the programme to design their study protocol. A hypothesis on low reporting of implementation fidelity may be that most of our studies were peer journal articles where other information was prioritised.

A non-negligible proportion of our sample of included studies provided analysis not only on the effectiveness of the interventions but also on the causal mechanism, barriers and enablers that drive the success or failure of interventions (n = 46). This reporting was primarily done through commentaries by the author (n = 33). Sixty-five studies covered questions of generalisability of their findings. On the other hand, we observed that only 25 of our studies reported on cost, which appears to contrast with the fact that a high portion of our interventions focused on financial inclusion and cash-based support.

Another element to note is that a vast majority of the included studies measure the impact of intervention directly after the end of the intervention: 66% (n = 86) of the included studies have a follow-up period of less than a month, 9% (n = 12) have a follow-up period of 1 to 6 months, 10% (n = 13) of 7-12 months, 1% (n = 1) of 13-18 months, 2% (n = 3) of 19-24 months, and only 3% (n = 4) over 25 months.

Outcomes

Our outcome framework was based on the three dimensions of women's empowerment and gender equality designed by Kabeer for a total of 44 unique outcomes as per Table 14 below:

Table 14: List of outcomes and number of studies covering them

Outcome categories	Outcomes
Resources (n = 71) material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services <ul style="list-style-type: none"> AA1 Women have access to rights, services and opportunities (n = 3) AA2 Increased employment of women in law and justice sector agencies (n = 1) AA3 Women have improved access to labour rights (n = 0)
	Economic livelihood related resources <ul style="list-style-type: none"> AB1 Increased capacity of women to understand and use financial, banking, and business services effectively (n = 15) AB2 Women have increased access and ownership to assets, credit and income (n = 64)

Outcome categories	Outcomes
	<ul style="list-style-type: none"> • AB3 Women and girls have equitable access to livelihood support services (n = 3) • AB4 Women and men benefit equitably from employment and livelihood opportunities during recovery and reconstruction (n = 0) • AB5 Durable and reliable housing for vulnerable populations, including women and girls (n = 1) Access to employment <ul style="list-style-type: none"> • AC1 More women engaged in other micro, small, and medium-sized enterprises (n = 6) • AC2 Women can access decent work (formal and informal employment) (n = 27) • AC3 Improved capacity of women entrepreneurs (n = 12)
Agency (n=87) ability to define one's goals and act upon them and operationalised decision-making	Individual agency <ul style="list-style-type: none"> • BA1 Women have improved success in the workplace (n = 3) • BA2 Women have more and better control over their bodies and sexual health (n = 21) • BA3 Women have increased freedom of movement and association (n = 28) • BA4 Women are more aware of their rights and the roles and responsibilities of duty bearers (n = 10) • BA5 women have more positive attitude towards taking action to claim their rights (n = 16) • BA6 Reduced percentage of women agreeing with certain reasons that justify violence against women and girls (n = 19) • BA7 Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them (n = 22) • BA8 Reduced support for or occurrence of child and forced marriage (n = 12) Community level agency <ul style="list-style-type: none"> • BB1 Increased participation in decision making by Women at the household or community level, including during crisis response (n = 48) • BB2 Women participate more in their community (n = 14) • BB3 Increasingly women-led organisations have the capacity or greater capacity to actively engage and influence other actors (n = 0) • BB4 Women increasingly form their own organisations (n = 2) Institutions supporting agency <ul style="list-style-type: none"> • BC1 The rights, safety, and security of women, men, girls, and boys are protected during relief, recovery, and

Outcome categories	Outcomes
	<p>reconstruction (n = 1)</p> <ul style="list-style-type: none"> BC2 Increased capacity of women to protect themselves from rights violations and advocate for and monitor service delivery including during relief, recovery, and reconstruction (n = 0) BC3 The rights of girls and boys are promoted through curricula and teaching methods (n = 0)
<p>Achievements (n=85) ways of being and doing which can be realised by different individuals</p>	<p>Improved systems</p> <ul style="list-style-type: none"> CA1 Power holders have improved awareness and responsiveness to the demands, claims, rights and inputs of women (n = 3) CA2 increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration (n = 15) CA3 Institutional mechanisms and structures monitor human rights violations and ensure the security, safety and health of women and girls (n = 1) CA4 Increased attention and focus on the needs and priorities of women and girls, and other vulnerable groups during relief and recovery in conflict and post-conflict settings (n = 0) CA5 Increased representation of women at all levels as decision-makers in post-conflict countries (n = 1) CA6 Effective prevention strategies supported to end violence against women and girls (n = 5) CA7 Women have improved and equitable access to justice (n = 3) CA8 Increased community support for women's and children's human, economic and legal rights (n = 11) CA9 Strengthened, accountable, and gender-responsive law and justice system (n = 0) CA10 CSOs have greater capacity to support and be responsive for women and marginalised groups to take collective action and hold power holders accountable (n = 0) CA11 CSOs are more inclusive to all genders (n = 1) <p>Norms and behaviour change</p> <ul style="list-style-type: none"> CB1 Increased awareness in communities of the issues affecting women (n = 3) CB2 Communities have a more positive attitude towards women/marginalised groups (n = 10) CB3 Women have improved attitudes, self-image and confidence (n = 29) CB4 Improved attitudes and increased support for women's

Outcome categories	Outcomes
	<p>economic, social and human rights by men, household and family members and community members (n = 25)</p> <ul style="list-style-type: none"> • CB5 Decreased violence/discrimination at the household level (n = 9) • CB6 Safer and more secure household, communities and areas/territories for women, girls, men, and boys (n = 15) • CB7 Reduced frequency and distribution of types of violence by an intimate partner (n = 23) • CB8 Improved quality of relationships between women and their household and community members (n = 31)
	<p>Empowerment index</p> <ul style="list-style-type: none"> • CC1 Empowerment/Equality Index (n = 27)

Study design

Across the entire review, 97, or 75 per cent, of the included studies were randomised controlled trials (including cluster RCTs) where random assignment was done at the individual or cluster level. The remaining 25 per cent were quasi-experimental designs. See Table 15 below for an overview

Table 15: List of outcomes and number of studies covering them

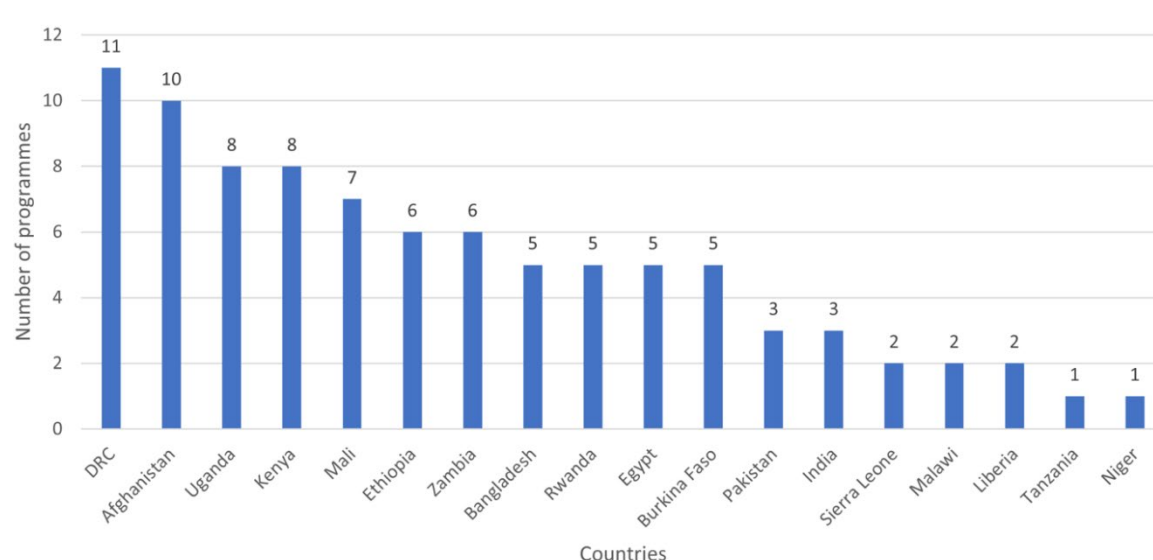
Category	Study design	Number	% of total
Experimental Designs	Randomised Controlled Trial (RCT)	29	21%
	Cluster Randomised Controlled Trial (Cluster RCT)	72	53%
Quasi Experimental Designs	Natural experiment: randomised or as-if randomised	1	1%
	Natural experiment: regression discontinuity (RD)	6	4%
	CBA (non-randomised assignment with treatment and contemporaneous comparison group, baseline and end line data collection) – individual repeated measurement	13	10%
	CBA pseudo panel (repeated measurement for groups but different individuals)	1	1%
	Interrupted time series (with or without contemporaneous control group)	0	0%
	Panel data, but no baseline (pre-test)	0	0%
	Comparison group with end line data only	14	10%

Characteristics of process evaluations and qualitative studies

Out of the 90 included qualitative studies, cash transfers (n = 25) were the most common intervention category. This was followed by life skills (n = 14), self-help groups and VSLAs (n = 12), asset transfers and sensitisation campaigns (both n = 10). The remainder of the intervention categories are as follows: inclusive community driven development (n = 6), TVET (n = 5), safe spaces (n = 4), discussion groups (n = 2), All Women Police Stations (n = 1) and community dialogues and reconciliation (n = 1). Finally, no qualitative studies and process evaluations were found for community-based services, institutional provision of loans and savings and quotas. Most of the qualitative evidence base comprises of descriptive quantitative studies (n = 47) while process evaluations account for the smallest study type category (n = 14). Only 29 qualitative research studies were included for the qualitative synthesis.

Figure 15 below shows the distribution of these qualitative studies by country. The qualitative studies covered 18 different countries. The qualitative evidence base was dominated by studies in Sub-Saharan Africa (77%) and the remainder are collectively from Afghanistan, Bangladesh, Pakistan and India. As the studies in the quantitative review component, qualitative studies were heavily concentrated in a small number of countries, including DRC and Uganda. DRC accounts for the greatest number of qualitative studies (n = 11), followed by Afghanistan (n = 10). Uganda and Kenya each have 8 qualitative studies focused on programmes from these two countries. The least number of qualitative studies found were based on programmes in Niger and Tanzania both account for only 1 study each.

Figure 15: Distribution of qualitative studies, by country



As indicated above, we explored the data reported in the 90 studies included in the qualitative review component according to four analytical lenses in the thematic synthesis: intervention design, intervention implementation, population characteristics, and contextual variables; and how they interact with programme effects. In this process, we generated total of 149 descriptive themes across all studies and 37 analytical themes. These analytical themes present the qualitative synthesis results and are outlined per intervention category below.

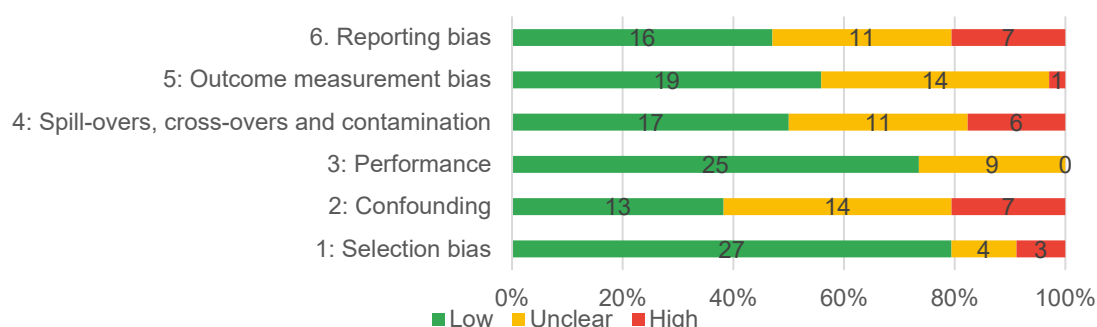
4.1.3 Risk of bias and critical appraisal of included studies

Risk of bias

The risk of bias tool aimed to assess the risk of bias in each included study based on an assessment of key features of study design and analysis. We undertook a RoB assessment using the Cochrane Collaboration Standards for QED and RCT (Cochrane Collaboration, n.d.). The figures 17 and 18 below provide a summary of the assessment of risk of bias in the included studies, with a breakdown for each domain for QED and for RCT. Online Appendix C gives a full summary of our risk of bias assessment at the study level. As we can see, the quality of the included impact evaluations varied.

Among the 34 quasi-experimental design (QED) studies, only 6% (n = 2) have an overall low risk of bias, while 41% (n = 14) have a high risk of bias and 53% (n = 18) raise concerns. By RoB domain, confounding and reporting bias are the categories with the most studies rating high risk of bias or some concerns. For confounding, only about a third of the studies were judged to be low risk (n = 13), which means that most studies were unclear or raised at least some concerns about results being confounded by some unobservable characteristic that was not or could not be accounted for during the analysis. For reporting bias, 45 per cent of included QED studies (n = 16) have a low risk, while the rest are either not clear or present concerns related to authors not using a credible analysis method, some evidence that the outcomes were selectively reported or lack of sensitivity analysis or robustness checks to add credibility to the results. Also, for spill-overs, crossovers and contamination, just above half of the included studies (n = 17) do not present implementation issues, comparisons are unlikely to receive the intervention, no external program can systematically affect treatment and control groups differently, and there are no survey effects, the rest of studies were unclear or raised some concerns related to any or all of these issues. The other three RoB domains (selection bias, performance and outcome measurement) raised less concerns that will be explored in more detail in each of the intervention sections.

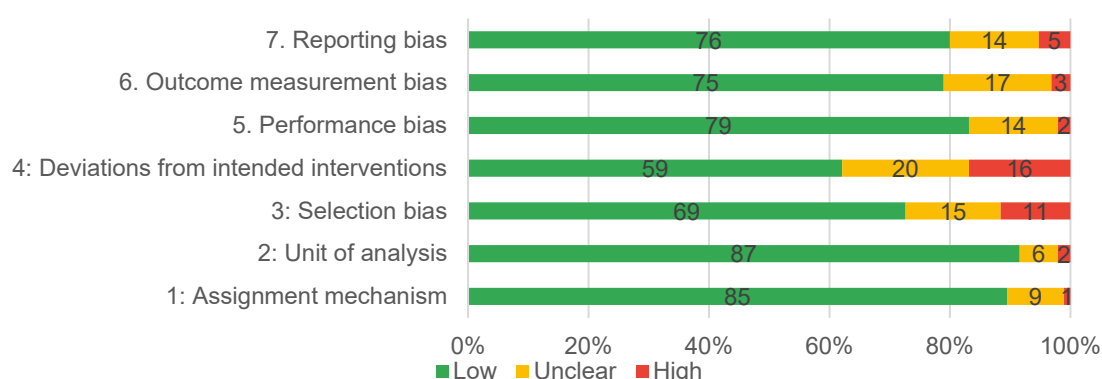
Figure 16: RoB Assessment of QED included studies



As expected, RCT studies raised less concerns regarding risks of bias. Among the 95 included studies using a RCT design, 26 per cent (n = 25) of the studies received an overall rating at low risk of bias and 28 per cent (n = 27) at high risk, reasons for concern vary across studies. If we look at the specific domains, only two have a relatively large number of studies scoring high risk or being unclear. The first is deviations from intended interventions because of implementation issues or spill-overs to comparisons (n = 16 studies scored high risk and n = 20 were unclear). We noticed in this domain that many papers, despite being prospective evaluations, do not discuss or consider potential contamination issues from other programs. The second domain with the largest number

of high-risk studies was selection bias related to issues with attrition or purposeful sampling (n = 11 high risk and n = 15 unclear). For all other domains (assignment mechanism, unit of analysis, performance, outcome measurement and reporting bias) at least 80 per cent of the studies scored low risk of bias.

Figure 17: RoB assessment of RCT included studies

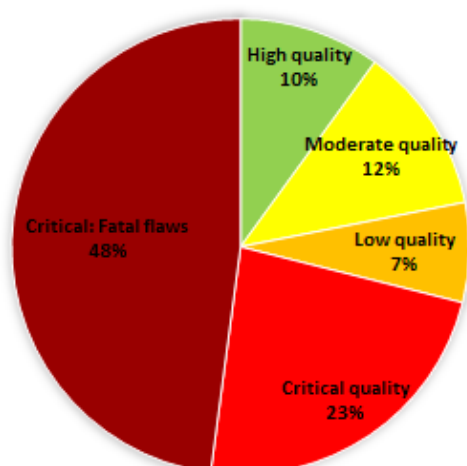


Critical appraisal of process evaluations and qualitative studies

All studies included in the qualitative review component were critically appraised for the trustworthiness of their contribution to the thematic synthesis. We applied a predefined mixed-methods appraisal tool developed by Langer et al. (2017) and applied in Snilstveit et al. (2019) (see Online Appendix A). The appraisal tool made judgments on the adequacy of reporting, data collection, presentation, analysis and conclusions drawn. The appraisal tool assessed the quality of the included qualitative studies, descriptive quantitative studies and process evaluations.

We rated studies on a scale from high quality, to moderate, low, and critical trustworthiness. Figure 18 provides the results of the critical appraisal on aggregate while Figure 19 presents the breakdown of appraisal ratings per appraisal category. Due to the mixed availability of qualitative documentation from the included studies, a few intervention groups did not have qualitative synthesis including community-based services; loans and savings; discussion groups; all women police stations; community dialogues and reconciliations; and quotas.

Figure 18: Summary of overall critical appraisal ratings across studies included in the qualitative synthesis

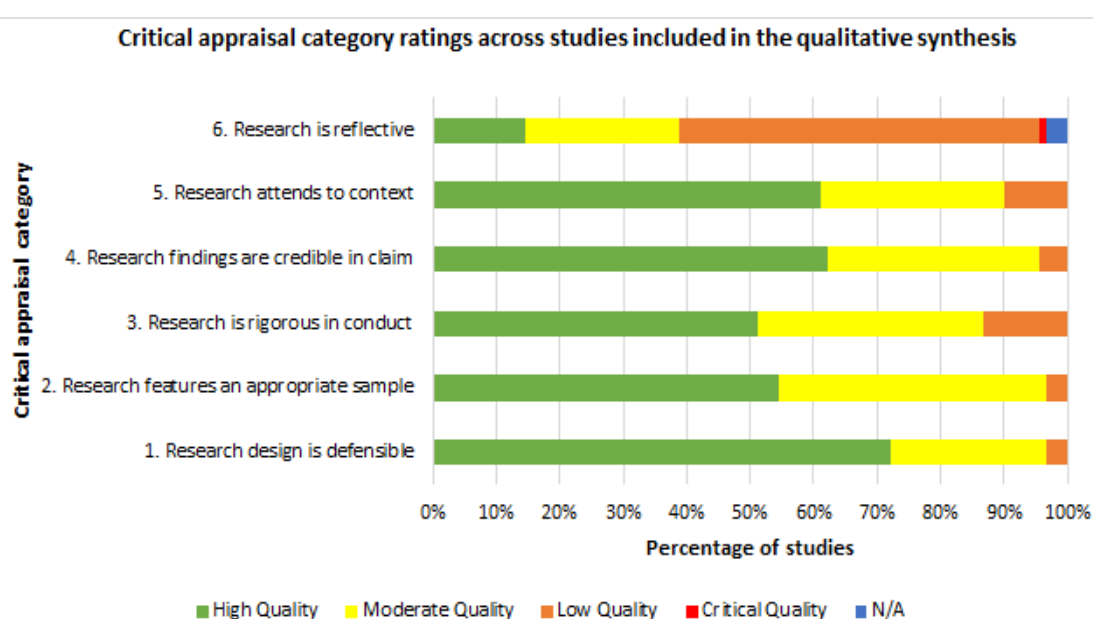


Of 311 critically appraised studies, 71% were rated as critical while the remaining 29% were either of high (10%), moderate (12%) and low quality (7%). This translates into a total of 221 studies being excluded from the synthesis due to critical trustworthiness. Of these, 149 were rated critical based on the presented fatal flaws (Online Appendix A) which refers to a significant lack of reported information within identified studies (e.g., no research question or objective reported; no evidence of empirical data being provided). These qualitative studies subject to fatal flaws constituted 48% of the critically appraised studies and provided a main driver of the high number of studies excluded from the synthesis. The second main reason for critical appraisal ratings referred to the appraisal domain of reflectiveness and lack of triangulation (15%); with the reminder of the critical appraisal rating being distributed equally across the remaining domains.

We included a total of 90 studies in our qualitative evidence base following the removal of critically rated studies. Out of these 90 studies, we ranked 31 as high empirical quality, 37 as moderate empirical quality and 22 as low empirical quality. This implies that just over a third of the included studies (34%) were of high trustworthiness while the remaining studies were of either moderate or low trustworthiness (66%). Overall, however, the qualitative evidence base was of poor quality as indicated by the large number of studies includes as of a critical appraisal rating.

For the 90 included studies, Figure 19 provides a breakdown of the critical appraisal ratings by domain. It indicates many high-quality ratings on the defensibility of the research design (72% of studies), the credibility in research claim (62% of studies) and the relevance for context (61% of studies). In contrast, most of the included studies rated low on trustworthiness on the depth of reflection and triangulation (57% of studies).

Figure 19: Critical appraisal category ratings across studies included in the qualitative synthesis⁶

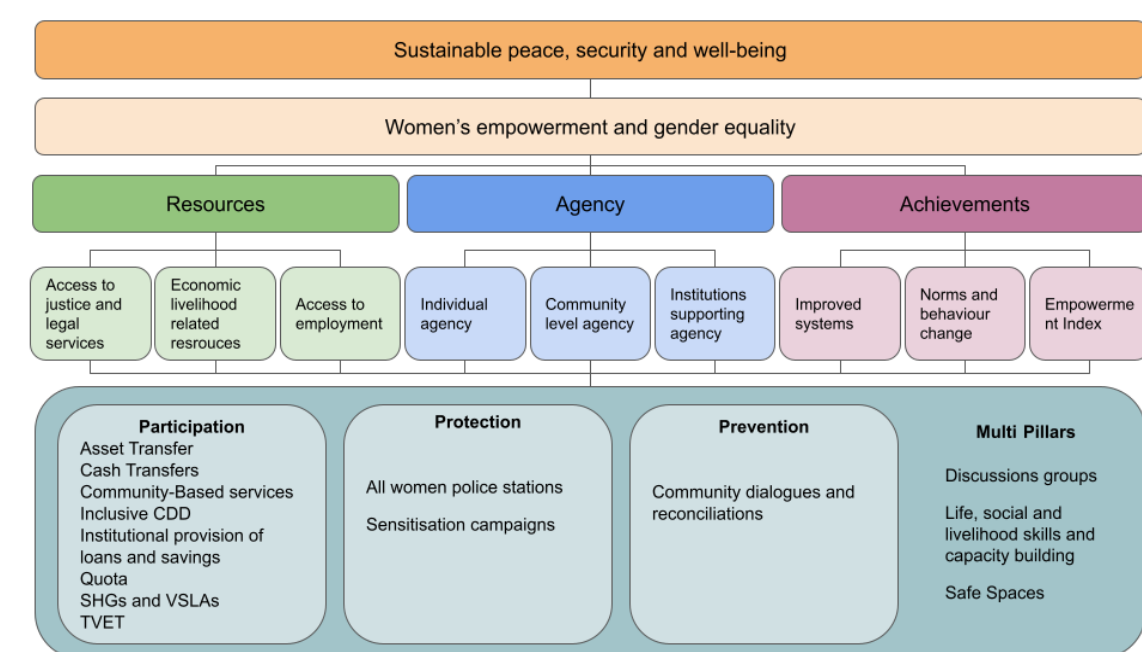


⁶ Excludes studies of critical overall quality.

4.2 Effects of interventions

In this chapter we report the results of the quantitative and qualitative analysis of the effects of our interventions in the included studies of this review. Our analysis is structured around the four pillars of UNSCR 1325 and includes interventions working across multiple pillars. None of the included studies' interventions primarily targeted the Recovery and Relief Pillar. Based on the framework developed in our protocol we identified 14 types of interventions contributing to the achievement of nine immediate outcomes, three secondary outcomes (Resources, Agency and Achievements) and one long term outcome: gender equality and women's empowerment for peaceful and inclusive societies. Figure 20 below provides an overview of the intended effects of the interventions of our included studies covering 290,484 beneficiaries:

Figure 20: Theory of change of included studies



The following sections will present the effects of the 14 intervention types at the outcome level and grouped into the UNSCR Pillars. Online Appendix O provides details on the characteristics of the studies and programmes included in our review.

4.3 Effects of interventions under the UNSCR participation pillar

The participation pillar encompasses a set of interventions that create opportunities for, build acceptance of or strengthen capacities for the equal participation and full involvement of women and girls in political, economic and social institutions and decision-making processes.

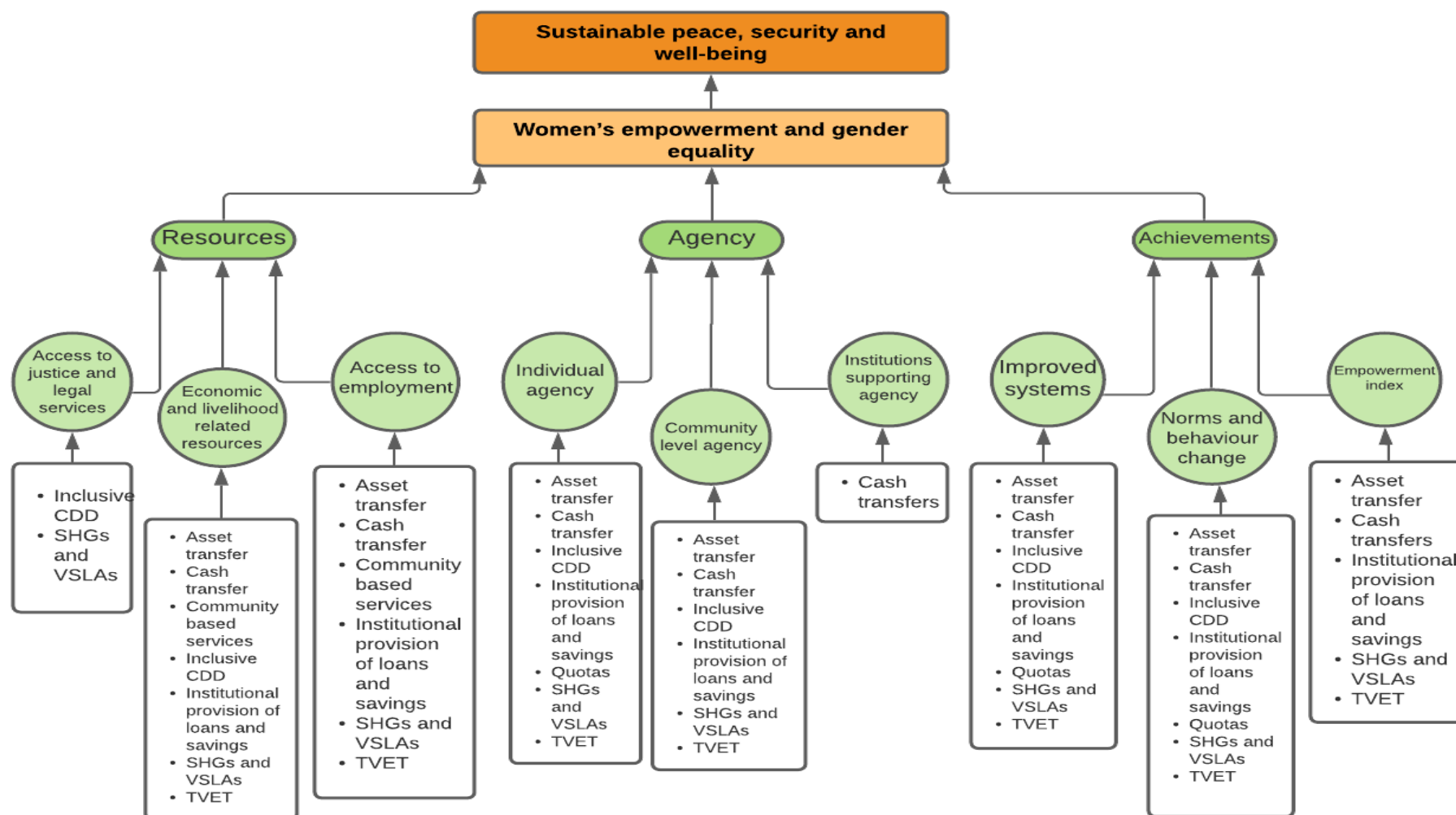
In our systematic review, this pillar gathered evidence from the following types of interventions covering 231,084 beneficiaries:

- Asset transfers
- Cash transfers
- Community-based services

- Inclusive community driven development (CDD)
- Institutional provision of loans and savings
- Quotas
- Self-help groups (SHGs) and village savings and loan associations (VSLAs)
- Technical and vocational education and training (TVET)

This section provides the findings of our synthesis of the 75 included studies evaluating the effect of these interventions on gender equality, women's empowerment and peaceful and inclusive societies. The section is organised by each intervention group. Each sub-section begins with a description of each intervention group, their activities and theory of change, followed by descriptive results and the findings addressing our research questions. Figure 21 provides a summary of the outcomes and targeted effects of the interventions included under the participation pillar:

Figure 21: Theory of change of included studies under the participation pillar



4.3.1 Asset transfer

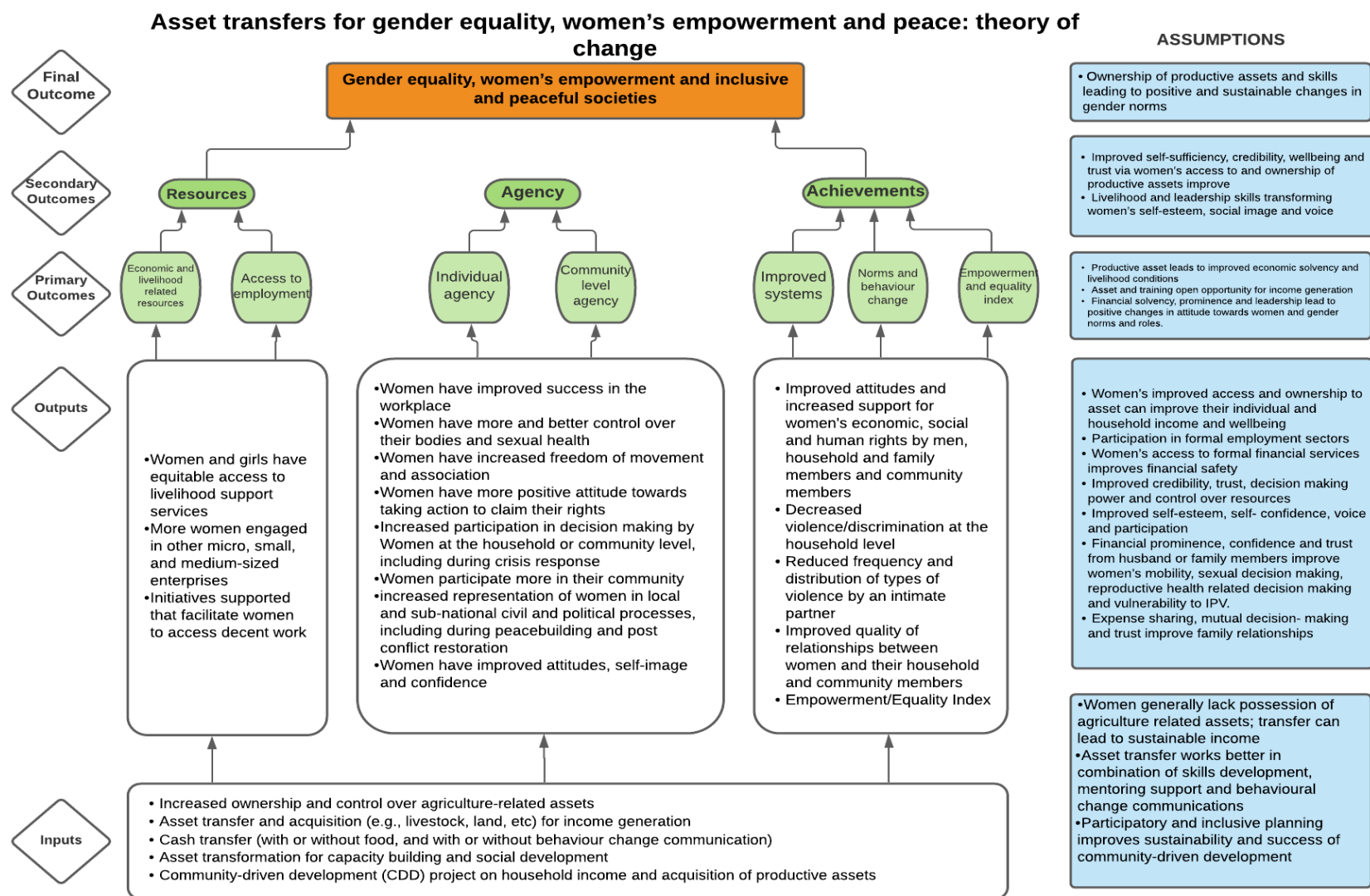
We define asset transfers interventions as the transfer of publicly owned land, buildings or goods (including materials, livestock, etc) to community organisations, individuals or groups (Das et al. 2013).

How do asset transfers affect gender equality, women's empowerment and peace outcomes?

Figure 22 below maps out the causal chain of how asset transfers may improve gender equality, women's empowerment and peace outcomes.

Improving women's and girls' access to productive assets creates multifaceted opportunities in terms of gender equity and women's empowerment. Assets, along with necessary skills development and leadership training, can improve women's access to employment in formal and informal sectors. Financial solvency, on one hand, can ensure personal financial freedom and on the other hand supports families' wellbeing, children's health and educational expenses (which in many cultures are considered a women's responsibility) along with other household needs. Shared responsibilities, mutual decision making, trust and confidence reduce conflict, violence and improves intra-family and social relation and cohesion. Financial prominence, control over assets and other relevant qualities can have positive impacts on socio-cultural norms, gender roles and attitudes unfavourable to women. Mutual respect and understanding can also have positive effects in reducing conflict and SGBV.

Figure 22: Asset transfers for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included 10 studies reported in 14 different papers evaluating the effect of eight programmes. We included more than one paper evaluating the same programme if the author(s) reported different outcomes over several papers.

Population

Asset transfer programmes in this review primarily targeted women and girls from poor households. However, a few of the programmes also engaged with men. Most of the programmes worked with either individual beneficiaries or households. Only one study (Nkonya 2012) evaluated a programme that entailed community level activities.

The included studies evaluated eight different programmes and trials in South Asia and Sub-Saharan Africa, totalling five countries: Afghanistan, Bangladesh, Burkina Faso, DRC and Nigeria.

Intervention, inputs and activities

The included studies evaluated a range of different asset transfer activities and inputs including:

- Improved land rights for increased ownership and control over agriculture assets for women (n = 1): The activities evaluated by the study under this category involved the establishment of village model farms (VMFs) run by women village farm leaders (VFLs). The VMFs served as training sites for women to learn about homestead food production and small animal rearing.
- Asset transfer for income generation and training/support (n = 2): Activities in this category included the transfer of productive assets such as livestock, assets for establishing and running small-scale retail operations, tree nurseries and vegetable growing. Often, assets were offered with a package of complementary training and support.
- Asset transformation for women's social development and skills training (n = 1): The included study evaluated a public safety net programme that supported women-headed poor households (between 18-50 years) with a monthly food ration (30 kg of rice) over a period of 24 months. In addition, to integrate food security with development objectives, the programme also provided training on income generating activities such as livestock rearing, vegetable gardening and awareness-raising on social, legal, and health and nutrition issues.
- Cash transfers for economic empowerment with or without food, and with or without behaviour change communication (n = 4): The interventions covered by the included studies included cash or food support, with or without intensive nutrition behaviour change communication (BCC).
- Community-driven development (CDD) on household income and acquisition of productive assets (n = 1): The included study evaluated a CDD project developed by local farmer groups who designed their own local development plans.
- Asset transfer and integration of skills training and learning to improve economic independence and reduce food insecurity (n = 1): Activities in this category included food and cash transfers along with training on livelihood skills, awareness on social, legal, health, and nutrition issues, basic literacy and numeracy training and access to credit; public works for community infrastructure development and productive asset creation.

- Livestock asset transfer and economic skills training (n = 1): Pigs for Peace (PFP) was a livestock transfer program wherein two piglets aged 2-4 months were provided to participating households under the condition that the household creates a pigpen and a compost pit. The programme also provided practical training for nutrition and care of the livestock asset and support by trained staff.
- Productive asset (livestock, for example) and cash transfer (n = 2): The programmes evaluated consisted of the transfer of a productive asset in the form of livestock, a monthly cash transfer/stipend, basic training on livestock rearing and entrepreneurship and a “health subsidy” which included the provision of a basic hygiene kit and reimbursement for medical expenses or latrine improvements. In one case, the transfer component was accompanied by mentoring visits.
- Provision of land and agricultural training for economic empowerment (n = 1): The included activity was an integrated agriculture and nutrition programme that provided land, input and training to women to promote consumption of nutrient rich foods and generate additional income.

Table 16 describes the activities associated with each included study.

Table 16: Asset transfer activities: features of included studies

Study	Activity/input	Length of treatment	Intervention frequency
Ahmed – 2009	Asset transfer and integration of skills training and learning to improve economic independence and reduce food insecurity.	66 months	Monthly
Bandiera – 2013	Asset transfer for income generation and training/support	Not reported	Not reported
Bandiera – 2017	Asset transfer for income generation and training/support	144 months	Monthly
Bedoya – 2019	Productive asset (livestock, for example) and a cash transfer	12 months	Weekly
Bedoya – 2020	Productive asset (livestock, for example) and a cash transfer	36 months	Annually
Das – 2014	Cash transfer for economic empowerment with or without food, and with or without behaviour change communication	Not reported	Not reported
Glass – 2017	Livestock asset transfer and economic skills training	7 months	Weekly
Nkonya – 2012	Community-driven development (CDD) project on household income and	24 months	Not reported

Study	Activity/input	Length of treatment	Intervention frequency
	acquisition of productive assets		
Olney – 2016	Provision of land and agricultural training for economic empowerment	24 months	Not reported
Pradhan – 2017	Asset transformation for women's social development and skills training	24 months	Monthly
Roy – 2015	Cash transfer for economic empowerment with or without food, and with or without behaviour change communication	Not reported	Not reported
Roy – 2013	Cash transfer for economic empowerment with or without food, and with or without behaviour change communication	Not reported	Not reported
Roy – 2019	Cash transfer for economic empowerment with or without food, and with or without behaviour change communication	24 months	Monthly
van den Bold – 2015	Access to land right for increase ownership and control over agriculture-related asset for women	24 months	Not reported

Comparison

All our included studies compared treated groups to comparison groups receiving no intervention. Six studies included multiple treatment arms.

Outcomes

The included studies reported on several relevant outcomes, including the following:

- Women have increased access and ownership to assets, credit and income (n = 11): Women can apply for, receive and manage assets/credit and income and have support to manage, claim and execute their assets without pressure or influence from external actors, including male family members, husbands and cultural leaders.
- Initiatives that facilitated women to access decent work (formal and informal employment), including people with disabilities (n = 7): Women were able to apply for, receive and work in jobs (and have support for the above), without discrimination for sex, gender or other identifying factors, including development of skills for improved access.
- Increased participation in decision making by Women at the household or community level, including during crisis response (n = 6): Women take part in all or any step of the decision-making process at the household community and district level, but also are able to meaningfully take part and have influence on the final decision, including in crisis response.
- Empowerment/Equality Index (n = 4): Indices using multiple outcomes of the list to aggregate them into a gender equality and/or women's empowerment score.

Our list of 44 outcomes reported in our included studies is grouped into nine immediate outcomes directly contributing to the three secondary outcomes as per Kabeer's dimensions of women's empowerment and gender equality (Resources, Agency and Achievement). The division into the immediate and secondary outcomes in asset transfer is reported in Table 17 below:

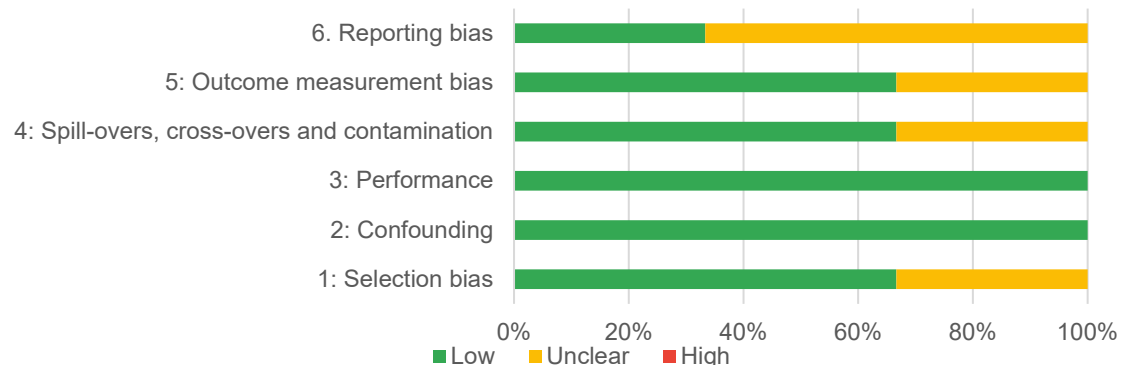
Table 17: Asset transfer summary of secondary and immediate outcomes

Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	0
	Economic and livelihood related resources	13
	Access to employment	6
Agency ability to define one's goals and act upon them and operationalised decision-making;	Individual agency	5
	Community level agency	6
	Institutions supporting agency	0
Achievement ways of being and doing which can be realised by different individuals	Improved systems	1
	Norms and behaviour change	6
	Empowerment index	3

Study design

Three of our asset transfer studies used a QED (Ahmed 2009; Nkonya 2012; Pradhan 2017). One of the three QEDs in this category was assessed as having a low risk of bias, the other two were assessed as having some concerns. As detailed in Figure 23 below, high risk of bias was not identified in any RoB domains. All three studies show no limitations regarding confounding (i.e. we did not identify any issues with selection in to the program being done according to clear rules, which are used for the matching even if there are still slight imbalances remaining after matching). Also, no major concern was observed regarding performance bias as we did not identify any issues regarding the monitoring processes. However, the quality assessment of the studies was limited by the lack of discussion by the authors about potential issues with the study design so that we could not discard limitations related to selection bias, data collection and treatment implementation processes were also not discussed in detail so issues related to spill-overs and outcome measurement could not be discarded. Finally, an unclear description of the methods used in two of the three studies do not allow us to discard any concerns related to potential reporting bias.

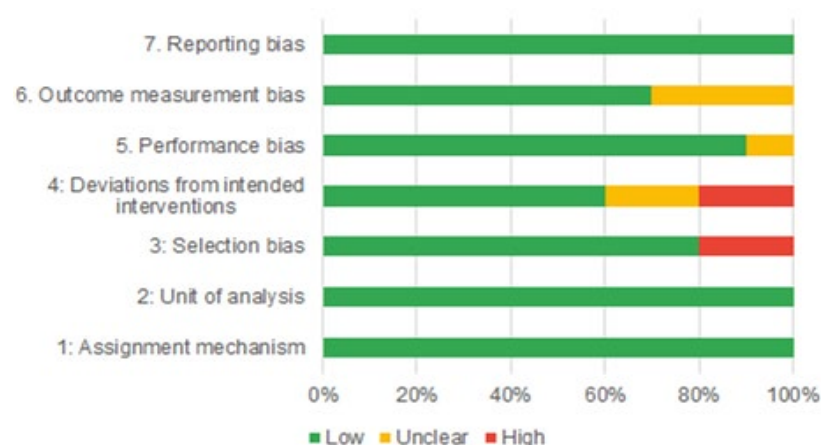
Figure 23: Asset transfer QED RoB assessment



Seven of our asset transfer studies used an experimental design (Das 2014; Bandiera 2017 and 2013; Bedoya 2019 and 2020; Glass 2017; Olney 2016; Roy 2015 and 2019; Van Den Bold 2015). Approximately 20 per cent of the papers scored a low risk of bias, 40 per cent a high risk of bias and the rest showed some concerns. As detailed in Figure 24 below, high risk of bias was identified in two RoB domains (selection, and deviation from intended interventions) and both represented 20 per cent of the sample of included studies.

Risk of selection bias concerns in the asset transfer interventions using RCT designs came from differences in attrition between treatment and control groups that were not adjusted for in the analysis (Olney 2016; van den Bold 2015). Further, risks of deviations from intended interventions were observed because of assignment at the household level within the same villages, where the lack of geographic separation between treatment and control groups poses a risk for contamination (Glass 2017; Bedoya 2019), in other cases spill-overs could not be discarded because participants in control villages were also showing changes in the outcome variables (van den Bold 2015) or authors do not report enough discussion to discard implementation issues (Bandiera 2103). No major concerns were observed for the rest of domains, but again we could not discard limitations related to outcome measurement in three studies (Roy 2015 and 2019; Das 2014) and to performance in Glass (2017). We did not identify limitations related to the assignment mechanism, the unit of analysis or selective reporting of outcomes in any of the included studies.

Figure 24: Asset transfer RCT RoB assessment



Qualitative studies, process evaluations and project documents

We identified ten additional documents related to seven programmes covered by the Asset Transfer group of included studies.

- Pigs for Peace (DRC): one qualitative study
- Enhanced-Homestead Food Production (E-HFP) programme (Burkina Faso): two process evaluations and one qualitative study
- Transfer Modality Research Initiative (TMRI) (Bangladesh): one process evaluation
- Targeting the Ultra Poor (TUP) programme (Afghanistan): one qualitative document and two descriptive quantitative documents
- Income-Generating Vulnerable Group Development (IGVGD) - Bangladesh: one process evaluation
- Targeted Ultra-Poor (TUP) Programme - Bangladesh: one descriptive quantitative document

The asset transfer qualitative documents were ranked as high-quality empirical research ($n = 4$), followed by moderate ($n = 5$) and low quality ($n = 1$).

Synthesis of findings

The following subsection presents the results of effectiveness of asset transfers on Gender Equality, Women's Empowerment and Peace outcomes.

Quantitative findings

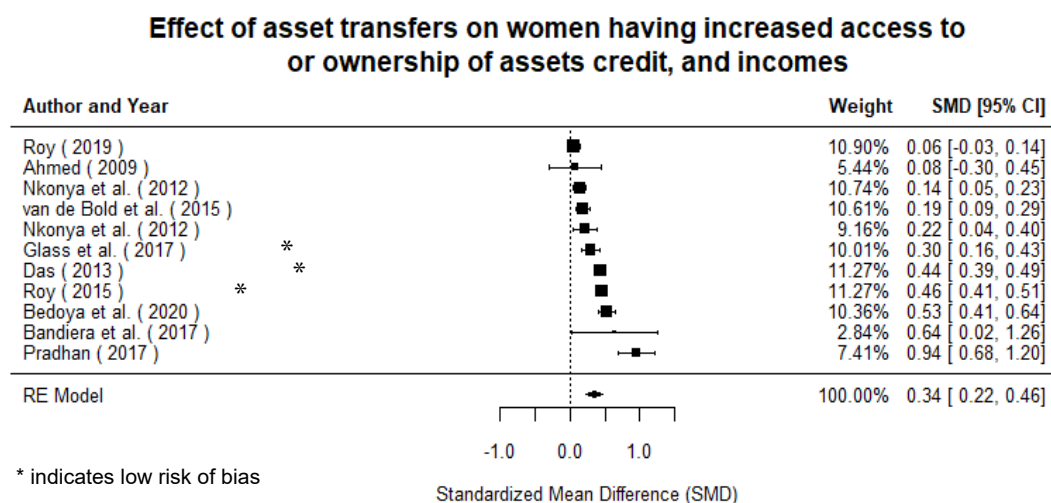
Effects of asset transfers on increased capacity of women to understand and use financial, banking and business services

Bedoya et al. (2020), an experimental study in Afghanistan, was the only study evaluating the impact of asset transfers on the increased capacity of women to understand and use financial, banking and business services. The study reported a large effect ($g = 0.62$, 95% CI [0.50, 0.74]). We assessed this study as having a low risk of bias.

Effects of asset transfers on women's access to and ownership of assets, credit and income

We included a total of $k = 11$ studies. We assessed three studies as low risk of bias, six as some concerns, and two as high risk of bias. The observed outcomes ranged from 0.06 to 0.94. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.34$ (95% CI: 0.22 to 0.46). Therefore, the average outcome differed significantly from zero ($z = 5.55$, $p < 0.001$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 25 AssetAB2.

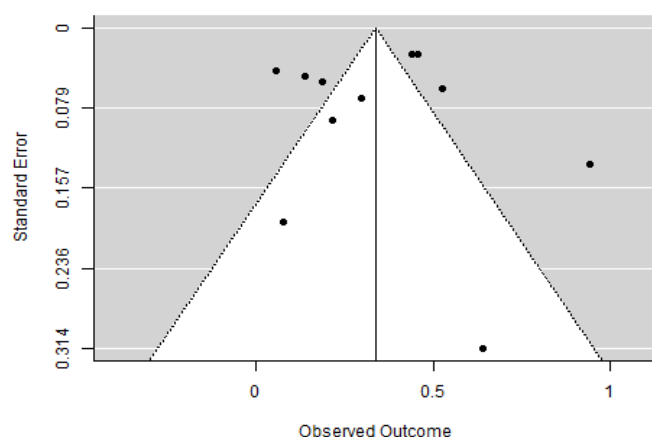
Figure 25: AssetAB2: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(10) = 139.89$, $p < 0.01$, $\hat{\tau}^2 = 0.03$, $I^2 = 92.85\%$). An examination of the studentised residuals revealed that one study (Pradhan 2017) had a value larger than ± 2.84 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

A funnel plot of the estimates is shown in Figure 26 AssetAB2b. Neither the rank correlation nor the regression test indicated any funnel plot asymmetry ($p = 0.76$ and $p = 0.36$, respectively). We had sufficient data to test most moderators for this intervention and outcome group (ROB score, exposure to intervention in months, evaluation period in months, study design (RCT vs QED), whether or not the model was adjusted for covariates, gender inequality score, and fragile states index score), but no moderators were significant. We were unable to test the number of items in the outcome measure given that most were single items related to annual income. Finally, for asset transfers we wanted to examine whether there was a difference between interventions with just an asset transfer versus those with an asset transfer plus an additional training component. However, only one programme offered an asset without an accompanying training, so we were unable to test this as a potential moderator of the size of the effect.

Figure 26: AssetAB2b: Funnel plot



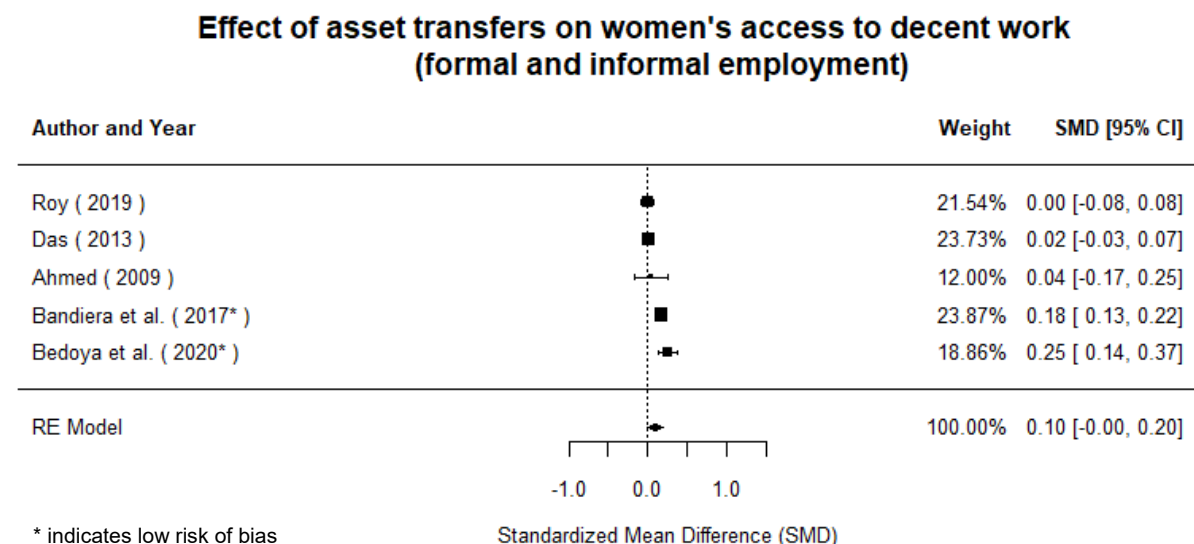
Effects of asset transfers on women's engagement in other micro, small, and medium-sized enterprises

Bedoya's (2019) experimental study in Afghanistan was the only study evaluating the impact of asset transfer on more women being engaged in other micro, small, and medium-sized enterprises. There was a medium, and statistically significant, effect ($g = 0.24$, 95% CI [0.13, 0.36]), however we assessed the study as having high risk of bias.

Effects of asset transfers on women's access to decent work (formal and informal employment)

We included a total of $k = 5$ studies in the analysis. We assessed two studies as low risk of bias, three as some concerns, and none as high risk of bias. The observed outcomes ranged from 0.00 to 0.25. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.10$ (95% CI: -0.00 to 0.20). Therefore, the average outcome did not differ significantly from zero ($z = 1.93$, $p = 0.05$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 25 AssetAC2.

Figure 27: AssetAC2: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(4) = 33.11$, $p < 0.01$, $\hat{\tau}^2 = 0.01$, $I^2 = 87.92\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.58 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

We had sufficient data to test most moderators for this intervention and outcome group (ROB score, exposure to intervention in months, evaluation period in months, study design (RCT vs QED), whether or not the model was adjusted for covariates, gender inequality score, and fragile states index score), but no moderators were significant. We were unable to test the number of items in the outcome measure given that most were single items related to annual income. We were also unable to test whether asset transfers with a training component produced greater effects than asset transfers alone because all of the interventions in this outcome category included the training component.

Effects of asset transfers on improved capacity of women entrepreneurs

Bedoya and colleagues' (2020) experimental study in Afghanistan was the only study evaluating the impact of asset transfer on improved capacity of women entrepreneurs. Their report included two effects that fell into this outcome category (primary woman participates in market work and primary woman is owner or manager of an HH non-agro business). The effects were positive and significant in both cases, the first was a large effect ($g = 0.46$, 95% CI [0.34, 0.58]) and the second a small one ($g = 0.15$, 95% CI [0.04, 0.27]). However, we assessed the study as having high risk of bias.

Effects of asset transfers on women having improved success in the workplace

Bandiera and colleagues' (2013) experimental study in Bangladesh was the only study evaluating the impact of asset transfer on women having improved success in the workplace. Their report included two effects that fell into this outcome category (labour productivity measured by hourly earnings – at 24 and 48 months). The point estimates were very small and not statistically different from zero, ($g = -0.03$, 95% CI [-0.08, 0.02]) and ($g = 0.09$, 95% CI [0.04, 0.14]), respectively. We assessed the study as having some risk of bias concerns.

Effects of asset transfers on women having more and better control over their bodies and sexual health

Ahmed's (2009) quasi-experimental study in Bangladesh was the only study evaluating the impact of asset transfer on women having more and better control over their bodies and sexual health. For each of the three interventions being evaluated (food, cash, and both), their report included four effects that fell into this outcome category (e.g., whether woman ever used birth control). The effects were null, ranging from $g = 0.00$ [95% CI: -0.22, 0.22] to $g = 0.14$ [95% CI: -0.10, 0.38]). We assessed the study as having some risk of bias concerns.

Effects of asset transfers on women's freedom of movement and association

Ahmed's (2009) quasi-experimental study in Bangladesh was the only study evaluating the impact of asset transfer on women having increased freedom of movement and association. For each of the three interventions being evaluated (food, cash, and both), their report included five effects that fell into this outcome category (e.g., whether woman decides by herself to go outside the community to visit friends or relatives). The effects ranged from medium, negative effects ($g = -0.24$, 95% CI [-0.46, -0.03]) to small, positive but non-significant point estimates ($g = 0.17$, 95% CI [-0.07, 0.41]). Only one of the 15 effect estimates (a negative one) was statistically significant. We assessed the study as having some risk of bias concerns.

Effects of asset transfers on reduced instances of child or forced marriage

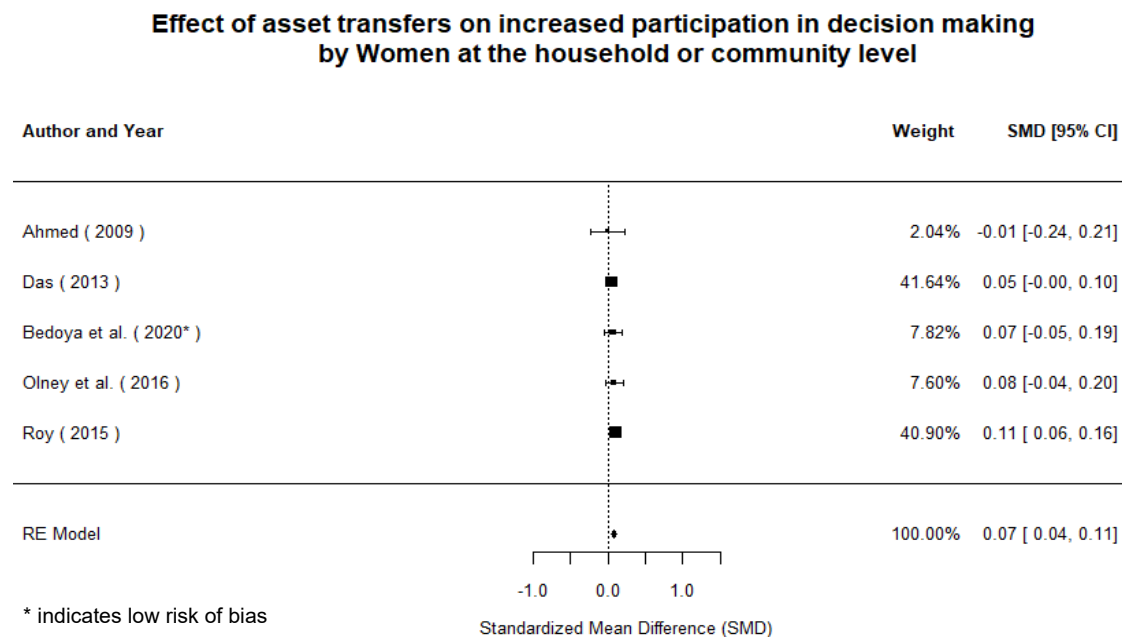
Bedoya and colleagues' (2020) experimental study in Afghanistan was the only study evaluating the impact of asset transfer on reduced instances of child or forced marriage. There was a small negative but non-significant point estimate ($g = -0.14$, 95% CI [-0.31, 0.02]), and we assessed the study as having low risk of bias.

Effects of asset transfers on participation in decision making by women at the household or community level

We included a total of $k = 5$ studies in the analysis. We assessed one study as low risk of bias, three as some concerns, and one as high risk of bias. The observed outcomes

ranged from -0.01 to 0.11 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.07$ (95% CI: 0.04 to 0.11). Therefore, the average outcome differed significantly from zero ($z = 4.51$, $p < 0.01$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 28 AssetBB1.

Figure 28: AssetBB1. Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(4) = 3.06$, $p = 0.55$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.58 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With no heterogeneity among these effects, moderator analyses were not appropriate.

Effects of asset transfers on women's participation in their community

Olney and colleagues' (2016) experimental study in Burkina Faso was the only study evaluating the impact of asset transfer on women participating more in their community. There was a medium, and statistically significant, positive effect ($g = 0.27$, 95% CI $[0.15, 0.39]$), however we assessed the study as having high risk of bias.

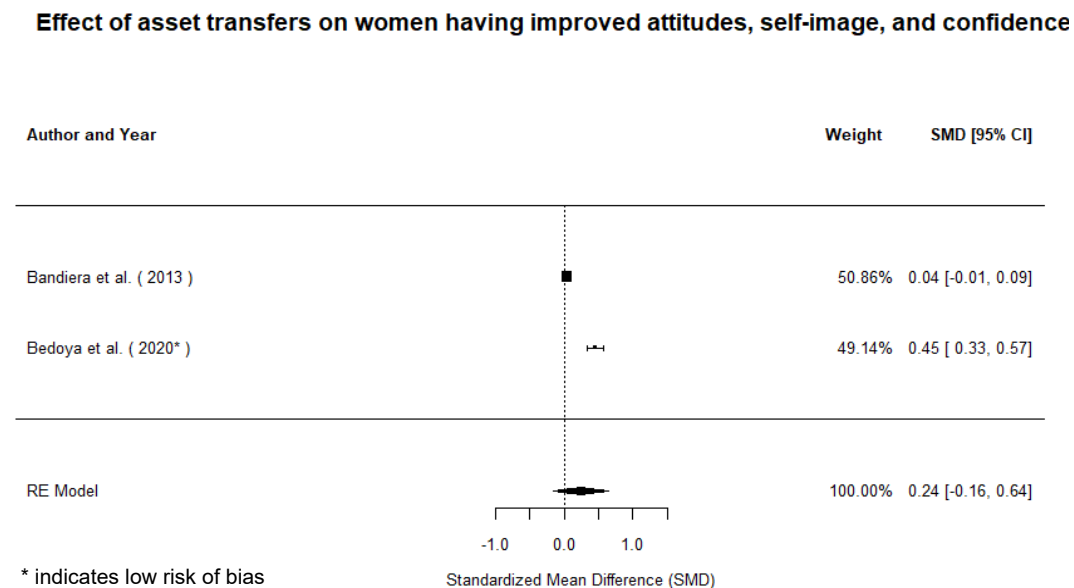
Effects of asset transfers on representation of women in local and sub-national civil and political processes

Bedoya and colleagues' (2020) experimental study in Afghanistan was the only study evaluating the impact of asset transfer on increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration. Their report included 5 effects that fell into this outcome category (e.g., primary woman belongs to a political party). The point estimates ranged from very small, positive and not statistically significant ($g = 0.02$, 95% CI $[-0.10, 0.14]$) to large, positive and significant effects ($g = 0.48$, 95% CI $[0.37, 0.60]$). We assessed the study as having low risk of bias.

Effects of asset transfers on women having improved attitudes, self-image and confidence

Only two studies reported the impact of asset transfers on women having improved attitudes, self-image, and confidence, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as low risk of bias and the other as having some concerns of risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.24$ (95% CI: -0.16 to 0.64). This positive effect did not differ significantly from zero ($z = 1.17$, $p = 0.24$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 29 AssetCB3. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 40.83$, $p < .001$, $\hat{\tau}^2 = 0.08$, $I^2 = 97.55\%$). With only two studies, moderator analyses were not possible, and tests of publication bias are not valid.

Figure 29: AssetCB3. Forest plot showing the observed outcomes and the estimate of the random-effects model



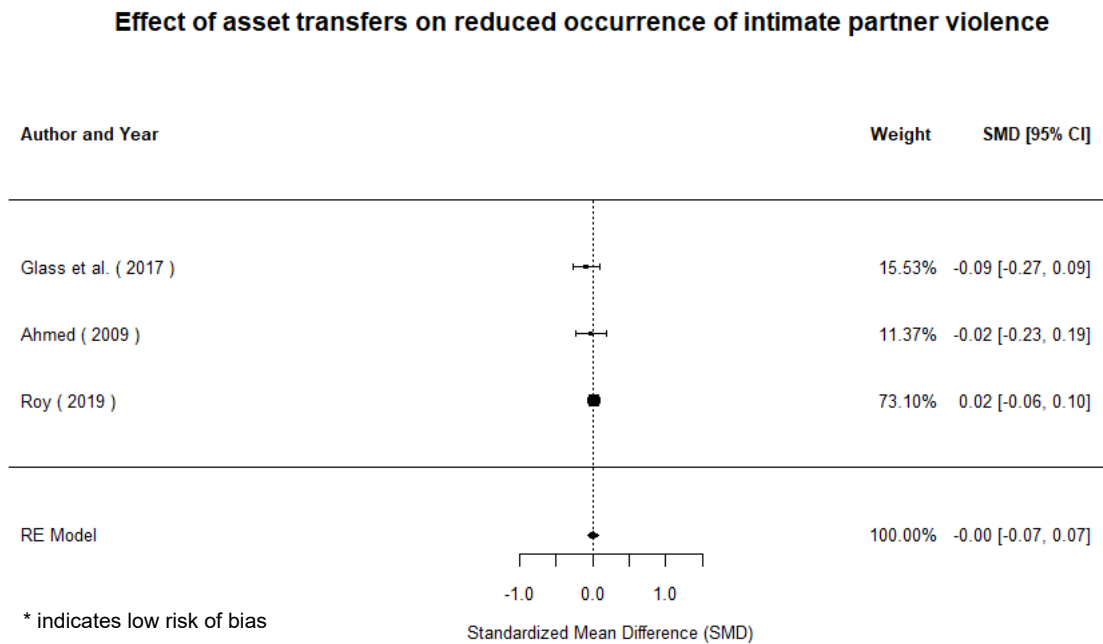
Effects of asset transfers on attitudes and support for women's economic, social and human rights by men, household and family members and community members

Bedoya's (2019) experimental study in Afghanistan was the only study evaluating the impact of asset transfer on improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members. There was a very small, but not statistically significant, point estimate ($g = -0.01$, 95% CI $[-0.14, 0.12]$), and we assessed the study as having high risk of bias.

Effects of asset transfers on reduced occurrence of intimate partner violence

We included a total of $k = 3$ studies in the analysis. We assessed none of the studies as low risk of bias, two as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.09 to 0.02 . The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.001$ (95% CI: -0.07 to 0.07). Therefore, the average outcome did not differ significantly from zero ($z = -0.04$, $p = 0.97$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 30 AssetCB7.

Figure 30: AssetCB7: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(2) = 1.16$, $p = 0.56$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.39 and hence there was no indication of outliers in the context of this model. According to the Cook’s distances, none of the studies could be considered to be overly influential. With only three studies and no heterogeneity, moderator analyses were not appropriate.

Effects of asset transfers on improved quality of relationships between women and their household and community members

We included a total of $k = 3$ studies in the analysis. We assessed none of the studies as low risk of bias, one as some concerns, and two as high risk of bias. The observed outcomes ranged from 0.003 to 0.06. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.05$ (95% CI: -0.03 to 0.13). Therefore, the average outcome did not differ significantly from zero ($z = 1.26$, $p = 0.21$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 31 AssetCB8.

Figure 31: AssetCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model

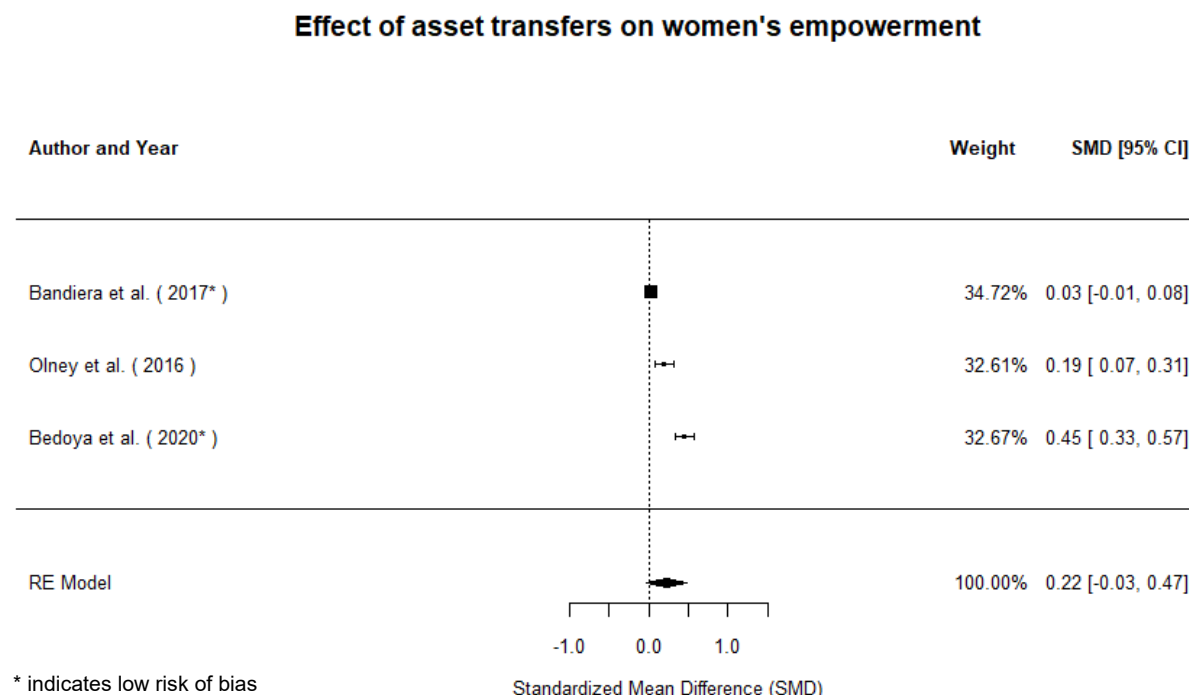


According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(2) = 0.21$, $p = 0.90$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.39 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies and no heterogeneity, moderator analyses were not appropriate.

Effects of asset transfers on women's empowerment index

We included a total of $k = 3$ studies in the analysis. We assessed two studies as low risk of bias, none as some concerns, and one as high risk of bias. The observed outcomes ranged from 0.034 to 0.449. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.22$ (95% CI: -0.03 to 0.47). Therefore, the average outcome did not differ significantly from zero ($z = 1.73$, $p = 0.08$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 32 AssetCC1.

Figure 32: AssetCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(2) = 44.13$, $p < 0.001$, $\hat{\tau}^2 = 0.05$, $I^2 = 95.47\%$). An examination of the studentised residuals revealed that one study (Bedoya 2020) had a value larger than ± 2.39 and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving out each study indicated that removing Bedoya (2010) would reduce the overall average effect ($\hat{\mu} = 0.19$ [95% CI: -0.05 to 0.26]). The effect would still be positive, but no longer significant ($z = 1.31$, $p = 0.19$). According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

Qualitative findings

We conducted a thematic synthesis on the ten linked qualitative studies to the included asset transfer interventions. As indicated above, this thematic synthesis aims to identify themes related to the interplay of intervention design, intervention implementation, target population, and contextual variables with intervention outcomes and effects. In total, we identified 15 descriptive themes, which we configured into four analytical themes (Online Appendix N). These four analytical themes present the synthesis results and are discussed in more detail below.

Theme 1: Water access issues, extreme weather conditions, lack of adequate inputs, animal and pest attacks negatively affect crop and livestock yields

Asset transfer programmes face challenges that could affect the ability of women to use land and to maximise yields (Olney et al. 2015; Van den Bold 2013). In Burkina Faso, findings from Van den Bold (2013) showed that even though there was an indication of greater ability for women to use land for agricultural purposes in the intervention villages, primary obstacles included rudimentary rainfall patterns, inadequate agricultural inputs

and general lack of land to cultivate. In the same country, Olney and colleagues (2015) findings showed whilst there were improvements in production, a significant number of the beneficiaries (83%) had faced challenges with crops they were growing. Specifically, 64% of the beneficiaries faced problems with insufficient water and whilst 31% did not have adequate materials for irrigation. Animal (goats, lizards) and pest (insects) attacks on gardens was an issue for 29% of beneficiaries. 27% of the beneficiaries indicated having insufficient or absence of fences to protect their gardens from pests. In the TUP programme in Bangladesh, beneficiaries reported insufficient nutritional support for animals that they had reared. This had implications on creating more work for beneficiaries to ensure that grass was available for these animals to consume (Sarkar and Karim 2018; Karim 2018).

There is evidence that livestock death and poor-quality seeds received may have affected the implementation of programme activities. Some beneficiaries of the HKI E-HFP programme in Burkina Faso mentioned that the chickens provided by the programme died soon after receipt and two people reported having problems with the seeds they received (Olney et al. 2015).

The TUP programme in Bangladesh reported negative spill over effects of the programme caused by animal destruction of non-member crops. Frequent destruction of neighbours' crops by some beneficiaries' livestock caused conflicts within the community. The issues were resolved by committees that organised and brokered compensation by beneficiaries (Sarkar and Karim 2018), potentially minimising programme gains from asset transfers in the process.

Theme 2: Traditional practices related to gender roles and women's rights to land, absence of spousal, family and/or community support can hinder success of assets transfer initiatives

Qualitative findings showed that asset transfers enhanced women's ability to own land (Van den Bold 2013), promoted women's household decision making (Ahmed, Sraboni and Shaba 2014), and supported marital and family relationships (Ahmed, Sraboni and Shaba 2014; Sarkar and Karim 2018). That said, such findings are prone to the mitigating effects of traditional practices (which stifle support) on gender roles and women's right to land. For instance, the HKI E-HFP programme in Burkina Faso observed changes in opinions regarding women's landownerships in villages and their potential to own and use the land. However, men continued to have majority control over and ownership of land and assets resulting from traditional practices of land ownership. A community level nutrition trainer from the HC pointed out "the weight of tradition" as a key barrier to the adoption of new community level practices (Van den Bold 2013).

In Bangladesh, women reported to continue to have little say in decision making and limited mobility within the community. Focus group and case study evidence from Ahmed, Sraboni and Shaba (2014) showed evidence of strict adherence to stringent gender roles and norms with minimum appreciation for the work that women carry out in the household. In some cases, there were no reported changes to gender divisions of labour within the household despite women taking on more work outside the household (Ahmed, Sraboni and Shaba 2014). In fact, access to new resources could trigger insecurities among other family members who try to maintain control over beneficiary women (Ahmed, Sraboni and Shaba 2014). Qualitative findings from Olney and

colleagues (2015) study on Burkina Faso's HKI E-HFP indicated lack of trust in women's activities by their husbands and other household members. Women also reported that husbands, other members of the household or community members viewed asset transfer related activities as diverting from traditional values. These factors acted as barriers to implementing learnt practices.

Lack of community support emerges as a barrier in asset transfers that brought conflicts between beneficiaries and non-beneficiaries in Bangladesh's TUP programme. For example, one respondent reported non-beneficiary refusal to allow beneficiaries to use paths that cross their lands as they felt excluded. This together with suspicions of beneficiary conversion to a different religion led to resentment and stigmatisation of some beneficiaries (Sarkar and Karim 2018). Other participants reported incidences of verbal abuse by other villagers for participating in these programmes as they considered it inappropriate for women to engage in manual labour (Ahmed, Sraboni and Shaba 2014) In Burkina Faso, however, community level support was expressed through accepting practices, encouragement to implement lessons learnt and by asking beneficiaries for advice or being seen as role models. One beneficiary mentioned that *"The people are interested; they often ask us questions and advice. It pleases us to inform others."*

Theme 3: Multicomponent asset transfer programmes can encourage more sustainable programme impact in improving well-being and beneficiary empowerment

To maximise the effect of asset transfers, interventions could incorporate different components such as legal support for beneficiaries, compulsory saving opportunities, microcredit access, and supplementary cash benefits. Training and capacity building form an important part of these components and is discussed in theme four below. The multiple intervention components complement one another to support livelihoods, social status, beneficiary confidence and awareness of beneficiary rights. In some cases, allowing women to work for their benefits instilled a greater sense of pride. An additional factor to consider is the size of the transfer benefit because it determines how beneficiaries utilise them. There are generally reported mixed preferences for asset and cash transfers. In the case of physical assets, there is less reported undue interference from male counterparts within the household (Olney et al. 2015).

Asset transfer initiatives that are designed to enhance women's well-being and empowerment may require additional components to fully realise this objective. For instance, the Income-Generating Vulnerable Group Development (IGVGD) programme in Bangladesh included a microfinance facility which allowed women to use loans for a multitude of purposes (Ahmed et al. 2007). One woman reported greater economic involvement because: *"After I had joined VGD, I received a 4,000-taka loan from BRAC. With that money I rented a small piece of land for 2 years. My husband and I grow potatoes, chilies, and vegetables on that land. We sell most of what we produce."* Another beneficiary enjoyed the same satisfaction as she proudly said: *"I became a BRAC member right after I had received the [IGVGD] card. I got training from BRAC on how to develop a nursery (to raise vegetable seedlings). After six months of training, I borrowed 3,000 taka from BRAC and started a nursery. Many people come to see my nursery. The current value of the nursery is 10,000 to 15,000 taka"* (Ahmed et al. 2007). In the Targeting the Ultra Poor (TUP) programme, women in focus group discussions

expressed that improvements in economic circumstances and training across a range of areas (including nutrition, hygiene and literacy) created a greater sense of empowerment (Sarkar and Karim 2018). Importantly, the design of programme activities such as compulsory training sessions could be tailored in a manner that does not interfere with standard income-generating activities such as usual labour participation (Ahmed et al. 2007). The additional provision of mobile phones for beneficiary communication with implementation staff for the Transfer Modality Research Initiative (TMRI) programme was reported as satisfying for most of the participants. Preferences on asset type across the included asset transfer programmes revealed diverse views from beneficiaries (Ahmed et al. 2007; Ahmed et al 2014). Across the qualitative evidence base, asset provision was often paired with cash benefits and the overall views on which options were preferred are mixed. For example, beneficiaries of the TMRI largely preferred their current benefit package while the majority of IGVGD preferred food packages (Ahmed et al. 2007; Ahmed et al. 2014).

Theme 4: The provision of quality training promotes knowledge and skills of participants to adequately carry out asset transfer related activities. Compensation of trainers is a key factor to consider improving the adequate delivery of quality training

A consistent theme that we identified was the adequate provision and quality of agricultural training to asset transfer beneficiaries. The embedding of a training and capacity-building component within asset transfer programmes was reflected as a positive design feature throughout the identified evidence base. In Burkina Faso, beneficiaries of the Enhanced-Homestead Food Production (E-HFP) conveyed that the prescribed training led to betterment in practical skills and business knowledge, knowledge of own rights, social status and social capital. One beneficiary of the E-HFP explained that: *“...the training is good quality because since receiving the training I no longer have problems with my vegetable production”* (Olney et al. 2015). Qualitative evidence on the impact of E-HFP on vegetable production was mainly positive as one woman explained that: *“Before we didn’t produce vegetables during the dry season. But now thanks to the programme we have vegetables all year long.”* A master agriculture trainer expressed: *“...we were able to touch on the points that I wanted to further understand like raising poultry and small animals.”* There were some exceptions who found the quality of the training to be poor as another trainer observed that: *“the training leader was not a specialist in gardening.”* Overall, participating farmers suggested that a more participatory and ‘hands-on’ pedagogy and more time for training would improve their learning and retention of information. Additionally, qualitative findings from the IGVGD initiative indicated that training activities should be arranged conveniently and on a frequent basis so that they do not conflict with existing income-generating activities (Ahmed et al. 2007).

Trainers’ compensation and motivation of trainers to carry out appropriate agricultural training should be considered as this may suppress the learning process of beneficiaries. Sentiments by two trainers indicated low remuneration which was not in tandem with the workload, *“What can I do about it? The tasks are heavy, but I have to do it; it’s my job.”* Another echoed a similar view: *“If you look at the work we do, it’s a small compensation. We train them and also follow-up with them; it’s not easy and everything is expensive these days”* (Olney et al. 2015).

Discussion

The quantitative evidence base contains 10 studies ($k = 10$). Results from six analysed outcomes from the quantitative meta-analysis are positive and statistically significant in three areas: 1) women's access to and ownership of assets, credit and income, 2) women's access to decent work (formal and informal employment) and 3) increased participation in decision making by women at the household or community level. While the improved quality of relationships between women and their household and community members and women's empowerment index were positive, these results were not statistically significant. There was no effect on the occurrence of intimate partner violence or women attitudes, self-image and confidence.

Some possible explanations that emerge from the qualitative findings can be summarised as follows:

- The provision of assets or tools for cultivation or livestock farming can generally support asset endowment, creditworthiness and income. If assets are utilised adequately, it opens pathways to credit use options and income-generating activities if used productively.
- The location of access to assets is an important factor for both participation and productive use. If training or farming facilities are inconveniently selected or costly to access for beneficiaries, it can hamper the active participation of beneficiaries.
- Poor quality of training for farmers can hinder willingness to participate. Similarly, the inadequacy of training (on both duration and content) may be a perverse problem – particularly for less experienced farmers (or farming communities).
- Substandard equipment quality and inadequate inputs can impede productivity for beneficiaries and delay programme activities
- Programme activities (including compulsory training) may affect time use preferences for women. Activities that disrupt usual day-to-day work obligations can be ineffective in recruiting and retaining target beneficiaries unless women see it as being beneficial.
- On the lack of impact on the prevalence of IPV, while asset transfer programmes can support asset endowment for beneficiaries more directly, it seems limited on behavioural and social changes. Moreover, if there are poverty reduction effects, household relationships can be less constrained.
- Due to the ownership and management of animals and crops, women can experience improved decision-making and social status within households and around communities.
- Multicomponent interventions can be an important design strategy to ensure the productive and sustained use of assets. The use of training coupled with the provision of loans and saving facilities can lead to more sustained benefits. Interventions should be tailored to target outcomes of interest.
- The qualitative evidence base suggests that beneficiaries tend to prefer the asset type that they have been assigned to or endowed with.

Summary of findings and discussion

We included 10 studies in five countries in South Asia and Sub-Saharan Africa that evaluated the effect of asset transfers. We examined effects on the following outcomes: Women have increased access and ownership to assets, credit and income, women and

girls have equitable access to livelihood support services, Initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities, Increased participation in decision making by women at the household or community level, including during crisis response and the Empowerment/Equality Index. Our included studies reported against all three secondary outcomes (Resources, Agency and Achievement) and seven of the nine immediate outcomes for our review.

The ten linked qualitative studies reported on four different themes that highlighted contextual and natural factors that impeded or promoted the success of the intervention. General climate barriers, including water access and arid climates can impede the success of an asset transfer, especially in an agricultural setting. Traditional gender roles in a community may also hinder or support a women's ability to capitalise on her assets. Lastly, multicomponent asset transfer interventions that combined the transfer with training or an additional component often had the most positive outcomes. The IGVGD programme in Bangladesh, for example, was successful by combining microfinance facilities and other linked components. Overall, the GRADE assessments generally indicate a range of very low to high certainty in this body of evidence.

From the quantitative and qualitative evidence base, we find that asset transfer was statistically significant for women accessing and owning assets, credit and income, which is buttressed by the above qualitative findings. The summary of quantitative findings along with the GRADE certainty of evidence assessments are presented in Table 18.

Table 18: GRADE summary of findings and certainty of evidence on asset transfers

Certainty assessment							Sample size	Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(AB1) Increased capacity of women to understand and use financial, banking, and business services effectively										
1	RCT	not serious	serious ^a	not serious	not serious	strong association	1053	SMD 0.62 SD higher (0.5 higher to 0.74 higher)	⊕⊕⊕⊕ HIGH	Important, but not critical
(AB2) Women have increased access and ownership to assets, credit and income.										
10	RCT – 7 QED – 3	serious ^b	very serious ^c	not serious	serious	none	20864	SMD 0.34 SD higher (0.22 higher to 0.46 higher)	⊕○○○ VERY LOW	Important, but not critical
(AC1) More women engaged in other micro, small, and medium-sized enterprises.										
1	RCT	very serious ^d	serious ^a	not serious	not serious	none	1146	SMD 0.24 SD higher (0.13 higher to 0.36 higher)	⊕○○○ VERY LOW	Important, but not critical
(AC2) Initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities										
5	RCT – 4 QED – 1	serious ^e	very serious ^f	not serious	serious ^g	none	16178	SMD 0.1 SD higher (0 to 0.2 higher)	⊕○○○ VERY LOW	Important, but not critical
(AC3) Improved capacity of women entrepreneurs										
1	RCT	not serious	serious ^a	not serious	serious ^h	none	1140	Two positive effect estimates with a 95% CI range of 0.04 to 0.58.	⊕⊕○○ LOW	Important, but not critical
(BA1) Women have improved success in the workplace.										
1	RCT	serious ⁱ	serious ^a	not serious	serious ^j	none	18387	One positive and one negative effect estimate with a 95% CI range of -0.08 to 0.14.	⊕○○○ VERY LOW	Important, but not critical
(BA2) Women have more and better control over their bodies and sexual health.										
1	QED	serious ^k	serious ^a	not serious	serious ^l	none	589	Five positive and one negative effect estimate with a 95% CI range of -0.22 to 0.38	⊕○○○ VERY LOW	Critical
(BA3) Women have increased freedom of movement and association.										
1	QED	serious ^m	serious ^a	not serious ⁿ	very serious ^o	none	589	Seven negative and three positive effect estimates with a 95% CI range of -0.46 to 0.28.	⊕○○○ VERY LOW	Limited Importance
(BA8) Reduced instances of child or forced marriage										
1	RCT	not serious	serious ^a	not serious	not serious	none	543	SMD 0.14 SD lower (0.31 lower to 0.02 higher)	⊕⊕⊕○ MODERATE	Critical
(BB1) Increased participation in decision making by women at the household or community level, including during crisis response										
5	RCT – 4 QED – 1	serious ^p	not serious	not serious	not serious	none	14638	SMD 0.07 SD higher (0.04 higher to 0.11 higher)	⊕⊕⊕○ MODERATE	Limited Importance
(BB2) Women participate more in their community.										
1	RCT	very serious ^q	serious ^a	not serious	not serious	none	1105	SMD 0.27 SD higher (0.16 higher to 0.39 higher)	⊕○○○ VERY LOW	Limited Importance
(CA2) Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration										
1	RCT	serious ^r	serious ^a	not serious	serious ^s	none	1147	Four positive effect estimates with a 95% CI	⊕○○○	Limited

								range of -0.09 to 0.60.	VERY LOW	Importance
(CB3) Women have improved attitudes, self-image and confidence.										
2	RCT – 2	serious ^w	serious ^x	not serious	serious ^y	none	20384	SMD 0.24 SD higher (0.16 lower to 0.64 higher)	⊕○○○ VERY LOW	Limited Importance
(CB4) Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members										
1	RCT	very serious ^z	serious ^a	not serious	not serious	none	900	SMD 0.01 SD lower (0.14 lower to 0.12 higher)	⊕○○○ VERY LOW	Limited Importance
(CB7) Reduced frequency and distribution of types of violence by an intimate partner										
3	RCT – 2 QED – 1	serious ^t	not serious	not serious	not serious	none	3256	SMD 0.00 SD (0.07 lower to 0.07 higher)	⊕⊕⊕○ MODERATE	Critical
(CB8) Improved quality of relationships between women and their household and community members										
3	RCT – 2 QED – 1	very serious ^u	serious ^v	not serious	not serious	none	2804	SMD 0.05 SD higher (0.03 lower to 0.13 higher)	⊕○○○ VERY LOW	Limited Importance
(CC1) Empowerment/Equality index										
3	RCT – 3	serious ^{aa}	serious ^{ab}	not serious	serious ^{ac}	none	8981	SMD 0.22 SD higher (0.03 lower to 0.47 higher)	⊕○○○ VERY LOW	Important, but not critical

CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

- All single studies downgraded once for inconsistency.
- Downgraded once because two studies present high risk of bias, and others present some concern for bias. That said, when the high risk of bias-studies were removed, the effect estimate only changed by 0.02.
- Downgraded twice for: 1. Variation of points estimates: points estimates on the same side of the threshold, 2. overlaps of CI: no systematic overlaps of CI, 3. $I^2 = 92.85\%$, 4. test of stat: the true outcomes appear to be heterogeneous.
- Based on a single study with a final risk of bias score rated as high. Deviations from intended intervention was a plausible risk.
- Downgraded once since all of the studies but one was an RCT and none of them were high risk of bias. Three studies have some concerns related to risk of bias
- Downgraded twice as 1. all studies but one are on the same side of the threshold but the CIs do not systematically overlap 2. $I^2 = 87.92\%$ and the Q-test concludes that the true outcomes appear to be heterogeneous.
- Downgraded once as the CI is quite large although $p = 0.05$ and sample size > 400 .
- Downgraded due to relatively wide variance of point estimates across the two effect sizes: one is large, the other small b) No overlap of confidence intervals (CI).
- Based on a single study with a final risk of bias score rated as medium for concerns with deviation from intended intervention.

- j. Downgraded due to wide variance of point estimates across the two effect sizes on both sides of the 0.00 threshold: -0.025 vs 0.086 b) No overlap of confidence intervals (CI).
- k. Downgraded because four of seven RoB criteria have been assessed as unclear: Randomness of assignment, performance bias, outcome measurement bias and reporting bias.
- l. Downgraded due to very wide range of CIs across the point estimates.
- m. Based on a single study with a final risk of bias score rated as medium. Four of seven RoB criteria have been assessed as unclear: Randomness of assignment, performance bias, outcome measurement bias and Reporting bias.
- n. Downgraded due to a wide range of point estimates across both sides of the threshold and CIs that do not overlap.
- o. Downgraded due to a wide range of point estimates across both sides of the threshold and CIs that do not overlap.
- p. Downgraded once because two studies are of high risk of bias. That said, it should be noted that those two studies only contribute a combined 15% of the weight of the meta-analysis. Other studies are either of uncertain or low risk.
- q. Based on a single study with high risk of bias due to not addressing selection bias to a satisfactory extent.
- r. Downgraded once as although the three studies are RCTs and two score low risk of bias one has a high risk of bias
- s. Downgraded because, despite being all positive, the effect estimates vary widely and have non-overlapping CIs.
- t. Downgraded because one study presents high risk of bias. Additionally, a study of uncertain risk of bias is weighted to contribute 73% to the meta-analysis.
- u. Downgraded twice because > 80% of the meta-analysis is coming from high risk of bias studies. If those studies are removed, the group is left with one QED of uncertain bias.
- v. Downgraded to account for the wide range of CIs across the three studies, crossing both sides of the threshold.
- w. Downgraded once since out of the 2 studies both of them are RCTs but one is medium and the other one is low risk of bias.
- x. Downgraded once 1. both studies are on the same side of the threshold but the CI are not overlapping 2. $I^2 = 97.55\%$ but there was no significant amount of heterogeneity in true outcomes.
- y. Downgraded once since $p = 0.24$ although sample size > 400.
- z. Based on a single study with a final risk of bias score rated as high. Deviations from intended intervention is a plausible risk.
- aa. Downgraded once as although the three studies are RCT and 2 score low risk of bias one has a high risk of bias.
- ab. downgraded twice 1. All the studies are on the same side of the threshold but the CI are not systematically overlapping 2. $I^2 = 95.468\%$ and Q-test concludes that the true outcomes appear to be heterogenous.
- ac. Downgraded once as the CI is very large and $p = 0.08$ although sample size > 400.

4.3.2 Cash transfers

Cash transfers are a form of social protection intervention that provide cash to vulnerable households or individuals to reduce vulnerability, poverty, and risk and to improve the human capital of the poor. They allow beneficiary communities, households and/or individuals to protect themselves from economic, demographic and seasonal shocks, improve nutrition and food security, and increase asset ownership and expenditure on basic necessities (Miller et al. 2008). They are interventions that either require beneficiaries to comply with specific conditions to be eligible for a money or cash transfer that will allow them to respond to specific needs of resources, or do not require any specific actions to be undertaken by targeted beneficiaries to be eligible for a money or cash transfer that will allow them to respond to specific needs of resources (Hemsteede 2018).

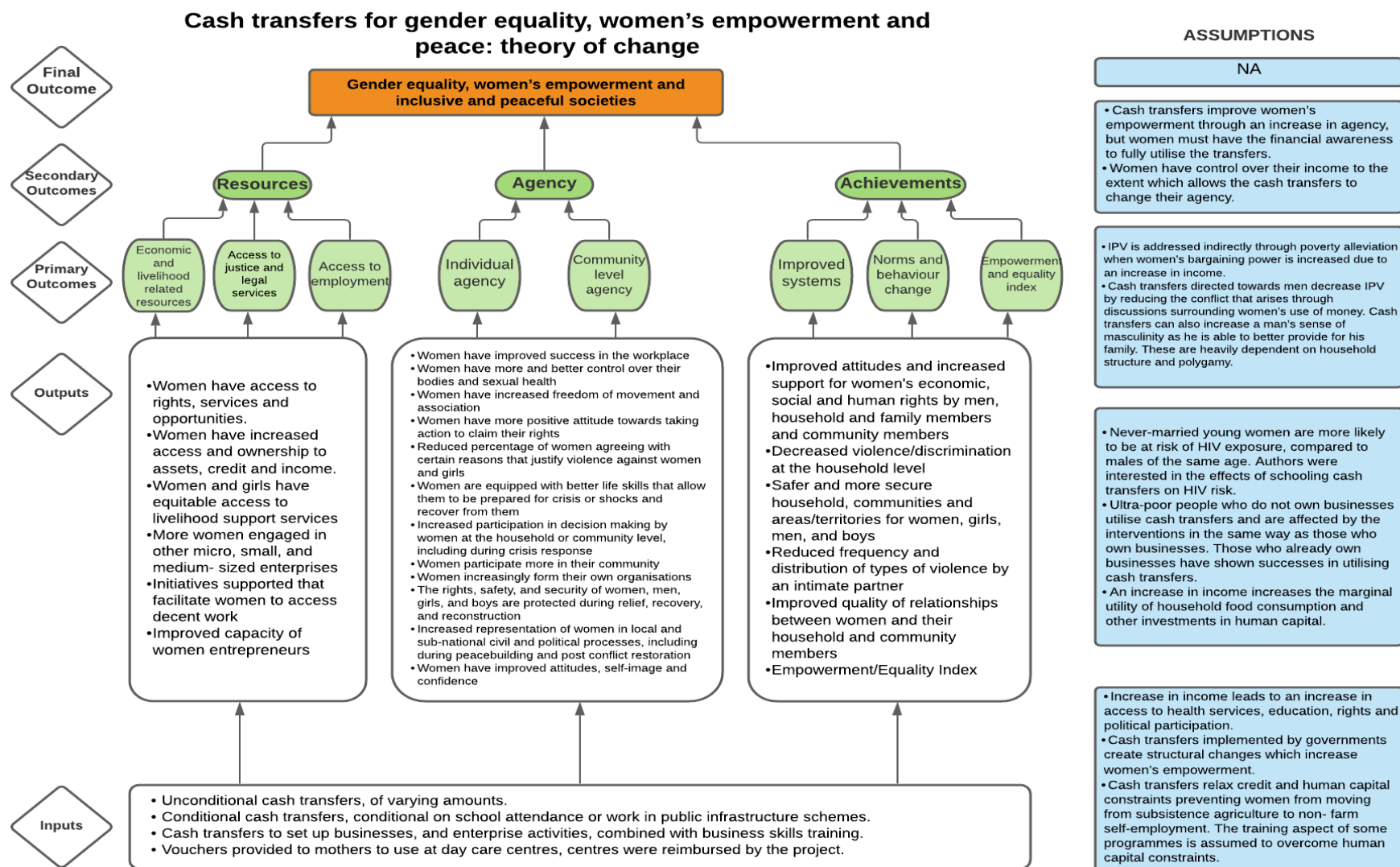
How do cash transfers affect gender equality, women's empowerment and peace outcomes?

Figure 33 below maps out the causal chain of how cash transfers may improve gender equality, women's empowerment and peace outcomes.

Cash transfers can work under two separate ToCs, depending on whether the transfer is directed towards the man or woman of the household. When cash transfers, conditional or unconditional, are directed towards the woman within the household, this can lead to two changes. First, the increase in income can empower women by increasing their agency and reducing their reliance on their husbands. An increase in their bargaining power within the household leads to greater agency. Secondly, cash transfer programmes lift two constraints faced by women. Firstly, it lifts the financial constraint which prevents women from moving from subsistence agriculture to non-farm self-employment. Secondly, programmes which incorporate business skills training lift the human capital constraints faced by women where they do not have the required knowledge to utilise the transfers to set up a business.

Cash transfers directed towards men work on a different theory of change. The increase of income can lead to a decrease in conflict between the man and woman within the house regarding the discussion and use of money, specifically when it comes to women asking for money. The increase of income can also increase the man's dominance as he is able to more easily provide for his household. The decrease in household conflict can lead to greater empowerment for women.

Figure 33: Cash transfers for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included 21 studies reported in 29 different papers evaluating the effect of 12 programmes. We included more than one paper that evaluated the same programme if the author(s) reported different outcomes over several papers.

Population

As previously mentioned, cash transfers can be targeted towards men or women as the main beneficiaries. The Benazir Income Support Programme (BISP) specifically targeted women in households which scored under a certain level on a proxy means test. Conditional cash transfers, conditional on school attendance target households with children, while in general, all transfers target the poor and ultra-poor.

The included studies evaluated 12 different programmes and trials in MENA, South Asia and Sub-Saharan Africa, a total of 11 countries including: Burkina Faso, Cameroon, Egypt, Ethiopia, Kenya, Malawi, Mali, Pakistan, South Sudan, Uganda and Zambia.

Intervention, inputs and activities

The included studies evaluated a range of different cash transfer activities and inputs including:

- Unconditional cash transfers (n = 13): Any programme that provided an unconditional cash transfer. These transfers are a fixed, but varying, amount and are not conditional on any requirements.
- Conditional cash transfers (n = 5): Any programme that provided a conditional cash transfer. This is a sum of money in exchange for meeting certain requirements such as a child's school attendance or participation in a public infrastructure work scheme.
- Day care vouchers (n = 1): Day care vouchers were provided to mothers who were able to use them to send their child to day care. The day care centres were reimbursed for the vouchers.

Table 19: Features of included cash transfer studies

Study	Activity/input	Length of treatment	Frequency of Cash Transfer
Ambler - 2017	Cash transfer of 1,000 PKR per month, paid in quarterly instalments of 3,000 PKR	Unclear	Quarterly
Baird - 2010	Conditional cash transfer of varying amounts, conditional on school attendance, as well as an unconditional cash transfer of varying amounts.	24 months	Monthly
Baird - 2019	Conditional cash transfer of varying amounts, conditional on school attendance, as well as an unconditional cash transfer of varying amounts.	24 months	Monthly
Baird - 2009	Conditional cash transfer of varying amounts, conditional on school attendance, as well as an unconditional cash transfer of varying amounts.	24 months	Monthly

Study	Activity/input	Length of treatment	Frequency of Cash Transfer
Baird - 2013	Conditional cash transfer of varying amounts, conditional on school attendance, as well as an unconditional cash transfer of varying amounts.	24 months	Monthly
Blattman – 2013	\$150 USD grants, as well as business skills training and planning, and on-going supervision to implement the business plan.	24 months	One Time Transfer
Blattman – 2016	\$150 USD grants, as well as business skills training and planning, and on-going supervision to implement the business plan.	24 months	One Time Transfer
Bonilla - 2017	\$24 USD bi-monthly cash transfers	48 months	Bi-Monthly
Breisinger – 2018	Cash transfers, the amount of which is dependent on the number of children and their schooling level.	Unclear	Monthly
Brooks - 2020	One-time cash transfer of 5,000 KES.	½ month	One Time Transfer
Clark - 2019	Mothers were offered vouchers to be used at day care centres, day care centres were also offered improvements.	14 months	Monthly
Eze - 2019	Distribution of cash transfers which totalled 654.1 billion FCFA.	96 months	Not Reported
Gelagay - 2019	Provision of a conditional cash transfer based on work in a public infrastructure scheme.	Unclear	Monthly
Gobin - 2017	\$100 USD cash transfer to set up a microenterprise, alongside business skills training. An additional \$50 USD transfer for support conditional on the microenterprise having been set up.	14 months	One time transfer with one additional transfer for qualifying participants.
Green - 2015	Business start-up grant of \$150 USD, as well as business skill training.	Unclear	One Time Transfer
Green - 2016	Business start-up grant of \$150 USD, as well as business skill training.	Unclear	One Time Transfer
Handa – 2015	Unconditional cash transfer of 1500 Kenyan shillings per household.	48 months	Monthly

Study	Activity/input	Length of treatment	Frequency of Cash Transfer
Haushofer – 2019	Unconditional cash transfer of \$709 USD per beneficiary.	24 months	Monthly
Haushofer 2013	Unconditional cash transfer of \$709 USD per beneficiary.	24 months	Monthly
Haushofer 2016	Unconditional cash transfer of \$709 USD per beneficiary.	24 months	Monthly
Haushofer 2018	Unconditional cash transfer of \$709 USD per beneficiary.	24 months	Monthly
Heath 2020	Cash transfers, group training sessions on different themes and nutrition packages.	36 months	Quarterly
Iqbal - 2020	Conditional cash transfers, the pilot also consisted of provision of training, microfinance and health insurance.	132 months	Quarterly
Iqbal - 2020	Conditional cash transfers, the pilot also consisted of provision of training, microfinance and health insurance.	132 months	Quarterly
Karimli - 2019	Four components: 1) savings, 2) technical skills training, 3) cash transfers with livelihood selection and 4) ongoing support.	24 months	One Time Transfer
Karimli - 2021	Four components: 1) savings, 2) technical skills training, 3) cash transfers with livelihood selection and 4) ongoing support.	24 months	One Time Transfer
Malaeb - 2017	Business start-up grant of \$150 USD, as well as business skill training.	Unclear	One Time Transfer
Mekonnen – 2017	Cash transfers to implement innovative agricultural production systems.	24 months	Unclear
Müller - 2019	Unconditional cash grant with business skill training.	14 months	Unclear
Natali - 2016	Unconditional cash transfer of \$12 USD per month.	34 months	Monthly

Comparison

All included studies compared treated groups to comparison groups receiving no intervention. Nine studies included multiple treatment arms.

Outcomes

The included studies reported on several relevant outcomes:

- Women have increased access and ownership to assets, credit and income (n = 13): Women can apply for, receive and manage assets/credit and income and have support to manage, claim and execute their assets without pressure or influence from external actors, including male family members, husbands and cultural leaders.
- Initiatives that facilitate women to access decent work (formal and informal employment), including people with disabilities (n = 7): Women are able to apply for, receive and work in jobs (and have support for the above), without discrimination for sex, gender or other identifying factors, including development of skills for improved access.
- Women have increased freedom of movement and association (n = 5): Women and girls can move freely and meet within their household, community, and area/territory without fear of attack or discrimination.
- Increased participation in decision making by women at the household or community level, including during crisis response (n = 11): Women take part in all or any step of the decision-making process at the household community and district level, but also are able to meaningfully take part and have influence on the final decision, including in crisis response.
- Women participate more in their community (n = 3): Women increase their influence and representation in the life of their community. This can include increased participation in local/district government, cultural institutions and representation in various sectors and having a greater influence and involvement in the activities and pursuits of the community, district, etc.

The majority of studies in this category (n = 14) measured economic and livelihood related resources through women having increased access and ownership to assets, credit and income and women and girls having equitable access to livelihood support services.

The majority of studies (n = 12) measured community level agency through increased participation in decision making by women at the household or community level, including during crisis response; women participate more in their community; women increasingly form their own organisations, and the rights, safety, and security of women, men, girls, and boys are protected during relief, recovery, and reconstruction.

The majority of studies (n = 11) measured individual agency through women have improved success in the workplace; women have more and better control over their bodies and sexual health; women have increased freedom of movement and association; women have more positive attitude towards taking action to claim their rights; reduced percentage of women agreeing with certain reasons that justify violence against women and girls and women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them.

The majority of studies (n = 11) also measured norms and behaviour change through women have improved attitudes, self-image and confidence; improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members; decreased violence/discrimination at the

household level; safer and more secure household, communities and areas/territories for women, girls, men, and boys; reduced frequency and distribution of types of violence by an intimate partner and improved quality of relationships between women and their household and community members.

The majority of studies (n = 7) measured access to employment through more women engaged in other micro, small, and medium-sized enterprises; initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities and improved capacity of women entrepreneurs.

Studies (n = 5) also measured empowerment/equality index through empowerment/Equality Index.

Other outcomes measured included: (n = 1) measured improved systems through increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration; (n = 1) measured access to justice and legal services through women have access to rights, services and opportunities.

Our list of 44 outcomes reported in our included studies was grouped into nine immediate outcomes directly contributing to the three secondary outcomes as per Kabeer's dimensions of women's empowerment and gender equality (Resources, Agency and Achievement). The division of the immediate and secondary outcomes is reported in Table 20 below:

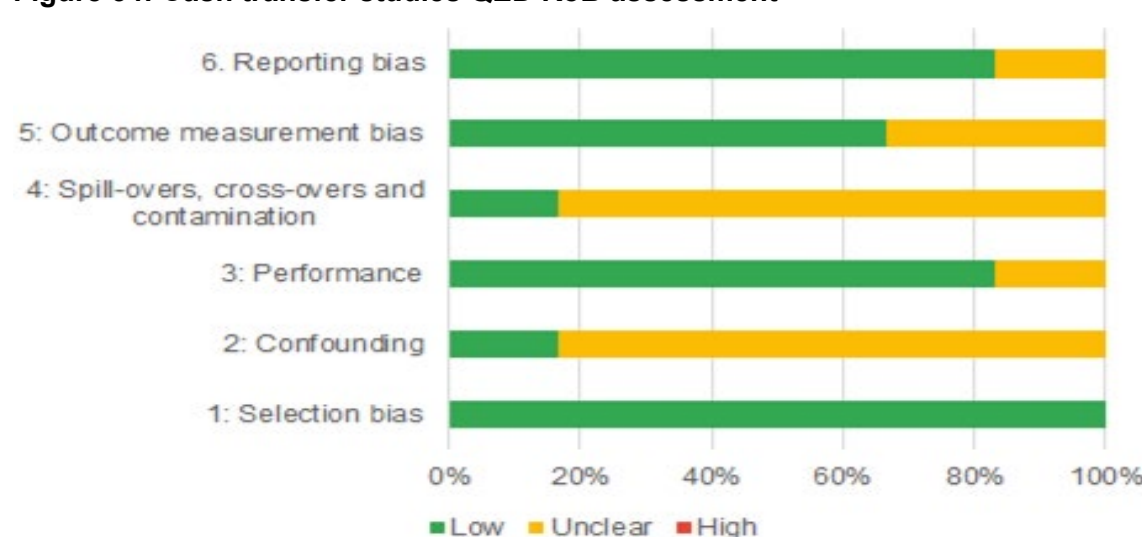
Table 20: Summary of secondary and immediate outcomes for cash transfers

Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	0
	Economic and livelihood related resources	13
	Access to employment	7
Agency ability to define one's goals and act upon them and operationalised decision-making;	Individual agency	12
	Community level agency	13
	Institutions supporting agency	1
Achievement ways of being and doing which can be realised by different individuals	Improved systems	3
	Norms and behaviour change	11
	Empowerment index	4

Study design

Four of our cash transfer studies used a QED (Ambler 2017; Brooks 2020; Eze 2019; Iqbal 2020). All four studies were assessed as having some concerns related to risk of bias. As detailed in Figure 34 below, high risk of bias was not identified in any RoB domains. However, the quality assessment of the studies was limited by the lack of discussion by the authors regarding potential issues with the study design that could have resulted in confounding of the results; not enough information was provided about data collection and treatment implementation processes, so that issues about implementation, such as spill-overs and crossovers, and about outcome measurement could not be discarded. This was the case in particular of Gelagay (2019) and Karimli (2019). We could not identify any limitation related to the assignment of the intervention and selection of the comparison group in any of the included studies, so that all included studies were assessed as low risk of selection bias.

Figure 34: Cash transfer studies QED RoB assessment



Twenty of the included cash transfer studies used an experimental design. This included Aker (2016), Baird (2009, 2010, 2013 and 2019), Blattman (2013 and 2016), Bonilla (2017), Clark (2019), Gobin (2017), Green (2015), Haushofer (2013, 2016, 2018 and 2019), Heath (2020), Karimli (2021), Malaeb (2017), Mekonnen (2017), Muller (2019) and Natali (2016).

Approximately 25% of the papers ($n = 5$) were assessed as having a high risk of bias and 30% ($n = 6$) as having some concerns, but we did not identify any major limitation for 45% of the included RCT papers evaluating cash transfers ($n = 9$). As detailed in Figure 35 below, the limitations that were associated with high risk of bias were identified in five RoB domains (unit of analysis, selection, deviation from intended intervention, outcome measurement and reporting), yet in all categories the number of papers presenting high risk or being unclear were four or less.

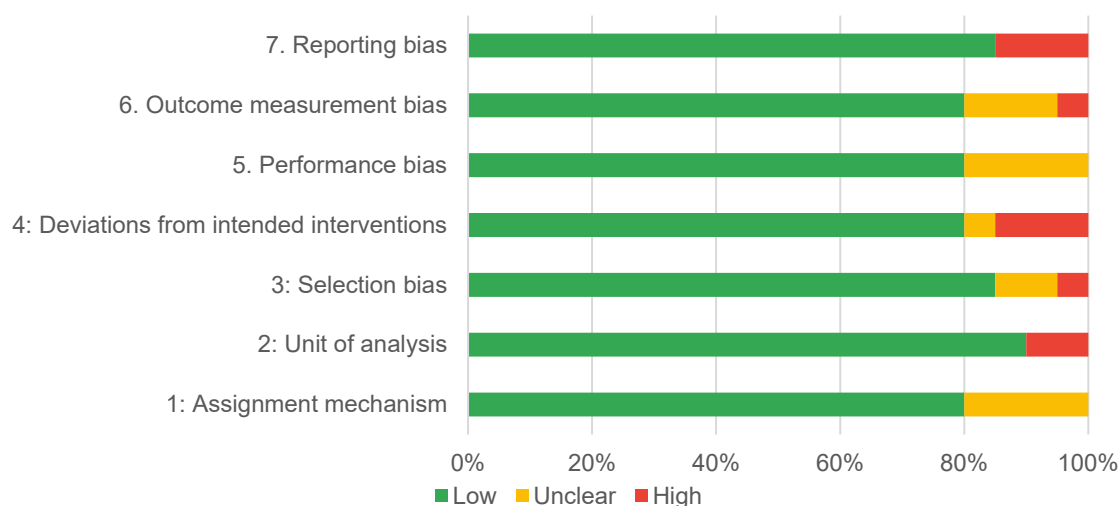
The most common issues in the RCT studies evaluating cash transfers were due to deviations from intended interventions and reporting bias. Deviations from intended interventions can occur when there are implementation issues that might lead the control participants to receive the treatment, as when vouchers delivered can be traded or transferred (Clark 2019) or when there is a risk of contamination because participants

could not be isolated from other programs (Karimli 2021). Reporting bias is observed when authors of RCT studies do not discuss baseline balance between treatment and control groups (e.g. Karimli 2021; Malaeb 2017), or when multiple treatment arms are not differentiated in the analysis (Clark 2019).

Other limitations that we identified in RCT studies evaluating cash transfers were the fact that a few had assigned treatment at the cluster level (e.g., village) but had analysed data at the individual level without adjusting for clustering (Clark 2019; Malaeb 2017); another study shows a large difference in attrition rates between treatment and control groups, which poses a risk of selection bias given that authors could not rule out that it was related to the intervention assignment (Clark 2019); further, one study was scored as high risk of bias because of measurement issues related to the use of self-report outcome data which is susceptible to recall and reporting bias (Blattman 2013).

The most common issues in the RCT studies evaluating cash transfers were due to deviations from intended interventions and reporting bias. Deviations from intended interventions can occur when there are implementation issues that might lead the control participants to receive the treatment, as when vouchers delivered can be traded or transferred (Clark 2019) or when there is a risk of contamination because participants could not be isolated from other programs (Karimli 2021). Reporting bias is observed when authors of RCT studies do not discuss baseline balance between treatment and control groups (e.g. Karimli 2021; Malaeb 2017), or when multiple treatment arms are not differentiated in the analysis (Clark 2019).

Figure 35: Cash transfer studies RCT RoB Assessment



Qualitative studies, process evaluations and project documents

We identified 25 additional documents related to nine programmes covered by the cash transfer studies.

- Child Grant Programme (Zambia): two process evaluations and one descriptive quantitative document
- Productive Safety Net Programme (PSNP) (Ethiopia): one qualitative and two descriptive documents
- Zomba Cash Transfer Programme (ZCTP) (Malawi): two descriptive quantitative documents

- Takaful and Karama programme (Egypt): three qualitative documents
- Benazir Income Support Programme (BISP) (Pakistan): two descriptive quantitative and one qualitative document
- Unnamed Cash Transfer Programme (Kenya): one qualitative document
- Women's Income Generating Support (WINGS) (Uganda): four descriptive quantitative documents
- Rural Entrepreneur Access Program (Kenya): four descriptive quantitative documents
- Trickle Up (Burkina Faso): two descriptive quantitative documents

The majority (n = 14) of studies under this intervention group were of moderate empirical quality, followed by high quality (n = 6) and low quality (n = 5).

Synthesis of findings

The following subsection presents the results of effectiveness of cash transfers on gender equality, women's empowerment and peace outcomes.

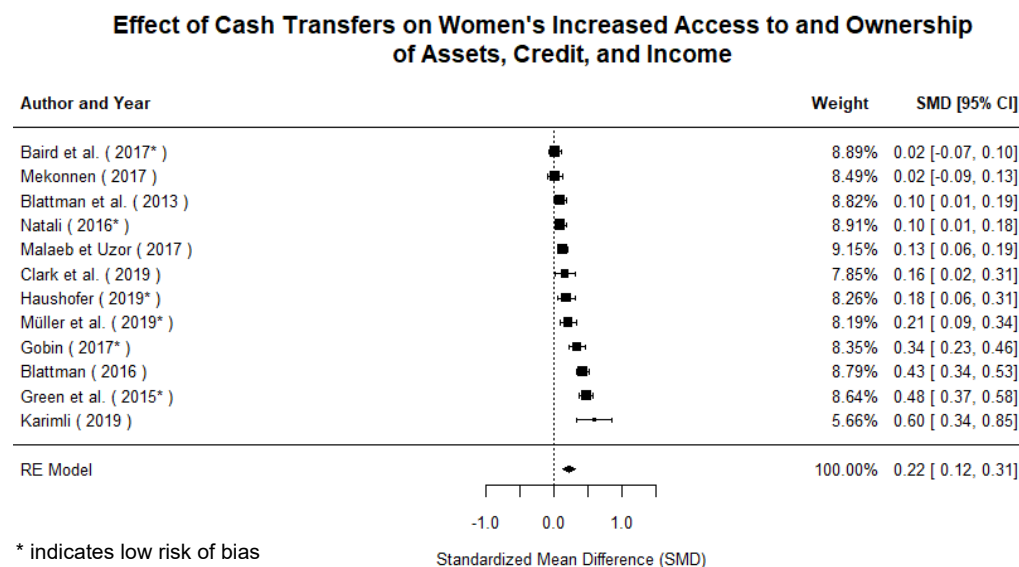
Quantitative findings

Effects of cash transfers on women's access to and ownership of assets, credit, and income

We included a total of $k = 12$ studies in the analysis. We assessed six studies as low risk of bias, three as some concerns, and three as high risk of bias. The observed outcomes ranged from 0.02 to 0.60. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.22$ (95% CI: 0.12 to 0.31). Therefore, the average outcome differed significantly from zero ($z = 4.50$, $p < 0.01$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 36 CTAB2a.

According to the Q -test, the true outcomes appear to be heterogeneous ($Q(11) = 109.31$, $p < 0.01$, $\hat{\tau}^2 = 0.02$, $I^2 = 89.94\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.87 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

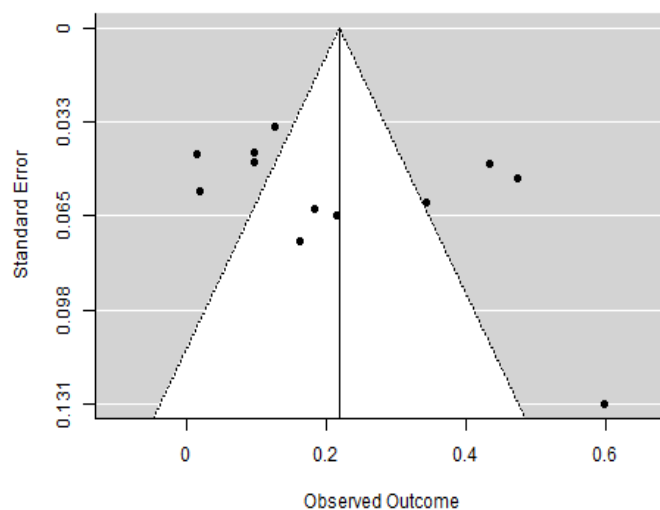
Figure 36: CTAB2a: Forest plot showing the observed outcomes and the estimate of the random-effects model



We were able to test several moderators for this outcome category. First, we tested whether effects from studies using conditional cash transfers differed from those using unconditional cash transfers. We found no difference between the two intervention conditions ($\hat{B} = 0.009, p = 0.94$ [95% CI: -0.20 to 0.22]). Studies with a high risk of bias were not significantly different than studies with some concerns or low risk of bias ($\hat{B} = -0.12, p = 0.29$ [95% CI: -0.35 to 0.10]). In addition, there was no difference in the size of effects between programmes that used conditional versus unconditional cash transfers ($\hat{B} = -0.02, p = 0.80$ [95% CI: -0.24 to 0.19]). There was also no moderation by the length of exposure to the intervention ($\hat{B} = -0.003, p = 0.32$ [95% CI: -0.01 to 0.003]), nor the length of the evaluation follow-up period ($\hat{B} = -0.004, p = 0.32$ [95% CI: -0.01 to 0.004]). Finally, publication year was also not a significant moderator ($\hat{B} = 0.006, p = 0.83$ [95% CI: -0.05 to 0.07]). Study design could not be tested as a moderator because all of the studies were experimental designs, and we also could not test for differences between adjusted and unadjusted models because only one study presented an unadjusted effect.

A funnel plot of the estimates is shown in Figure 37 CTAB2b. Neither the rank correlation nor the regression test indicated any funnel plot asymmetry ($p = 0.25$ and $p = 0.05$, respectively).

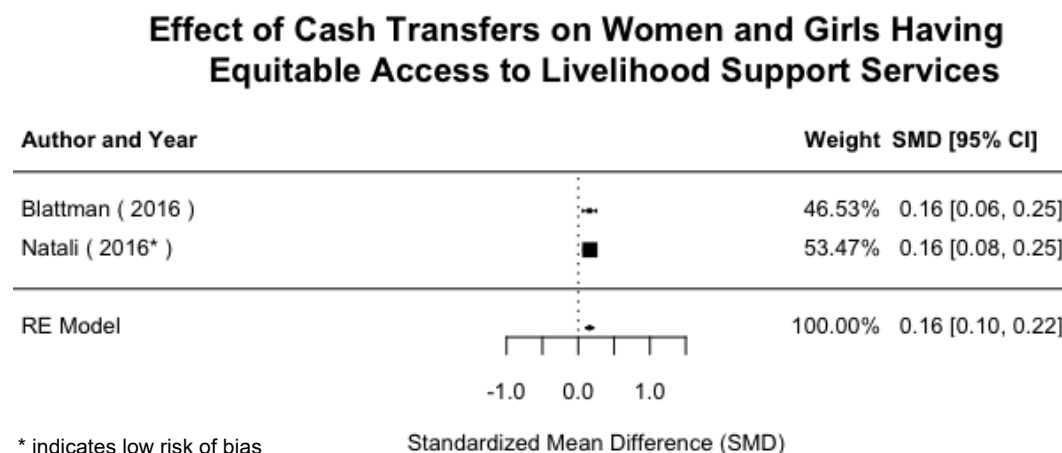
Figure 37: CTAB2b: Funnel plot



Effects of cash transfers on women and girls having equitable access to livelihood support services

Only two studies reported the impact of cash transfers on equitable access to livelihood support services, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as low risk of bias and the other as some concerns. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.16$ (95% CI: 0.10 to 0.22). This positive effect differed significantly from zero ($z = 4.96, p < 0.01$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 38 CTAB3. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.00, p = 0.94, \hat{\tau}^2 = 0.00, I^2 = 0.00\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid.

Figure 38: CTAB3. Forest plot showing the observed outcomes and the estimate of the random-effects model



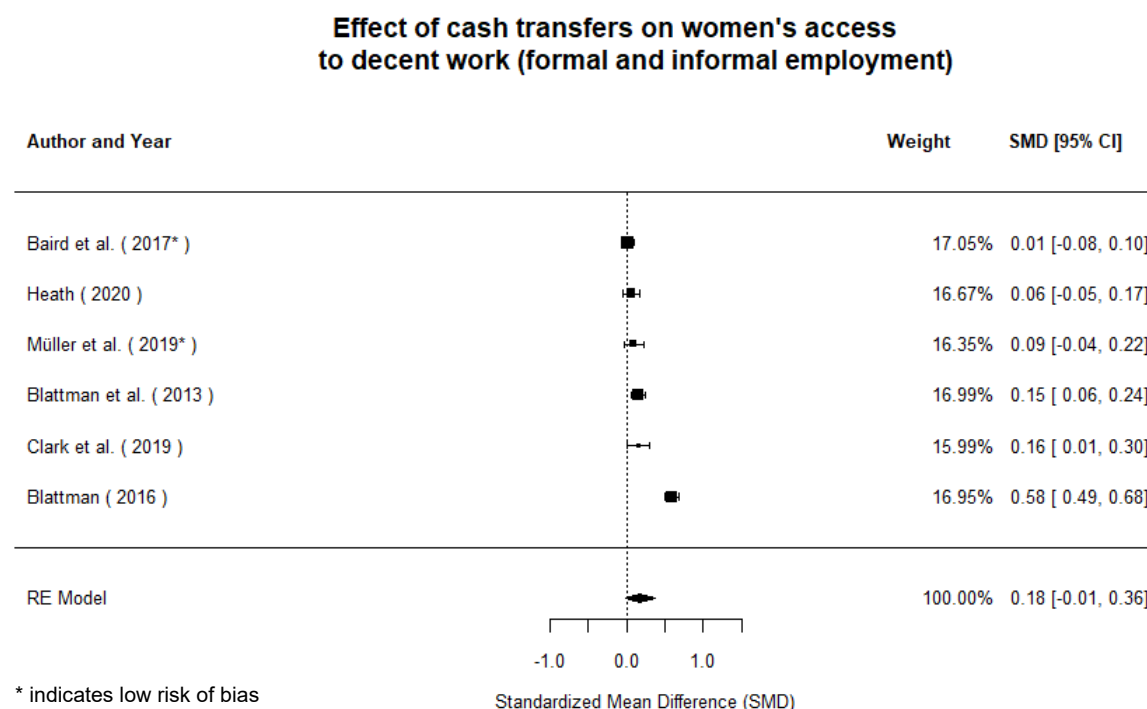
Effects of cash transfers on more women engaged in other micro, small and medium-sized enterprises

Only two studies reported the impact of cash transfers on more women engaged in other micro, small, and medium-sized enterprises. These studies were linked, both examining the impact of WINGS in Uganda. Both studies reported large positive and significant impacts after both 6 months of exposure to the intervention $g = 0.91$ (95% CI: 0.71 to 0.92) and 16 months of exposure to the intervention $g = 0.81$ (95% CI: 0.82 to 1.01). We assessed one study as having low risk of bias and the other as some concerns.

Effects of cash transfers on initiatives that help women access decent work (formal and informal employment)

We included a total of $k = 6$ studies in the analysis. We assessed two studies as low risk of bias, two as some concerns, and two as high risk of bias. The observed outcomes ranged from 0.01 to 0.58. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.18$ (95% CI: -0.01 to 0.36). Therefore, the average outcome did not differ significantly from zero ($z = 1.86$, $p = 0.06$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 39 CTAC2.

Figure 39: CTAC2: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(5) = 93.61$, $p < 0.01$, $\hat{\tau}^2 = 0.05$, $I^2 = 94.66\%$). An examination of the studentised residuals revealed that one study (Blattman 2016) had a value larger than ± 2.69 and may be a potential outlier in the context of this model. According to the Cook's distances, one study (Blattman 2016) could be considered to be overly influential. Indeed, sensitivity analyses leaving each study out indicated that removing Blattman (2016) would reduce the overall average effect ($\hat{\mu} = 0.08$ (95% CI: 0.03 to 0.15), but the effect would still be positive and would become statistically significant ($z = 2.86$, $p = 0.004$).

Effect from studies with high risk of bias were not different than effects from studies with some concerns or low risk of bias ($\hat{\beta} = -0.03$, $p = 0.88$ [95% CI: -0.43 to 0.37]). There was also no moderation by exposure to intervention ($\hat{\beta} = -0.01$, $p = 0.41$ [95% CI: -0.03 to 0.01]) nor by evaluation period ($\hat{\beta} = -0.002$, $p = 0.71$ [95% CI: -0.01 to 0.01]). Publication year was also not a significant source of variation among effects ($\hat{\beta} = -0.03$, $p = 0.49$ [95% CI: -0.10 to 0.05]). We were not able to test for moderation by study design (all RCTs), post-intervention versus change from baseline, or whether the model was adjusted for covariates (all were adjusted). In addition, only one effect was conditional, so we were unable to test for differences between conditional and unconditional transfers.

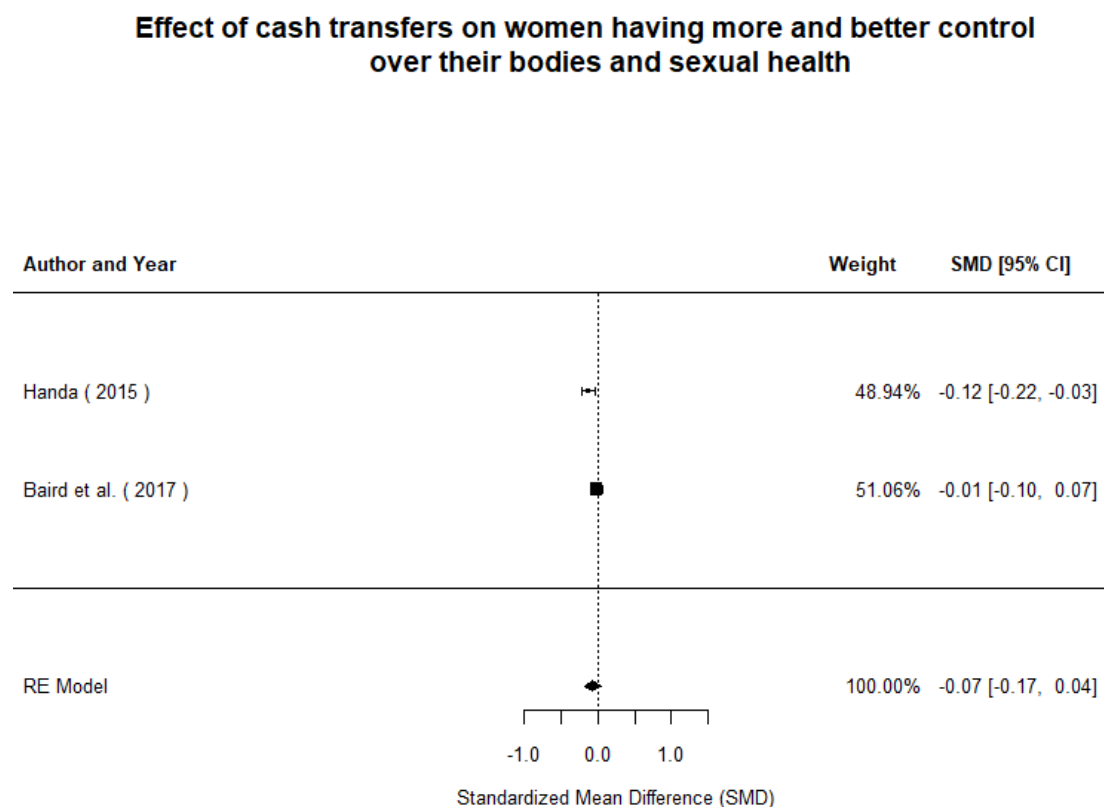
Effects of cash transfers on improved capacity of women entrepreneurs

Blattman and colleagues' (2013) experimental study in Uganda was the only study evaluating the impact of cash transfers on improved capacity of women entrepreneurs. There was a large, and statistically significant, effect ($g = 0.40$, [95% CI: 0.32 to 0.49]), but we assessed the study as having high risk of bias.

Effects of cash transfers on women have more and better control over their bodies and sexual health

Only two studies reported the impact of cash transfers on improved life skills for women, thus we included $k = 2$ studies in the analysis. We assessed one study as having low risk of bias (Baird et al. 2017) and the other as some concerns related to risk of bias (Handa 2015). The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.07$ (95% CI: -0.17 to 0.04). Therefore, the average outcome did not differ significantly from zero ($z = -1.22$, $p = 0.22$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 40 CTBA2. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 2.90$, $p = 0.09$, $\hat{\tau}^2 = 0.004$, $I^2 = 65.56\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid.

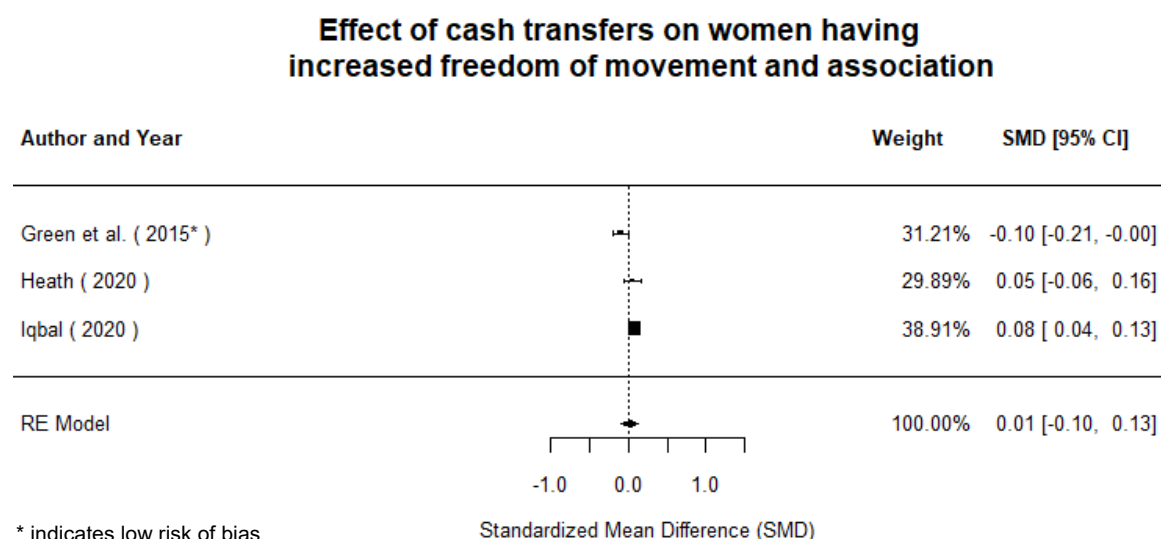
Figure 40: CTBA2: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of cash transfers on women have increased freedom of movement and association

We included a total of $k = 3$ studies in the analysis. We assessed one study as low risk of bias, two as some concerns, and none as high risk of bias. The observed outcomes ranged from -0.10 to 0.08 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.01$ (95% CI: -0.10 to 0.13). Therefore, the average outcome did not differ significantly from zero ($z = 0.25$, $p = 0.80$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 41 CTBA3.

Figure 41: CTBA3: Forest plot showing the observed outcomes and the estimate of the random-effects model

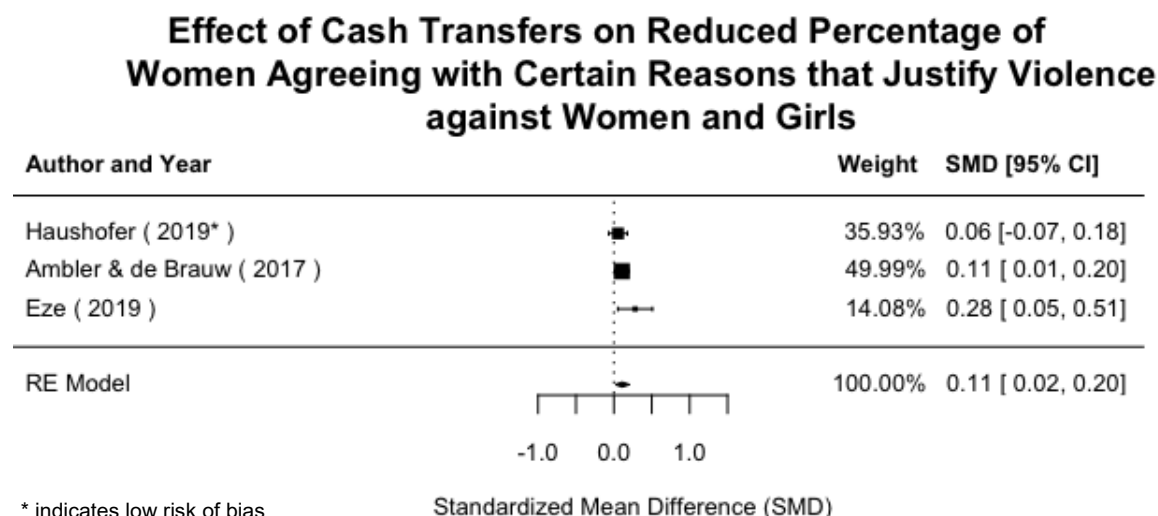


According to the Q -test, the true outcomes appear to be heterogeneous ($Q(2) = 11.02$, $p < 0.01$, $\hat{\tau}^2 = 0.01$, $I^2 = 81.85\%$). An examination of the studentised residuals revealed that one study (Green et al. 2015) had a value larger than ± 2.39 and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Green and colleagues (2015) would increase the overall average effect ($\hat{\mu} = 0.08$ (95% CI: 0.04 to 0.12), and the effect would still be positive, but would be significant ($z = 3.72$, $p < .001$). According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

Effects of cash transfers on effect of cash transfers on reduced percentage of women agreeing with certain reasons that justify violence against women and girls

We included a total of $k = 3$ studies in the analysis. We assessed one study as low risk of bias, two as some concerns, and none as high risk of bias. The observed outcomes ranged from 0.06 to 0.28, with all estimates being positive. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.11$ (95% CI: 0.02 to 0.20). Therefore, the average outcome differed significantly from zero ($z = 2.38$, $p = 0.02$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 42 CTBA6.

Figure 42: CTBA6: Forest plot showing the observed outcomes and the estimate of the random-effects model

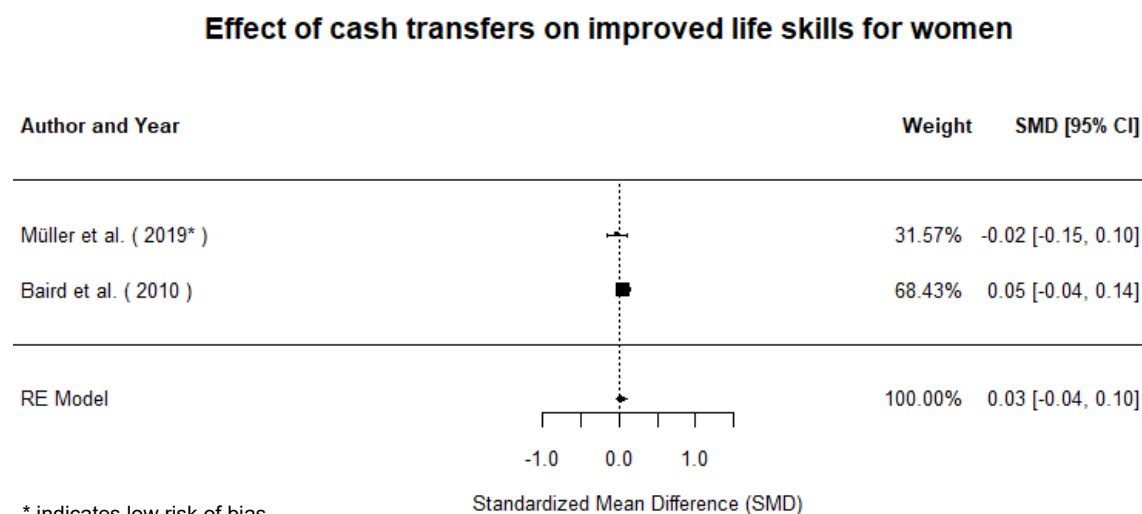


According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(2) = 2.93$, $p = 0.23$, $\hat{\tau}^2 = 0.00$, $I^2 = 31.64\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.39 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

Effects of cash transfers on improved life skills for women

Only two studies reported the impact of cash transfers on improved life skills for women, thus we included $k = 2$ studies in the analysis. One study was assessed as having low risk of bias, the other as having high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.03$ (95% CI: -0.04 to 0.10). Therefore, the average outcome did not differ significantly from zero ($z = 0.74$, $p = 0.46$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 43 CTBA7. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.89$, $p = 0.46$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid.

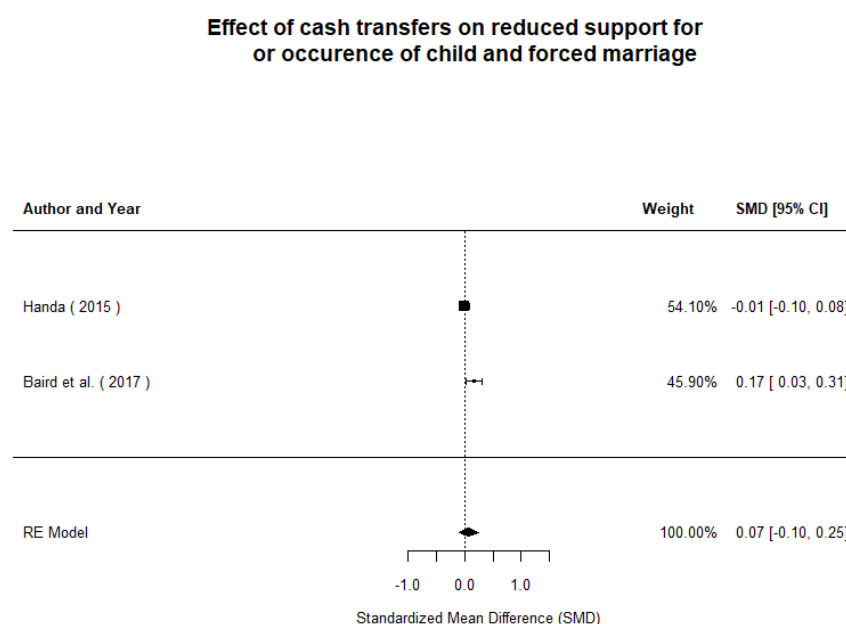
Figure 43: CTBA7: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of cash transfers on reduced support for, or instances of, child or forced marriage

Only two studies reported the impact of cash transfers on reduced support for, or instances of, child or forced marriage, thus we included $k = 2$ studies in the analysis. We assessed one study as having low risk of bias (Baird et al. 2017) and the other as some concerns (Handa 2015). The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.07$ (95% CI: -0.10 to 0.25). Therefore, the average outcome did not differ significantly from zero ($z = 0.80$, $p = 0.42$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 44 CTBA8. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 4.59$, $p = 0.03$, $\hat{\tau}^2 = 0.01$, $I^2 = 78.22\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid.

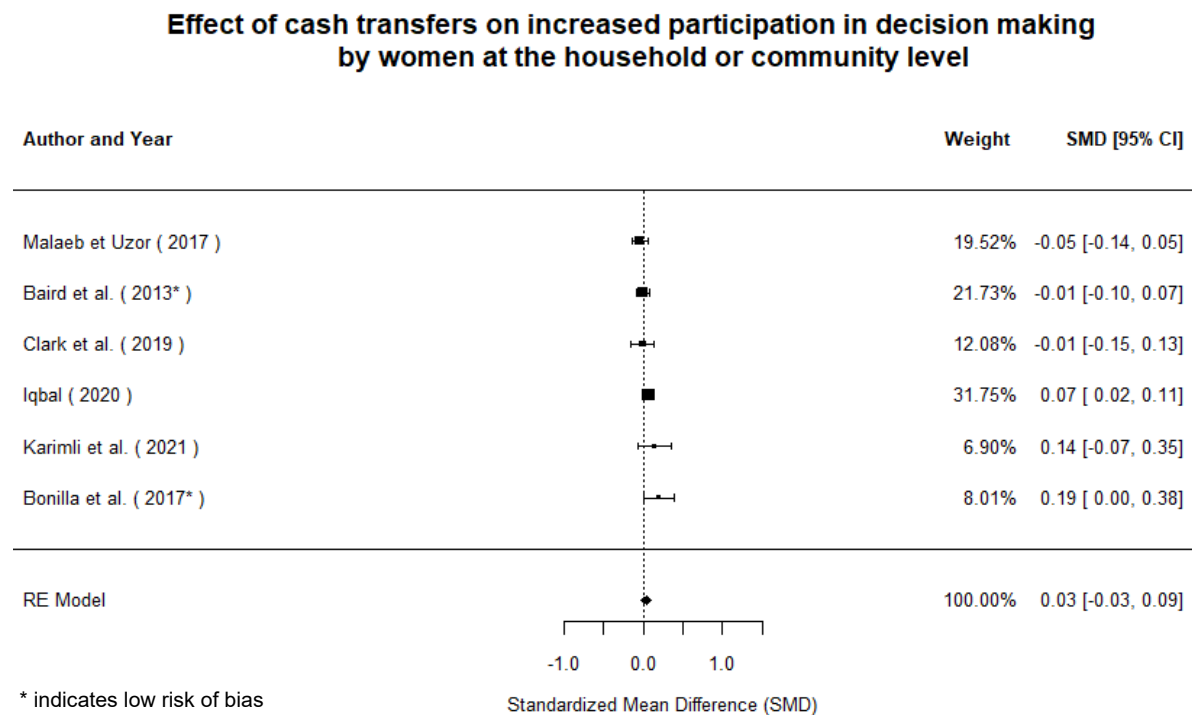
Figure 44: CTBA8: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of cash transfers on increased participation of women in decision making at the household or community level

We included a total of $k = 6$ studies in the analysis. We assessed two studies as low risk of bias, one as some concerns, and three as high risk of bias. The observed outcomes ranged from -0.05 to 0.19 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.03$ (95% CI: -0.03 to 0.09). Therefore, the average outcome did not differ significantly from zero ($z = 1.11$, $p = 0.27$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 45 CTBB1.

Figure 45: CTBB1: Forest plot showing the observed outcomes and the estimate of the random-effects model



The Q -test for heterogeneity was not significant, but some heterogeneity may still be present in the true outcomes ($Q(5) = 9.79$, $p = 0.08$, $\hat{\tau}^2 = 0.00$, $I^2 = 48.92\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.77 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. If the high risk of bias studies were removed, the average effect increases slightly and is still significantly different from zero.

For decision making outcomes, effects from conditional cash transfers were smaller than effects from unconditional cash transfers by .10 standard deviation units, a small but significant difference ($\hat{B} = -0.10$, $p = 0.01$ [95% CI: -0.18 to -0.03]). Effect from studies with high risk of bias were not different than effect from studies with some concerns or low risk of bias ($\hat{B} = -0.06$, $p = 0.35$ [95% CI: -0.18 to 0.06]). There was also no moderation by exposure to intervention ($\hat{B} = 0.001$, $p = 0.44$ [95% CI: -0.001 to 0.004]). Publication year was also not a significant source of variation among effects ($\hat{B} =$

0.01, $p = 0.20$ [95% CI: -0.01 to 0.03]). Effects did not differ by whether the study was examining post intervention changes versus changes from baseline ($\hat{\beta} = 0.07$, $p = 0.30$ [95% CI: -0.06 to 0.20]) nor by whether or not the model was adjusted for covariates ($\hat{\beta} = -0.06$, $p = 0.28$ [95% CI: -0.16 to 0.04]). We were not able to test for moderation by study design (only one study used a QED), nor by evaluation period (all studies examined effects immediately after exposure to intervention).

Effects of cash transfers on increased on women participate more in their community.

We included a total of $k = 3$ studies in the analysis. None of the studies were appraised as low risk of bias, and so these findings should be applied cautiously. We assessed two studies as having some concerns of risk of bias and one as high risk of bias. The observed outcomes ranged from -0.07 to 0.26 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.10$ (95% CI: -0.07 to 0.28). Therefore, the average outcome did not differ significantly from zero ($z = 1.17$, $p = 0.24$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 46 CTBB2.

Figure 46: CTBB2: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(2) = 25.55$, $p < 0.01$, $\hat{\tau}^2 = 0.02$, $I^2 = 92.17\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.39 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

Effects of cash transfers on the rights, safety, and security of women, men, girls, and boys being protected during relief, recovery, and reconstruction

Brooks and colleagues' (2020) experimental study in Kenya was the only study evaluating the impact of cash transfers on the rights, safety, and security of women, men, girls, and boys being protected during relief, recovery, and reconstruction. Their report included two effects that fell into this outcome category (Personal protective equipment spending and protective measures). The effects were medium size and positive for both PPE spending ($g = 0.29$, 95% CI $[0.04, 0.53]$) and protective measures ($g = 0.36$, [95% CI: 0.12 to 0.61]). We assessed the study as having high risk of bias.

Effects of cash transfers on increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration

We included a total of three studies in Pakistan reporting on the impact of cash transfers on the representation of women in political processes (Ambler and de Brauw 2017; Iqbal 2020a; Iqbal 2020b). All three were studying the Benazir Income Support Programme (BISP) after different periods of exposure to the intervention (20 months, 72 months, and 90 months). We assessed all three studies as having some concerns of risk of bias. The observed outcomes ranged from $g = 0.07$ (95% CI: 0.02 to 0.11) to $g = 0.22$ (95% CI: 0.13 to 0.30), with the largest impact reported after 90 months of exposure to the intervention. All three reports were appraised as having some concerns related to risk of bias.

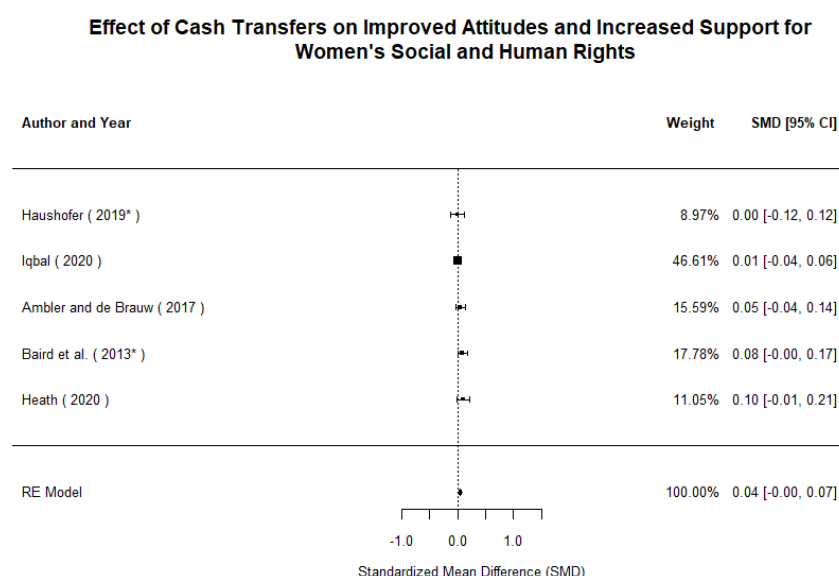
Effects of cash transfers on communities' attitudes towards women/marginalised groups

Blattman's (2016) experimental study in Uganda was the only study evaluating the impact of cash transfers on communities' attitudes towards women/marginalised groups. Without group training, there was a small, negative and statistically significant effect ($g = -0.11$, 95% CI $[-0.20, -0.01]$), and with group training there was a very small, positive but not statistically significant effect ($g = 0.02$, [95% CI: -0.08 to 0.11]). We assessed the study as having some risk of bias concerns.

Effects of cash transfers on improved attitudes and increased support for women's social, economic, and human rights by men, family member, household members and the community

We included a total of $k = 5$ studies that examined the impact of cash transfer on improved attitudes towards women. We assessed two studies as low risk of bias, three as some concerns, and none as high risk of bias. The observed outcomes ranged from 0.00 to 0.10. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.04$ (95% CI: -0.00 to 0.07). Therefore, the average outcome did not differ significantly from zero ($z = 1.88$, $p = 0.06$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 47 CTCB4.

Figure 47: CTCB4: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(4) = 4.27, p = 0.37, \hat{\tau}^2 = 0.00, I^2 = 6.43\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.58 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, one study (Iqbal 2020) could be considered to be overly influential. Indeed, sensitivity analyses leaving each study out indicated that removing Iqbal (2020) would increase the overall average effect ($\hat{\mu} = 0.06, (95\% \text{ CI: } 0.01 \text{ to } 0.11)$), and the effect would become statistically significant ($z = 2.47, p = .01$).

For decision making outcomes, effects from conditional cash transfers were no different than effects from unconditional cash transfers ($\hat{\beta} = 0.05, p = 0.21 [95\% \text{ CI: } -0.03 \text{ to } -0.13]$). There was also no moderation by exposure to intervention ($\hat{\beta} = -0.001, p = 0.14 [95\% \text{ CI: } -0.002 \text{ to } 0.0003]$). Effects from experimental studies were no different than effects from quasi-experimental studies ($\hat{\beta} = -0.05, p = 0.16 [95\% \text{ CI: } -0.13 \text{ to } 0.02]$). Publication year was also not a significant source of variation among effects ($\hat{\beta} = -0.01, p = 0.19 [95\% \text{ CI: } -0.02 \text{ to } 0.004]$). Effects did not differ by whether the study was examining post intervention changes versus changes from baseline ($\hat{\beta} = -0.05, p = 0.16 [95\% \text{ CI: } -0.13 \text{ to } 0.02]$) nor by whether or not the model was adjusted for covariates ($\hat{\beta} = 0.07, p = 0.052 [95\% \text{ CI: } -0.001 \text{ to } 0.14]$). Finally, we found no moderation by scores on neither the gender inequality index ($\hat{\beta} = 1.26, p = 0.06 [95\% \text{ CI: } -0.08 \text{ to } 2.60]$), nor the fragile states index ($\hat{\beta} = -0.01, p = 0.08 [95\% \text{ CI: } -0.01 \text{ to } 0.001]$). We were not able to test for moderation by study design (only one study used a QED), and there were no high risk of bias studies in this outcome group, so we did not test moderation by high risk of bias. We also could not test by evaluation period because all studies evaluated effects immediately after the end of the intervention.

Effects of cash transfers on more women engaged in other micro, small, and medium-sized enterprises

Only two studies reported the impact of cash transfers on more women engaged in other micro, small, and medium-sized enterprises, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as low risk of bias and the other as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.09 (95\% \text{ CI: } -0.04 \text{ to } 0.22)$. This effect was not significantly from zero ($z = 1.38, p = 0.17$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 48 CTCB6. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 2.78, p = 0.09, \hat{\tau}^2 = 0.01, I^2 = 64.10\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid.

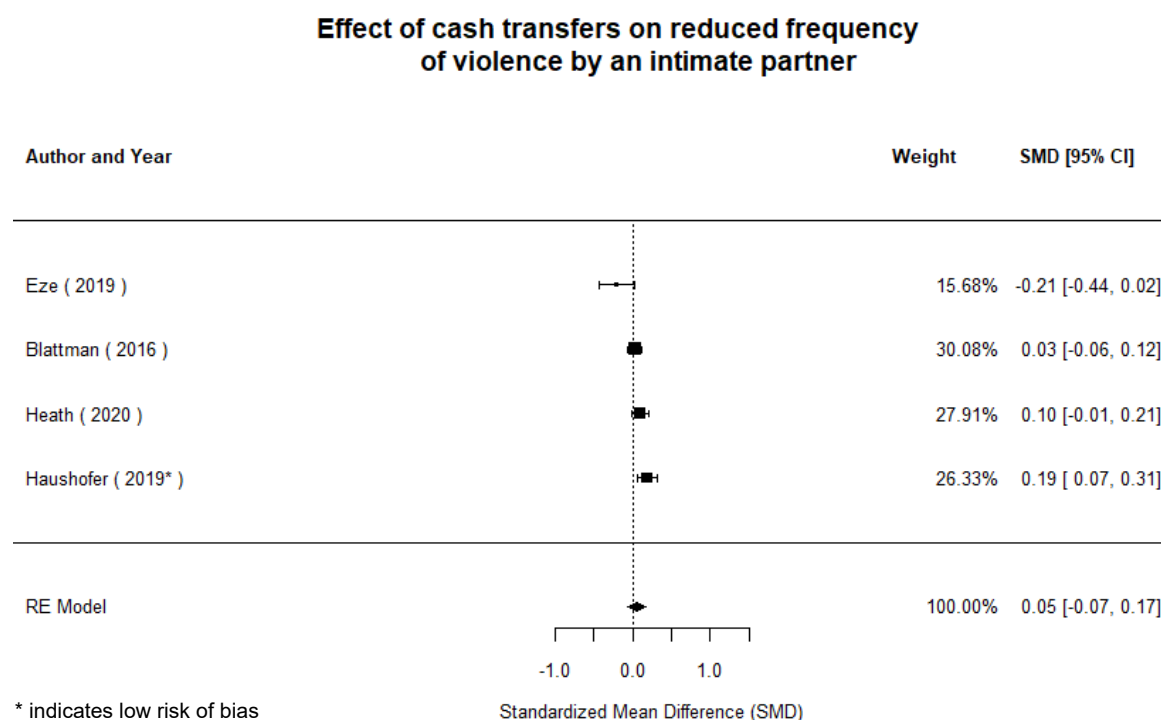
Figure 48: CTCB6: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of cash transfers on frequency and distribution of types of violence by an intimate partner

We included a total of $k = 4$ studies in the analysis. We assessed one study as low risk of bias, three as some concerns, and none as high risk of bias. The observed outcomes ranged from -0.21 to 0.19 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.05$ (95% CI: -0.07 to 0.17). Therefore, the average outcome did not differ significantly from zero ($z = 0.88$, $p = 0.38$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 49 CTCB7.

Figure 49: CTCB7: Forest plot showing the observed outcomes and the estimate of the random-effects model



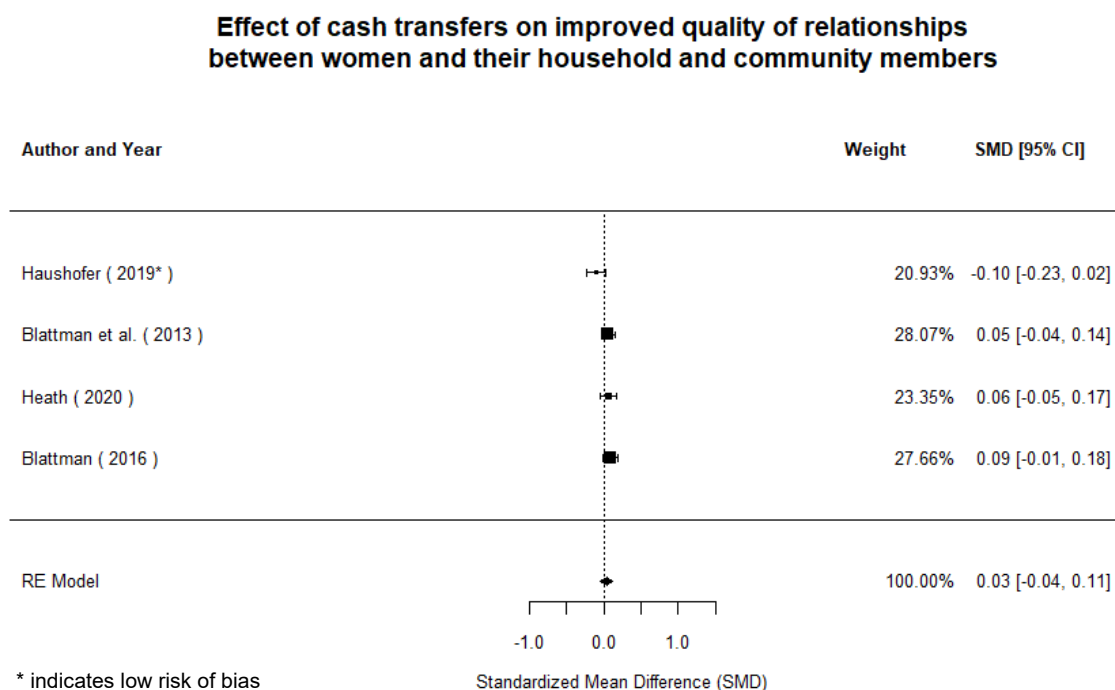
According to the Q -test, the true outcomes appear to be heterogeneous ($Q(3) = 10.52$, $p = 0.01$, $\hat{\tau}^2 = 0.01$, $I^2 = 71.49\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.58 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

There was also no moderation by exposure to intervention ($\hat{B} = 0.003$, $p = 0.79$ [95% CI: -0.02 to 0.03]), and publication year was also not a significant source of variation among effects ($\hat{B} = 0.01$, $p = 0.86$ [95% CI: -0.11 to 0.13]). Effects did not differ by whether or not the model was adjusted for covariates ($\hat{B} = 0.05$, $p = 0.79$ [95% CI: -0.32 to 0.42]). We were not able to test for moderation by study design (only one study used a QED), and there were no high risk of bias studies in this outcome group, so we did not test moderation by high risk of bias. We also could not test by evaluation period because only one study did not evaluate effects immediately after the end of the intervention, nor for moderation by whether the study was examining post-intervention changes or changes from baseline. Finally, we could not examine differences between conditional and unconditional transfers because only one used a conditional transfer.

Effects of cash transfers on improved quality of relationships between women and their household and community members

We included a total of $k = 4$ studies in the analysis. We assessed one study as low risk of bias, two as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.10 to 0.09 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.03$ (95% CI: -0.04 to 0.11). Therefore, the average outcome did not differ significantly from zero ($z = 0.81$, $p = 0.42$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 50 CTCB8.

Figure 50: CTCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model



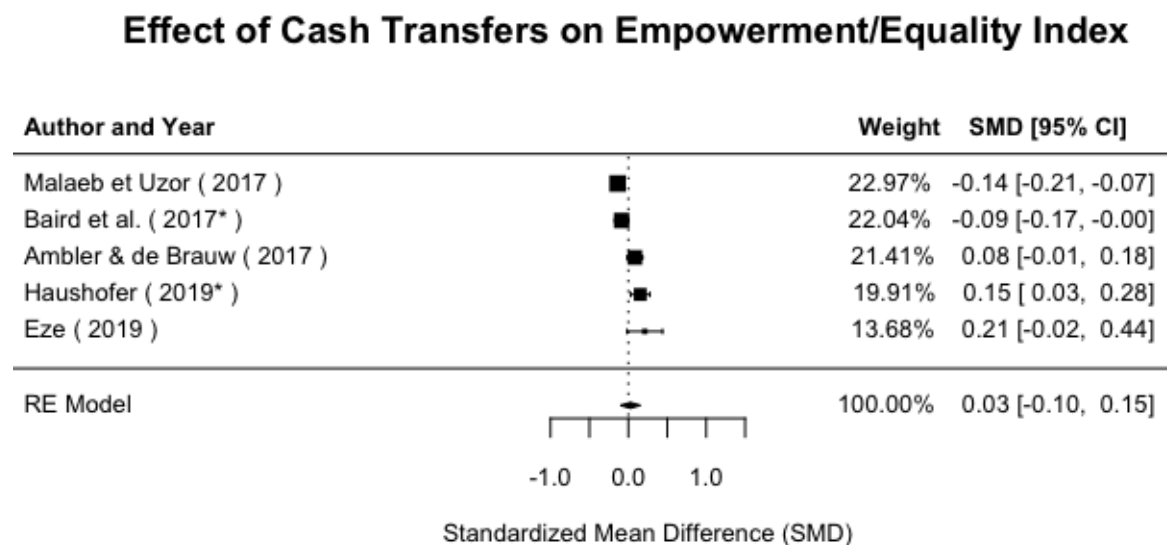
The Q -test for heterogeneity was not significant, but some heterogeneity may still be present in the true outcomes ($Q(3) = 6.45$, $p = 0.09$, $\hat{\tau}^2 = 0.00$, $I^2 = 53.50\%$). An examination of the studentised residuals revealed that one study (2) had a value larger than ± 2.58 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

Most moderators could not be tested because of a lack of variation among the studies, but of the two that could be tested, effects did not differ by publication year ($\hat{\beta} = -0.001$, $p = 0.89$ [95% CI: -0.02 to 0.01]), nor by length of exposure to intervention ($\hat{\beta} = -0.01$, $p = 0.53$ [95% CI: -0.04 to 0.02]).

Effects of cash transfers on women's empowerment index

We included a total of $k = 5$ studies in the analysis. We assessed two of the studies as low risk of bias, two as some concerns and one as high risk of bias. The observed outcomes ranged from -0.14 to 0.21 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.03$ (95% CI: -0.10 to 0.15). Therefore, the average outcome did not differ significantly from zero ($z = 0.41$, $p = 0.68$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 51 CTCC1.

Figure 51: CTCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(4) = 29.32$, $p < 0.01$, $\hat{\tau}^2 = 0.02$, $I^2 = 86.36\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.58 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

Effects from experimental studies were not different than effects from quasi-experimental studies ($\hat{\beta} = 0.16$, $p = 0.21$ [95% CI: -0.09 to 0.42]). They were also no different between studies using conditional versus unconditional cash transfers ($\hat{\beta} = 0.004$, $p = 0.98$ [95% CI: -0.30 to 0.31]). However, effects did differ by length of exposure to

intervention ($\hat{B} = -0.07, p = 0.045$ [95% CI: -0.01 to -0.001]) such that each additional month decreased the effect by .07 standard deviation units. Finally, effect from studies whose models were adjusted for covariates were larger than those from studies using unadjusted models by .23 standard deviation units ($\hat{B} = -0.07, p = -0.23$ [95% CI: -0.43 to -0.02]).

Qualitative findings

We conducted a thematic synthesis on 25 linked qualitative studies to the included cash transfer interventions. As indicated above, this thematic synthesis aims to identify themes related to the interplay of intervention design, intervention implementation, target population, and contextual variables with intervention outcomes and effects. In total, we identified 18 descriptive themes, which we configured into five analytical themes (Online Appendix N). These five analytical themes present the synthesis results and are discussed in more detail below.

Theme 1: Gender norms, culture and religious beliefs can preserve men's dominance in the household which affects programme uptake and participation and may result in undesirable outcomes in poor households.

Cash transfer programmes can support joint decision-making processes within beneficiary households and enhance the autonomy of women when making financial decisions. However, religion and traditional beliefs regarding gender roles were often used to justify men's superior status which leads to the perpetuation of control and inequality within poorer households. Cash transfers to married women may lead to men's dissatisfaction with women earning money and could cause household conflicts. In some communities, women are not allowed to overshadow their husbands as the provider (El Didi et al. 2018). The identified themes suggest that these cultural nuances of household power dynamics can undermine positive changes for women's empowerment outcomes.

Gender norms and cultural and religious beliefs thus featured heavily in cash transfer programmes and their uptake. For instance, qualitative findings from the WINGS programme in Uganda suggested that the role of supportive partners enhanced the well-being of target women and other benefits from the programme (Blattman, Fiala and Martinez 2013). Notwithstanding household heterogeneity, WINGS' women beneficiaries reported modest improvements in decision-making power within the household, even in the face of rigid traditional beliefs (Bonilla et al. 2017). Furthermore, an improvement in autonomy may mean that women felt more confident to make decisions without their partner's approval.

Household relationships are a more complex, yet influential, factor in how cash transfers administered were perceived and utilised. Gendered norms were seen to influence the control and use of cash within the households, often in favour of men. The participants' views were mixed, as secondary perspectives suggested that disagreements on how to use the money were often divisive and conflict-inducing in the Zambian child grant scheme (Bonilla et al. 2017). In Egypt, one Takaful and Karama beneficiary expressed improved relationship quality with her partner: *"We are able to afford a shared meal of meat more often. We became closer to each other,"* (El Didi et al. 2018). In Ethiopia, women participating in the Productive Safety Net Programme (PSNP) felt more empowered but were similarly restricted by existing household duties and mandatory public works participation (Gelagay and Lecoutere 2019).

Smaller cash benefits and those designated for women may only exacerbate tensions within beneficiary households. In the Takaful and Karama programme, a woman beneficiary, when asked if any joint decisions were taken between her and her husband, remarked: *“None. The money is with my husband to barely cover our food.”* The woman felt no improvement in her decision-making contribution within the household. Another noted that relations between her and her partner had deteriorated slightly adding that: *“Yes, [the transfers] can make some problems because a man can get mad... Some men can get annoyed because the money is coming to the woman, so he would think she’ll spend it on herself.”* Moreover, men’s dominance in traditional societies seemed to stifle programme impact when it came to household relationships and empowerment.

Theme 2: Targeting presented a key tool to generate benefits for the most vulnerable households and individuals

Cash transfer programmes typically targeted the most vulnerable households and women who struggle to meet their basic needs and those who do not have access to credit or capital to start businesses. Targeting procedures aimed to minimise beneficiary selection errors, maximise trust in administrative processes and focus on low-income groups. Accurate targeting requires consideration of social and economic contexts so that it works to empower the historically marginalised and improve their well-being. Women felt more empowered when targeted as beneficiaries than if men were. As one programme officer of the Takaful and Karama programme observed: *“We cannot conclude whether these effects would have been different if the transfers were targeted to men, but it seems unlikely that doing so would have improved women’s control over household decisions”* (El Didi 2018). There was additional support by both men and women of the choice to prioritise women (El-Enbaby et al 2019).

An inclusive targeting approach is important to help policymakers reach those in socioeconomic need and families as a unit rather than individuals (El Didi et al. 2018; Berlie 2018). Participants in the Takaful and Karama Programme reflected that for severely poor households, the cash injection was largely spent on basic needs such as food, clothing and medical expenses (El Didi et al. 2018). Most of these beneficiaries had high levels of financial debt from an array of sources. Rural Egyptian communities with rigid social norms were difficult to access and to target women, who were blocked from participating in work (El-Enbaby et al. 2019).

Beneficiaries of the WINGS programme in Uganda suggested that targeted participants should be of a determined eligible age to enhance the way cash is used in the household. Indeed, some perceptions suggested that younger cohorts of the beneficiary pool were undeserving of the cash because they had an income to rely on (Baird 2013). There was evidence on meeting of conditional requirements to qualify for cash injections; participants in the Takaful and Karama programme suggested that it was a justified strategy to ensure that the needy are prioritised and that cash is not misused (Zaki 2018). Poorer households were also often indifferent on meeting the requirement of minimum school attendance and clinic visit targets for pre-schoolers which was stipulated by the Takaful and Karama programme (El-Enbaby et al. 2019).

Qualitative results from the PSNP urged that the graduation model for beneficiaries was perceived to prematurely drop qualifying beneficiaries (Gelagay 2018). Furthermore, strict criteria on graduation requirements might be a problem. As one beneficiary adds: *“At*

present, I graduated from safety nets programme but I could not feed my family for 12 months from my own production,” (Berlie 2018). Some of the key issues alluded to the absence of a full family targeting strategy as well as low asset formation, as cash was largely used for basic needs. Due to these two factors, the graduation eligibility criteria in some instances excluded qualifying beneficiaries prematurely (Berlie 2018).

Theme 3: Mentorship and support from either programme mentors or family members encourages beneficiaries to engage in programme activities.

Mentorship, training and support from family members encourages adherence and compliance to programme requirements, especially in conditional cash transfer programmes. Mentorships encourage beneficiaries to meet their financial obligations and for parents to enrol their children in schools, for instance. Family coaching is perceived as useful for the achievement of financial autonomy and improved marital relations for women. Stringent requirements for beneficiary compliance can stifle programme implementation. To maintain eligibility status, cash transfer programmes targeted at younger women often require parents to adhere to conditions. The feasibility of these conditions requires careful consideration of the extent to which they may curtail positive outcomes.

From the qualitative evidence, an integral part of cash transfer programmes was the role of interactions between participants, programme staff and third-party support structures. The Rural Entrepreneur Access Project (REAP) in Kenya revealed some crucial insights. For both monitoring and evaluation (MandE) and programme support purposes, BOMA Project mentors supported REAP participants in ensuring programme adherence and coaching for two years to enhance school attendance, financial literacy and beneficiaries' knowledge (BOMA 2018).

Additional results from Karimli (2020) indicate that the mentoring and coaching component embedded within the Trickle Up cash transfer programme in Burkina Faso made various positive contributions, largely related to resource access and control. Evidence from the Takaful and Karama scheme suggested that women preferred being the main beneficiaries of cash programmes with one female beneficiary expressing: *“It’s good this way (that the transfers are going to women) [...] the state’s caring for me. In other words, it’s given me dignity,”* (El-Enbaby et al. 2019).

Theme 4: Awareness-raising initiatives support programme coverage and sensitise target beneficiaries to programme objectives to encourage participation.

Cash transfer programmes design factors included a sufficient range of awareness-raising strategies and prior support to understand the objectives and properties of the programme. Sensitising target populations is important to consider as a design factor that can support the realisation of desired impacts. This is convenient in trying to influence perceptions of prospective beneficiaries, so that they fully understand issues of eligibility, conditionality and full compliance. Prior knowledge on the exact size of cash transfers is a relevant factor for programme design. When cash transfers and programme coverage were perceived to be insufficient, beneficiaries were less satisfied. The role of inflation in diminishing purchasing power and limiting beneficiary options for cash use presents a further design consideration.

In the Zomba Cash Transfer Programme (ZCTP) in Malawi, slow uptake and lack of awareness in communities of interest motivated the implementation team to design and execute a formal sensitisation process. Participation rates improved and were reported by programme staff as successful in helping prospective participants understand and accept the intervention. However, the additional time and resources spent on this was costly to the overall programme (Baird 2013). For the ZCTP, the investigation team utilised the village structure to ensure that willing participants were successfully identified. Another key barrier to implementation was that potential participants were seemingly only aware of the study but not the specification that a cash incentive might be offered to chosen beneficiaries.

In the case of Trickle Up in Burkina Faso, Karimli, Bose and Kagotho (2020) found that because prior sensitisation was not feasible due to lack of resources, the implementation team resolved to limit their intended programme reach. The evidence suggested that there might be value in raising awareness and educating target populations prior to full implementation. Takaful and Karama continued to struggle with raising awareness as a key member of the government implementation team expressed: *"No booklets or brochures have been given to the applicants,"* apart from printed posters, which included the eligibility criteria for prospective beneficiaries (Zaki 2018).

Theme 5: Adverse programme operations, untimely delivery of cash and negative perceptions of administrative challenges are bottlenecks to the successful implementation of cash transfer programmes.

The timely and predictable delivery of cash and the low cost of access to beneficiaries are key design contributors to enhance cash transfer programme success. Obstacles to access the transfer and perceptions of administrative difficulties may temper programme satisfaction. Adherence and participation rely heavily on beneficiary knowledge on and confidence in cash transfer programmes.

Zambia's cash transfer scheme was logistically satisfactory for beneficiaries. However, most study participants suggested a more intimate, door-to-door delivery model for payments (Seidenfeld, Handa and Tembo 2013). In addition to this, beneficiaries felt that the cash amount received was insufficient to adequately cover their basic household needs. A Takaful and Karama beneficiary expressed that: *"What the government gives with the right hand, it takes with the left one,"* while another added that: *"This isn't the sum of money that would let the man of the household sleep,"* (Zaki 2018). Moreover, nearly all surveyed recipients of the cash transfer were satisfied with the timely availability of their benefits and had confidence that their next payments would also be available on time (Handa et al. 2019). Another important factor from the qualitative evidence is related to how beneficiaries perceived the administration and dissemination of cash benefits. Negative perceptions on programme staff were an additional barrier to implementation.

Additional qualitative findings from the Takaful and Karama cash transfer programme revealed that many beneficiaries were unhappy with the selection process; those who did not qualify were reported to often have been erroneously included. Eligibility criteria and ineffective adherence monitoring were reportedly often manipulated by beneficiaries who were able to underreport their earnings without consequence (El Didi et al. 2018). Furthermore, perceived corruption in the provision of cash benefits and staff mistreatment of beneficiaries were cited as common issues in the Egyptian context. One Takaful and

Karama beneficiary complained that: *“They were pushing us and dealing with the people in a very bad way”* while another shared a different sentiment on staff treatment: *“...they were good to us and dealt with us in a good way”*. Ultimately, hostility from government staff was met with hostility from the beneficiaries (Zaki 2018). In some cases, cash transfer programmes may not be adequately marketed to intended beneficiaries.

Lack of knowledge on programme operations and systems presents an impediment to uptake and adherence because participants often have little understanding on how to maintain their benefits and involvement in the cash transfer programmes. Beneficiaries in the Takaful and Karama programme expressed scant knowledge on the programme rules and actual value of the benefits while implementing partners confessed that there was no reliable system in place to monitor such failures or limitations (Zaki 2018). Other recipients of the transfer knew the programme by a different name which curbed their interest in learning the rules and qualification criteria (*“Maash El-Sisi”* was the name that many participants used). Some of the monitoring failures were captured when one Takaful and Karama informant complained that: *“Someone called me and asked me a number of questions and then told me that if anyone asked you if a researcher visited you say yes.”* To ensure that beneficiaries were compliant, another beneficiary was of the view that: *“The government has the right to enter the homes of the applicants,”* to assess applicant living conditions and confirm eligibility as a monitoring technique (Zaki 2018).

Discussion

Across 20 total studies ($n = 20$), quantitative meta-analysis findings from cash transfers tested a wide range of outcomes covering domains such as empowerment, well-being and relationship quality. While several outcomes such as those related to access to assets, credit, and income as well as engagement in small, medium and micro enterprises are positive, there are more qualitative insights to discuss the empowerment-related outcomes. Summarily, there are smaller positive quantitative effects on empowerment-related outcomes such women’s decision-making power in communities or households and gender equality (albeit statistically significant) outcomes. A theme that emerges quite commonly from the qualitative evidence base ($n = 25$) is the centrality of gender norms and roles when designing programmes. Below are a few possible explanations from the qualitative evidence:

- Social norms, existing gender roles and rigid traditional beliefs can preserve men’s dominance regardless of the cash injection in a typical household. While the cash transfer can accrue to women and/or girls within the household, men can control the use of these benefits or even discourage the uptake of these programmes by women and girls within the household.
- The design of cash transfer programmes that consider the harmful role of structural norms in society can aid improved outcomes to empowerment. Cash benefits should not be assumed to directly spill over to addressing structural norms.
- Programmes can be enhanced by adding components such as a couple life, social and livelihood skills and capacity building component that aims to improve household relationships and conflict resolution skills. Additionally, mentoring components can deepen peer-to-peer exchanges and collective building to cast the net for learning and interaction wider.
- Cash transfers can advance the financial standing of women in society as asset ownership, access to credit and income improve over time if benefits are used well.

- Women's improved access to livelihood support services can be attributed to their improved well-being and financial position as they can access key amenities and services that are central to their livelihood.
- There are various points of focus that present opportunities for programme designers to refine cash transfer programmes:
 - More accurate targeting efforts and clear awareness-raising tactics can boost uptake and the identification of the most relevant target group.
 - Instead of targeting beneficiaries individually, there can be better consideration for family unit targeting to mitigate conflict about benefits.
 - According to the qualitative data, there is evidence that beneficiaries would prefer more cash as a benefit due to inflationary pressures and household size.
 - Programme staff emphasised that this cash amount should not replace salaries and wages from work.

Summary of findings and discussion

We included 20 studies in 12 countries in MENA, South Asia and Sub-Saharan Africa that evaluated the effect of cash transfers. We were able to examine effects on the following outcomes: Women have increased access and ownership to assets, credit and income, Initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities, women have increased freedom of movement and association, Increased participation in decision making by women at the household or community level, including during crisis response, women participate more in their community. Our included studies report against all three of the secondary outcomes (Resources, Agency and Achievement) and eight of the nine immediate outcomes for our review. Overall, the GRADE assessments generally indicate a very low to high certainty in this body of evidence.

For 25 linked qualitative studies, five analytical themes emerged that buttress the quantitative findings and give a contextual lens to the outcomes of cash transfer programmes in FCAS. Gender norms were a significant barrier to success for women or individuals receiving cash transfers due to roles and permissions related to the management of money in the household. Targeting specific vulnerable populations supported progress against specific outcomes. However, project management and implementation could hinder the ability of cash transfers to positively impact a target population. Lastly, awareness raising and mentorship initiatives complement cash transfers by sensitising beneficiaries and their household and community members on the timeline and use of cash transfers.

Table 21 below presents the GRADE review of our findings:

Table 21: GRADE summary of findings and certainty of evidence on cash transfers

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(AB2) Women's access to assets, income and savings (Assessed with: self-reports)										
12	RCT- 11 QED - 1	serious ^a	serious ^b	not serious	not serious	strong association	22126	SMD 0.22 SD higher (0.12 higher to 0.31 higher)	⊕⊕⊕○ MODERATE	Important, but not critical
(AB3) Women and girls have equitable access to livelihood support services.										
2	RCT - 2	not serious	not serious	not serious	not serious	none	3869	SMD 0.16 SD higher (0.1 higher to 0.22 higher)	⊕⊕⊕⊕ HIGH	Important, but not critical
(AC1) More women engaged in other micro, small, and medium sized enterprises.										
2 (Linked)	RCT	not serious	serious ^h	not serious	not serious	none	3288	Two linked studies reporting effect estimates of 0.81 and 0.91.	⊕⊕⊕○ MODERATE	Important, but not critical
(AC2) Initiatives supported to facilitate women's access to decent work										
6	RCT - 6	serious ^d	serious ^e	not serious	serious ^f	none	8651	SMD 0.18 SD higher (0.01 lower to 0.36 higher)	⊕○○○ VERY LOW	Important, but not critical
(AC3) Improved capacity of women entrepreneurs										
1	RCT	very serious ^g	serious ^h	not serious	not serious	strong association	1997	SMD 0.4 SD higher (0.32 higher to 0.49 higher)	⊕⊕○○ LOW	Important, but not critical
(BA2) Women have more and better control over their bodies and sexual health.										
2	RCT	serious ⁱ	serious	not serious	not serious	none	6888	SMD 0.07 SD lower (-0.17 lower to 0.04 higher)	⊕⊕○○ LOW	Critical
(BA3) Women have increased freedom of movement and association.										
3	RCT – 2 QED - 1	serious ^l	serious ^m	not serious	serious ⁿ	none	4021	SMD 0.01 SD higher (-0.10 lower to 0.13 higher)	⊕○○○ VERY LOW	Limited Importance
(BA6) Reduced percentage of women agreeing with certain reasons that justify violence against women and girls										
3	RCT – 1 QED - 2	serious ^o	not serious	not serious	not serious	none	3098	SMD 0.11 SD higher (0.02 higher to 0.2 higher)	⊕⊕⊕○ MODERATE	Critical
(BA7) Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them.										
3	RCT - 3	not serious	not serious	not serious	serious ^p	none	5054	SMD 0.03 SD higher (0.04 lower to 0.1 higher)	⊕⊕⊕○ MODERATE	Important, but not critical
(BA8) Reduced instances of child or forced marriage										
2	RCT	not serious	not serious	not serious	serious ^q	none	2109	SMD 0.07 SD higher (0.1 lower to 0.25 higher)	⊕⊕⊕○ MODERATE	Critical
(BB1) Women's decision-making at the household and community levels (Assessed with: self-reports)										
6	RCT – 5 QED - 1	Very serious ^r	very serious ^s	not serious	serious	none	31619	SMD 0.03 SD higher (0.03 lower to 0.9 higher)	⊕○○○ VERY LOW	Limited Importance

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(BB2) Women participate more in their communities.										
3	RCT – 2 QED - 1	very serious ^u	very serious ^v	not serious	serious ^w	none	6571	SMD 0.1 SD higher (0.07 lower to 0.28 higher)	⊕○○○ VERY LOW	Limited Importance
(BC1) The rights, safety, and security of women, men, girls, and boys are protected during relief, recovery, and reconstruction										
1	RCT	serious ^x	serious ^h	not serious	not serious	none	4112	Two positive effect estimates with a 95% CI range of 0.04 to 0.61.	⊕⊕○○ LOW	Critical
(CA2) Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration										
3 (Linked)	QED - 3	serious ^y	very serious ^z	not serious	not serious	undetected	12557	Three positive effect estimates with a 95% CI range of 0.02 to 0.3	⊕○○○ VERY LOW	Limited Importance
(CB2) Communities have a more positive attitude towards women/marginalised groups.										
1	RCT	very serious ^{aa}	serious ^h	not serious ^{ab}	serious ^{ac}	none	1800	One positive and one negative effect estimate with a 95% CI range of -0.20 to 0.11.	⊕○○○ VERY LOW	Limited Importance
(CB3) Women have improved attitudes, self-image and confidence.										
3	RCT – 2 QED - 1	not serious	serious ^{ad}	not serious	serious ^{ae}	none	4863	SMD 0.09 SD higher (0.01 lower to 0.18 higher)	⊕⊕○○ LOW	Limited Importance
(CB4) Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members										
5	RCT – 3 QED - 2	serious ^{af}	serious ^{ag}	not serious	not serious	none	11360	SMD 0.04 SD higher (0 to 0.07 higher)	⊕⊕○○ LOW	Limited Importance
(CB6) Safer and more secure household, communities and areas/territories for women, girls, men, and boys										
2	RCT - 2	serious ^{ah}	serious ^{ai}	not serious	serious ^{aj}	none	2813	SMD 0.09 SD higher (0.04 lower to 0.22 higher)	⊕○○○ VERY LOW	Critical
(CB7) Reduced frequency and distribution of types of violence by an intimate partner										
4	RCT – 3 QED - 1	serious ^{ak}	very serious ^{al}	not serious	not serious	none	4368	SMD 0.05 SD higher (0.05 lower to 0.17 higher)	⊕○○○ VERY LOW	Critical
(CB8) Improved quality of relationships between women and their household and community members										
4	RCT – 5	serious ^{am}	very serious ^{an}	not serious	serious ^{ao}	none	7405	SMD 0.03 SD higher (0.04 lower to 0.11 higher)	⊕○○○ VERY LOW	Limited Importance
(CC1) Empowerment/Equality index										
5	RCT – 3 QED - 2	very serious ^{ap}	very serious ^{aq}	not serious	serious ^{ar}	publication bias strongly suspected	8315	SMD 0.03 SD higher (0.1 lower to 0.15 higher)	⊕○○○ VERY LOW	Important, but not critical

Note. CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

- a. This group has been downgraded because three of twelve studies present high risk of bias. That said, when the highest risk of bias study is removed, the effect size is only very marginally impacted.
- b. One downgrade has been applied here because there is inconsistency inasmuch as there are estimate points whose 95% CI do not overlap at all. That said, this is not "very serious" because there is overlap across eight of twelve studies, and those which do not are still presenting positive effect sizes (i.e., the results vary, but are all positive).
- c. Funnel plot asymmetry detected.
- d. Downgraded once because, while four of six studies are of low or some concerns related to risk of bias, two present high risk. There is particular risk and uncertainty surrounding outcome measurement bias.
- e. Downgraded once to capture the wide range of point estimates, including across both sides of the threshold, and the CIs of several studies that do not overlap.
- f. Downgraded once because of wide CIs that cross both sides of the threshold.
- g. Downgraded to very serious because this study presents significant risk of outcome measurement bias as well as a lack of clarity across other criteria.
- h. All single study groups downgraded once for inconsistency.
- i. This analysis has been downgraded once due to one study having a high risk of bias and the other presenting some concerns related to deviation from intended intervention although they are all RCTs.
- j. Downgraded once: 1. Variation of points estimates: data on both side of the threshold, 2. Overlaps of CI: systematic overlap, 3. $I^2 = 0\%$, 4. test of stat: According to the Q-test, there was no significant amount of heterogeneity in the true outcomes
- l. This analysis has been downgraded once due to a mix of RCT and QED with only one study with low risk and the rest with some concerns related to risk of bias
- m. Downgraded to reflect CIs widely on both sides of the threshold and a lack of systematic overlap of CIs across different effect sizes.
- n. Downgraded once as p value of 0.14 and Sample size > 400.
- o. This analysis has been downgraded once due to a mix of RCT and QED with RCT with low risk and the rest with some concerns related to risk of bias.
- p. Downgraded once as p value of 0.46 and Sample size > 400.
- q. Downgraded because the CI's of the effect estimate cross the threshold considerably.
- r. This group has been downgraded on risk of bias because three studies present high risk of bias across a number of domains (reporting bias, selection bias). That said, when high risk of bias studies were removed from analysis, the finding only marginally changes and remains positive.
- s. Downgraded twice because 1) There are effect sizes whose confidence intervals do not systematically overlap, and who present both positive and negative impacts, and 2) Q-score and I^2 indicate significant heterogeneity which cannot be explained by moderators.
- u. Downgraded twice 1. Mix of QED and RCT 2. Two studies are low risk and one RCT is high risk of bias 3. Removing either study would have an impact on the meta-analysis.
- v. Downgraded twice: 1. Variation of points estimates: data on both side of the threshold, 2. Overlaps of CI: no systematic overlap, 3. $I^2 = 92.17\%$, 4. test of stat: According to the Q-test, the true outcomes appear to be heterogeneous.
- w. Downgrade once because of a p-value of 0.26 and Sample size > 400.
- x. BC1 The rights, safety, and security of women, men, girls, and boys are protected during relief, recovery, and reconstruction.
- y. Downgraded once because the body of evidence is comprised three QED studies two of three of them raised some concerns related to risk of bias.
- z. Downgraded once to reflect the range of point estimates.
- aa. Downgraded because of risk of bias relating to the unit of analysis and a lack of clarity across other criteria.
- ab. Downgraded once because the point estimates of this study cross both sides of the threshold and have a wide 95% CI.
- ac. Downgraded for imprecision because the CI of the meta-analytic effect size crosses the threshold while also crossing a non-trivially beneficial level.
- ad. Downgraded once: 1. Variation of points estimates: data on same side of the threshold, 2. Overlaps of CI: systematic overlap, 3. $I^2 = 59.86\%$, 4. test of stat: According to the Q-test, some heterogeneity may still be present in the true outcomes.
- ae. Downgraded once as p value of 0.07 and Sample size > 400.
- af. Downgraded once because some studies presented unclear risk of bias, and less than half of the studies included are of low risk of bias.

- ag. Downgraded once because although there is overlap in the confidence intervals of effect sizes, some are positive while others are negative. Additionally, $I^2 = 6.43\%$.
- ah. Downgraded once because half of the studies are rated as having high risk of bias with particular concern for outcome measurement bias.
- ai. Downgraded once because of widely varying point estimates that do not fall with the CIs of the other study.
- aj. Downgraded once for wide CIs around the point estimate that cross both sides of the threshold.
- ak. This group has been downgraded once because several studies are "unclear" in a number of risks of bias categories (namely in performance bias and outcome measurement bias). That said, no studies in this group present high risk of bias.
- al. Two downgrades have been applied for inconsistency because 1) There are effect sizes whose confidence intervals do not overlap, and who present both positive and negative impacts, and 2) The $I^2 = 68.34\%$ suggests that some inconsistency can only be interpreted as due to chance, and not heterogeneity.
- am. This group has been downgraded once 1. All studies are RCTs but one of them has a high risk of bias, two with some concerns related to risk of bias and two low risk 2. The study with high risk of bias represents a high proportion of results of the meta-analysis
- an. Downgraded twice: 1. Variation of points estimates: data on both side of the threshold, 2. Overlaps of CI: no systematic overlap, 3. $I^2 = 56.28\%$, 4. test of stat: According to the Q-test, some heterogeneity may still be present in the true outcomes.
- ao. Downgraded once as p value of 0.15 and Sample size > 400.
- ap. This group has been downgraded twice 1. Mix of RCT and QED 2. One study is high risk, two are medium and two are low risk 3. Removing the high risk would affect the results of the meta-analysis.
- aq. Downgraded twice: 1. Variation of points estimates: data on both side of the threshold, 2. Overlaps of CI: no systematic overlap, 3. $I^2 = 86.36\%$, 4. test of stat: According to the Q-test, the true outcomes appear to be heterogeneous.
- ar. Downgraded once as p value of 0.68 and Sample size > 400

4.3.3 Community-based services

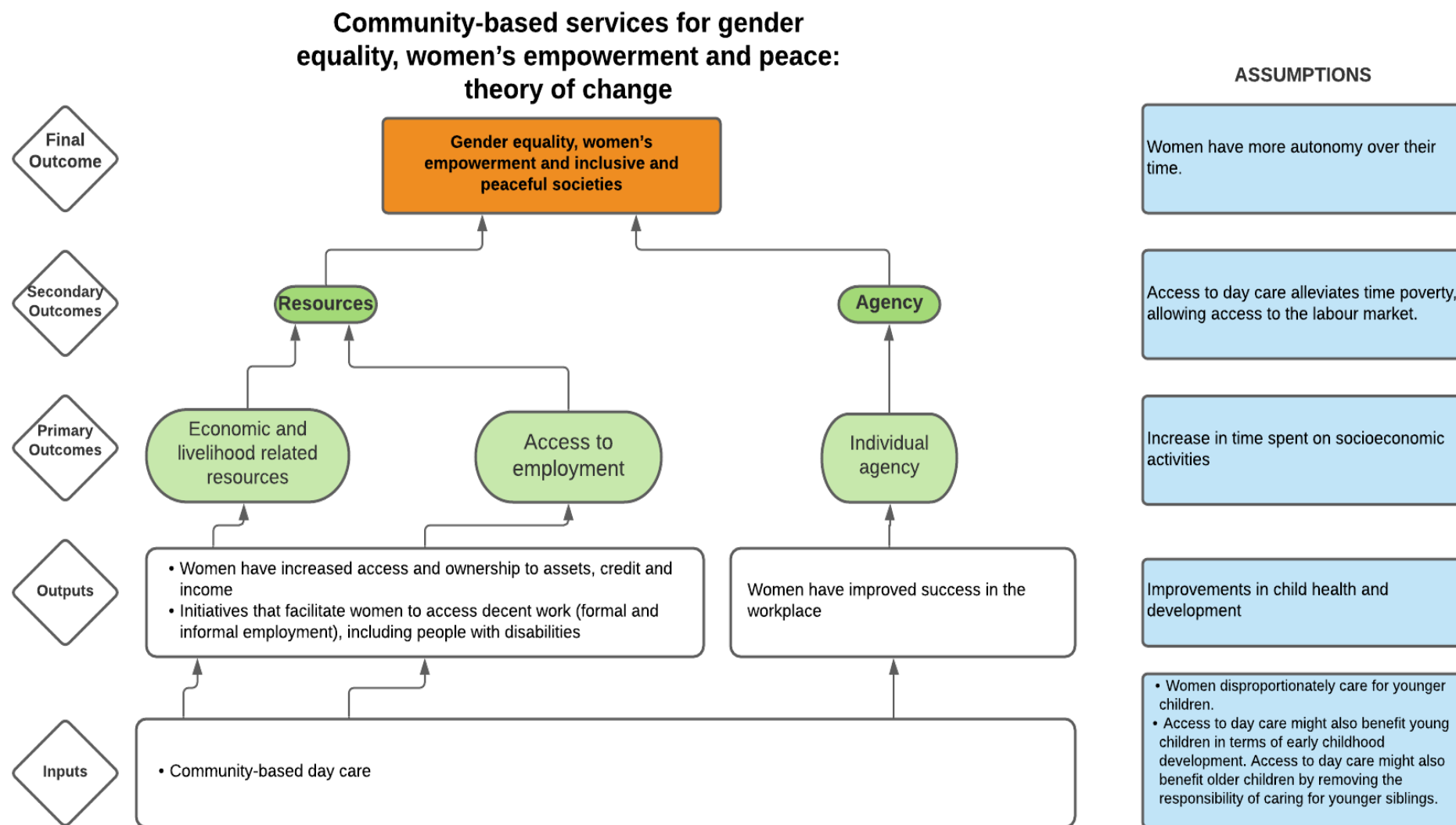
Community-based services are the improved public provision of services and support to families and individuals at the local and/or community level. This can include health, education, economic service provision and/or counselling and networking support.

How do community-based services affect gender equality, women's empowerment and peace outcomes?

Figure 52 below maps out the causal chain of how community-based services may improve gender equality, women's empowerment and peace outcomes.

As illustrated in figure 52 below, community-based service initiatives that are inclusive (such as community-based day care services to reduce the burden of childcare) can help societies to achieve high levels of societal resources and agency, which are essential to the overall empowerment of women and the fulfilment of an inclusive and peaceful society. These resources and individual agency rely on access to economic and livelihood related resources, access to employment, and individual agency, and are fostered by immediate outcomes that are generated from inclusive community-based service programmes. Immediate outcomes of such programmes include women having increased access and ownership to assets, credit and income, facilitating women to access decent work (formal and informal employment) and improving their success in the workplace.

Figure 52: Community-based services for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included only one study that evaluated the effect of community-based services (Table 22).

Population

The included study evaluated one programme in South Asia, in one country: India.

Intervention, inputs and activities

The included study evaluated the following community-based service activity:

- Community-based day care (n = 1): Programme implementation was structured around one major intervention at the village level, the provision of childcare, nutritious food and supplements, basic medicines and preschool education to children one to six years old.

Table 22: Community-based services features of included studies

Study	Activity/Input	Length of treatment	Intervention frequency
Nandi (2020)	Community-based day care services	24 months	Daily

Comparison

The included study compared treated groups to comparison groups receiving no intervention. No study included multiple treatment arms.

Outcomes

The included study reported on a number of relevant outcomes:

- Women have increased access and ownership to assets, credit and income (n = 1): Women can apply for, receive and manage assets/credit and income and have support to manage, claim and execute their assets without pressure or influence from external actors, including male family members, husbands and cultural leaders.
- Initiatives that facilitate women to access decent work (formal and informal employment), including people with disabilities (n = 1): Women are able to apply for, receive and work in jobs (and have support for the above), without discrimination for sex, gender or other identifying factors, including development of skills for improved access.

The division of the immediate and secondary outcomes is reported in Table 23 below:

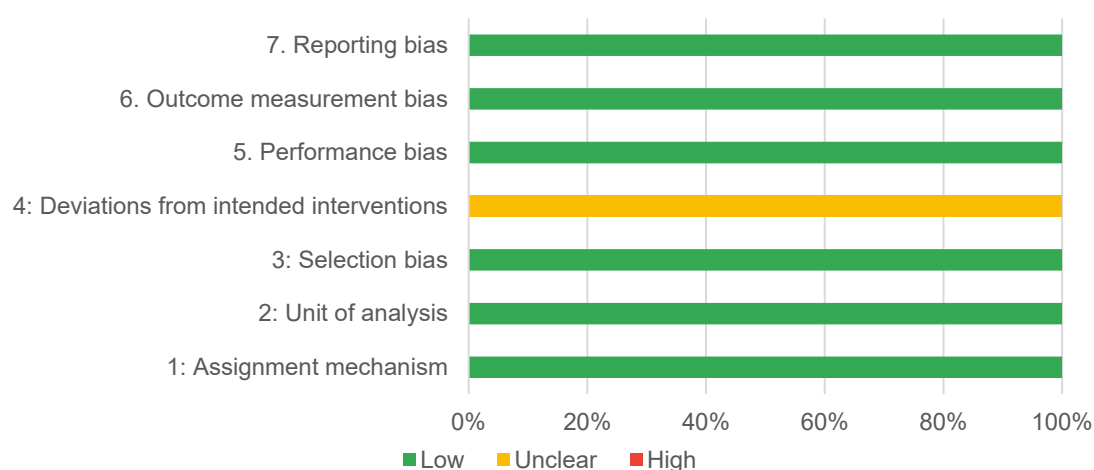
Table 23: Summary of secondary and immediate outcomes for community-based services

Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	0
	Economic and livelihood related resources	1
	Access to employment	1
Agency ability to define one's goals and act upon them and operationalised decision-making	Individual agency	1
	Community level agency	0
	Institutions supporting agency	0
Achievement ways of being and doing which can be realised by different individuals	Improved systems	0
	Norms and behaviour change	0
	Empowerment index	0

Study design

Our only community-based services study was Nandi (2020). It used an experimental design with no major concern but potential contamination issues where children from control villages attended day-care services in neighbouring treatment villages. We did not identify any other limitation related to any of the risk of bias domains (assignment mechanism, unit of analysis, selection, performance, outcome measurement and reporting), attrition was relatively low and similar between treatment and control groups, baseline characteristics were balanced for both groups, and allocation was concealed until after the baseline survey to minimise opportunities for bias in recruitment of participants. Figure 53 below presents the results of the RoB assessment:

Figure 53: RCT Community based services RoB assessment



Qualitative studies, process evaluations and project documents

We identified no linked qualitative studies and were unable to conduct a qualitative evidence synthesis related to the community-based services intervention.

Synthesis of findings

The following subsection presents the results of effectiveness of community-based services on gender equality, women's empowerment and peace outcomes.

Quantitative findings

Effects of community-based services on women accessing decent work (formal and informal employment)

Nandi and colleagues' (2020) experimental study in India also evaluated the impact of community-based services on women to access decent work (formal and informal employment). Their report included three effects that fell into this outcome category (e.g., worked in past seven days). The effects ranged from very small, negative estimates ($g = -0.02$, 95% CI $[-0.09, 0.05]$) to very small, positive estimates ($g = 0.02$, [95% CI: -0.05 to 0.09]), but none were statistically significant. We assessed the study as having some risk of bias concerns.

Qualitative findings

Summary of findings and discussion

We included only one study in India that evaluated the effect of community-based services. We examined effects on two outcomes: women have increased access and ownership to assets, credit and income and women's access decent work (formal and informal employment). Our included study reported against one of the three secondary outcomes (Resources, Agency and Achievement) and two of the nine immediate outcomes for our review. We found no evidence of the community day care service being effect for increasing outcomes related to women's empowerment and gender equity. Overall, the GRADE assessments (Table 24) generally indicate a low certainty in this body of evidence.

We identified no linked qualitative studies and were unable to conduct a qualitative evidence synthesis related to the community-based services intervention.

Table 24: GRADE summary of findings and certainty of evidence on community-based services

Certainty assessment							Sample size	Effect	Certainty	Importance
Ne of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(AB2) Women have increased access and ownership to assets, credit and income										
1	RCT	serious ^a	serious ^b	not serious	not serious	none	2859	Two (identical) positive effect estimates of 0.02 (-0.05, 0.09).	⊕⊕○○ LOW	Important, but not critical
(AC2) Initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities										
1	RCT	serious ^a	serious ^b	not serious	not serious	none	2859	Two negative and one positive effect estimate with a 95% CI range of -0.09 to 0.09.	⊕⊕○○ LOW	Important, but not critical

Note. CI: Confidence interval

Explanations:

a. Downgraded once due to issues of bias relating to deviation from intended intervention.

b. All single studies downgraded once for inconsistency.

4.3.4 Inclusive community-driven development

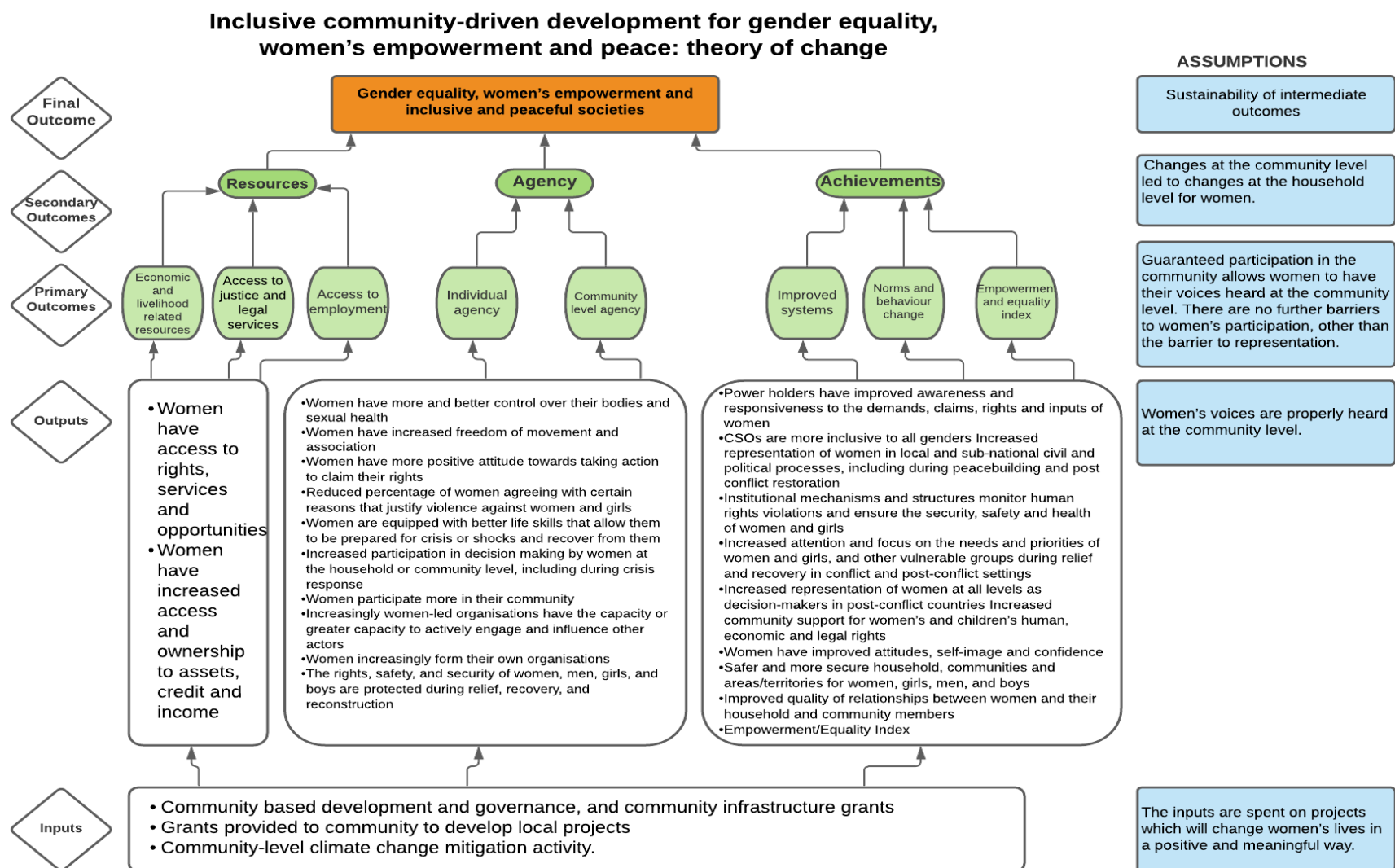
We define inclusive community-driven development as development initiatives that put control of the development process, resources and decision-making authority directly towards groups in a community. Community-driven development (CDD) programmes operate on principles of transparency, participation, local empowerment, demand-responsiveness, greater downward accountability and enhanced local capacity (World Bank 2021). This enhances participatory involvement at the local level (White et al. 2018). In our case, CDD can include a marginalised group in the decision-making process or as beneficiaries. It is a bottom-up model of development that aims to increase participation among the beneficiaries of development.

How does inclusive community-driven development affect gender equality, women's empowerment and peace outcomes?

Figure 54 below maps out the causal chain of how inclusive community-driven development may improve gender equality, women's empowerment and peace outcomes.

As illustrated in Figure 54 below, community-driven development initiatives that are inclusive (such as community-based governance as well as provision of resources to tackle community challenges) can help societies to achieve high-level resources, agency and achievements, which are essential to the overall empowerment of women and inclusive and peaceful society. These resources, agency and achievements rest on access to justice and legal services, livelihood related resources, individual and community agency, norms and behaviour change, among others, and are fostered by immediate outcomes that are generated from inclusive community driven development programmes. Immediate outcomes of such programmes include women's access to rights and services, improved participation in decision making, improved freedom, among many others.

Figure 54: Inclusive community-driven development for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included three programme studies reported in eight different papers that evaluated the effect of inclusive community-driven development programmes. We included more than one paper that evaluated the same programme if the author(s) reported different outcomes over several papers.

Population

The included studies focussed on individuals as well as entire communities, with special attention given in many cases to opinion leaders such as village chiefs. The included studies evaluated three different programmes in Latin America, East Asia and the Pacific, South Asia and Sub-Saharan Africa in eight countries: Brazil, Cameroon, Indonesia, Peru, Tanzania, Vietnam, Afghanistan and DRC.

Intervention, inputs and activities

The included studies evaluated a range of different inclusive community-driven development activities and inputs including the followings:

- Community-based development and governance and community infrastructure grants (n = 3): Programme implementation was structured around two major interventions at the village level: (i) the election of a gender-balanced community development council (CDC) through a secret-ballot, universal suffrage election centred on democratic processes and women's participation; and (ii) the provision of 'block grants' – valued at \$200 USD per household, up to a village maximum of \$60,000 USD and averaging \$33,000 USD – to fund village-level projects designed and selected by CDCs in consultation with villagers. Projects generally fall into one of six categories: transport; water and sanitation; irrigation; power; literacy and vocational training; and others.
- Community-level climate change mitigation activity (n = 1): REDD+ initiatives vary across countries but typically involve work to prevent deforestation by working with community-level NGOs.
- Grant provided to community to develop local projects (n = 4): Tuungane 1 had two components: (i) an intervention at the village development committee (VDC) level that involved \$3,000 USD grants, and subsequently; (ii) an intervention at the CDC level that involved \$50,000 to \$70,000 USD sub-grants to undertake infrastructure projects such as the construction of schools and hospitals. A core element of Tuungane 1 was a focus on women's empowerment and on championing the rights and roles of women in collective decision-making.

Table 25: Inclusive community-driven development activities design features of included studies

Study	Activity/input	Length of treatment	Intervention frequency
Beath (2008)	Community-based development and governance and community infrastructure grants	50 months	N/A
Beath (2015)	Community-based development and governance and community infrastructure grants	50 months	N/A
Beath (2013)	Community-based development and governance and community infrastructure grants	50 months	N/A
Larson (2018)	Community-level climate change mitigation activity	48 months	Annually
Laudati (2018)	Grant provided to community to develop local projects	48 months	N/A
Mvukiyehe (2020)	Grant provided to community to develop local projects	48 months	N/A
Van Der Windt (2018)	Grant provided to community to develop local projects	48 months	N/A
Van Der Windt (2018) (linked paper)	Grant provided to community to develop local projects	48 months	N/A

Comparison

All included studies compared treated groups to comparison groups receiving no intervention. One study included multiple treatment arms.

Outcomes

The included studies reported on a number of relevant outcomes, including the following:

- Women have increased freedom of movement and association (n = 5): Women and girls are able to move freely and meet within their household, community, and area/territory without fear of attack or discrimination.
- Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration (n = 6): Women are represented in the institutions, organisations and other decision-making processes and have capacities to influence the decision and direction of these processes. Women's specific needs and requests are taken into account and are reflected in the final decision, giving a full role to women in civil and political processes.

- Women have improved attitudes, self-image and confidence (n = 2): Women feel more entitled to claim their rights and needs in community and change social norms and behaviours. They are aware of the importance of their status in society and are empowered to take this role and make use of the available resources to guarantee their rights.
- Safer and more secure household, communities and areas/territories for women, girls, men, and boys (n = 4): We report the division of the immediate and secondary outcomes in Table 26 below:

Table 26: Inclusive community-driven development summary of secondary and immediate outcomes

Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	1
	Economic and livelihood related resources	2
	Access to employment	0
Agency ability to define one's goals and act upon them and operationalised decision-making;	Individual agency	6
	Community level agency	4
	Institutions supporting agency	0
Achievement ways of being and doing which can be realised by different individuals	Improved systems	6
	Norms and behaviour change	5
	Empowerment index	0

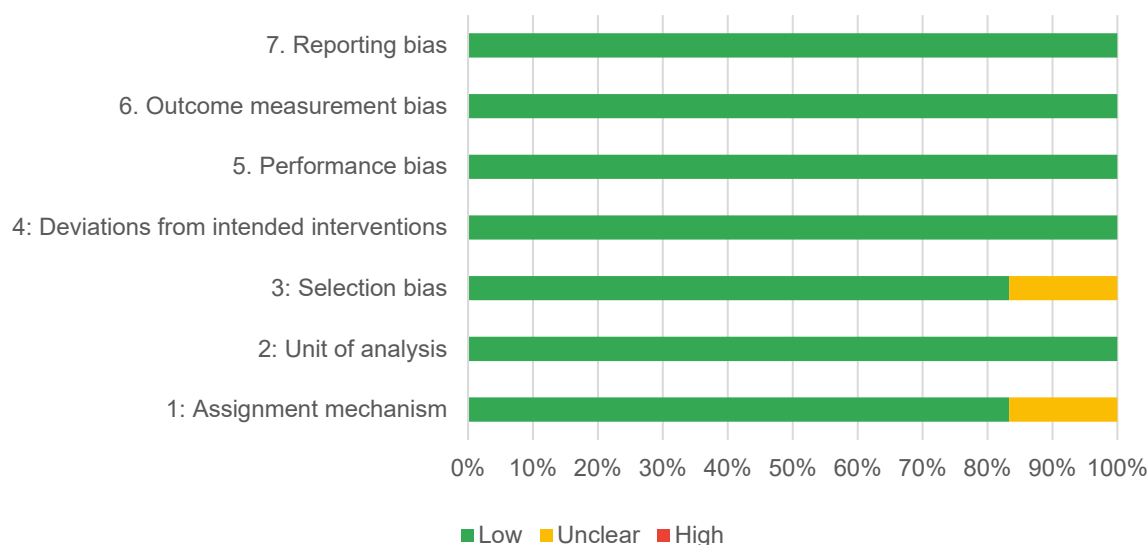
Study design

One of our inclusive community-driven development studies used a QED (Larson 2018). It was assessed as having high risk of bias because of reporting issues related to a lack of details provided on the matching methods used for identifying a valid comparison group. The extent to which matching methods were able to control for a comprehensive set of time-varying characteristics that might confound the impact of the program was unclear, tests for hidden bias were not reported, nor were sensitivity analyses presented, and balance after matching was not tested. Further, no reasons were discussed for eliminating 25 villages from the dataset. Therefore, we assessed the study as having high risk of confounding, outcome measurement and reporting bias. Yet, we did not identify any implementation issues or limitations because of spill-overs, crossovers or contamination.

Six of our inclusive community-driven development studies used an experimental design. These include Beath (2008, 2013 and 2015); Laudati (2018); Mvukiyehe (2020); and Van Der Windt (2018). Overall, no major limitation was identified for this set of studies where

two thirds were assessed as having a low risk of bias. As detailed in Figure 55 below, high risk of bias was not identified in any RoB domains of the sample of included studies. The only concerns observed in RCT studies evaluating inclusive community driven development interventions were the lack of a balance table in the report (Beath 2015) and risk of selection bias because of a large proportion of villages were not surveyed during the follow-up due to security conditions (Beath 2013). Other than that, no limitations were identified with regards to unit of analysis, deviations from intended intervention, performance, outcome measurement and reporting.

Figure 55: Inclusive CDD studies RCT RoB assessment



Qualitative studies, process evaluations and project documents

We identified six additional documents related to two programmes covered by the inclusive CDD group.

- National Solidarity Programme (NSP) (Afghanistan): four qualitative documents
- Tuungane 1 (DRC one qualitative and one descriptive quantitative documents)

Most (n = 5) of the included qualitative studies were marked as high empirical quality, with the rest (n = 1) marked as moderate.

Synthesis of findings

Quantitative findings

Effects of inclusive community driven development on women's access to rights, services and opportunities

Beath and colleagues' (2013) experimental study in Afghanistan was the only study evaluating the impact of inclusive community driven development on women's access to rights, services and opportunities. Their report included five effects that fell into this outcome category (e.g., girl aspires to work in professional occupation). The effects ranged from very small, negative estimates ($g = -0.004$, [95% CI: -0.07 to 0.06]) to very small, estimates ($g = 0.09$, [95% CI: -0.02 to 0.20]). We assessed the study was assessed as having low risk of bias.

Effects of inclusive community driven development on women's access and ownership to assets, credit and income

Beath and colleagues' (2013a) and Beath and colleagues' (2013b) experimental study in Afghanistan was the only one evaluating the impact of inclusive community driven development on women's access and ownership to assets, credit and income. Their reports included four effects that fell into this outcome category (e.g., generated income in past year). The effects ranged from very small, negative estimates ($g = -0.01$, [95% CI: -0.05 to 0.04]) to small, positive effects ($g = 0.11$, [95% CI: 0.05 to 0.17]). We assessed the studies as having low risk of bias and some concerns, respectively.

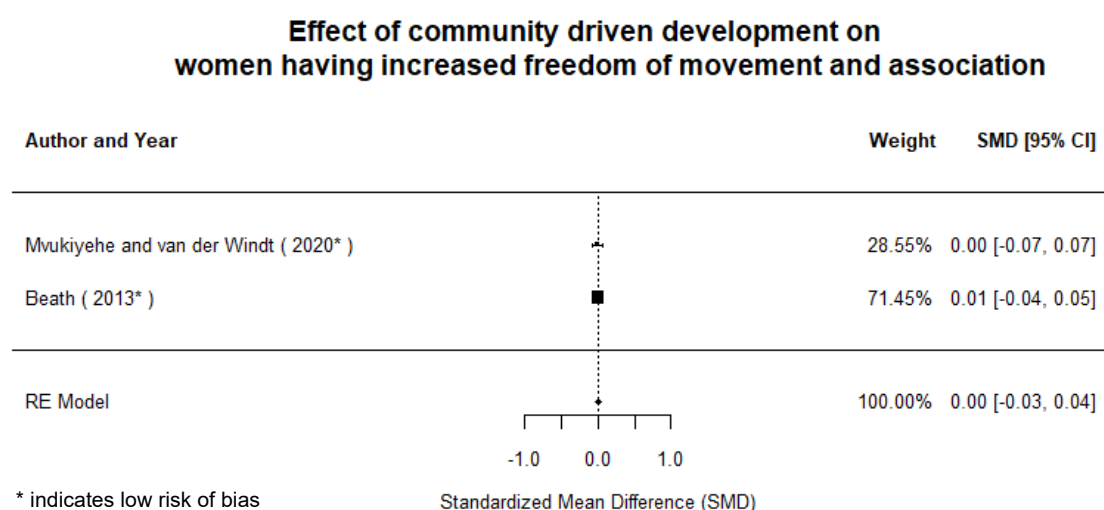
Effects of inclusive community driven development on women having more and better control over their bodies and sexual health

Beath and colleagues' (2013a) experimental study in Afghanistan was the only study evaluating the impact of inclusive community driven development on women having more and better control over their bodies and sexual health. There was a very small, but not statistically significant, estimate ($g = -0.02$, [95% CI: -0.06 to 0.03]), and we assessed the study as having low risk of bias.

Effects of inclusive community driven development on women having increased freedom of movement and association

For increased freedom of movement and association, two studies reported disaggregated data for women, thus we included a total of $k = 2$ studies in the analysis. We assessed both studies as having low risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.004$ (95% CI: -0.03 to 0.04). Therefore, the average outcome did not differ significantly from zero ($z = 0.21$, $p = 0.83$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 56 CDDBA3. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(2) = 0.02$, $p = 0.89$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). Given the small number of studies and that there was no heterogeneity among the effects, moderator analyses were not possible.

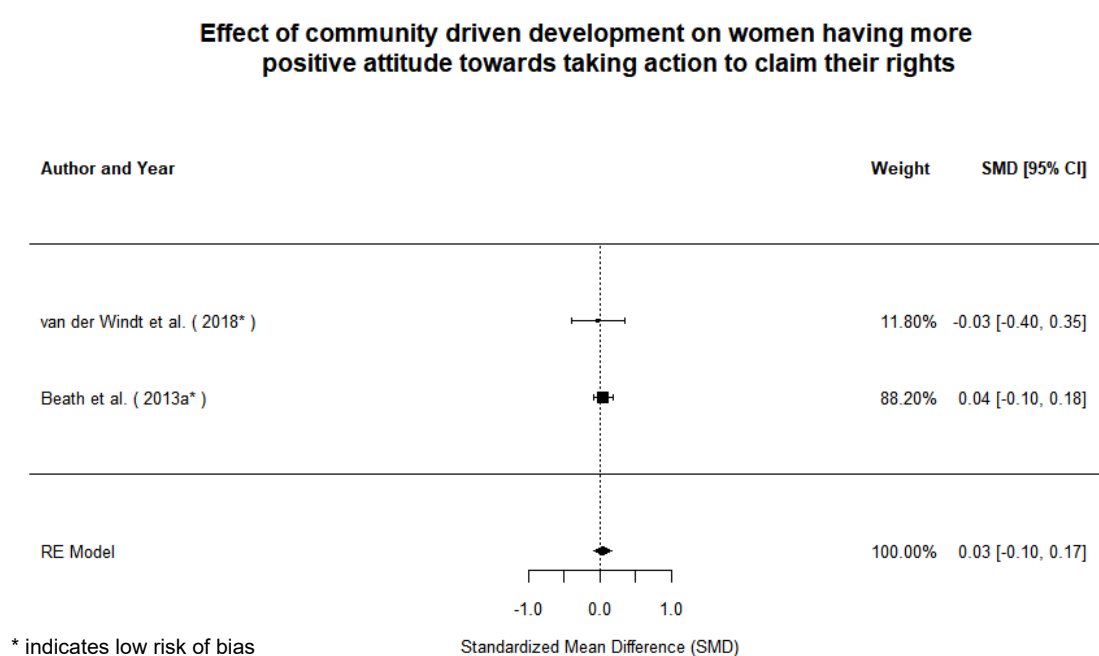
Figure 56: CDDBA3: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of inclusive community driven development on women having more positive attitude towards taking action to claim their rights

For women having more positive attitude towards taking action to claim their rights, two studies reported disaggregated data, thus we included $k = 2$ studies in the analysis. We assessed both studies as having low risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.03$ (95% CI: -0.10 to 0.17). Therefore, the average outcome did not differ significantly from zero ($z = 0.52$, $p = 0.60$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 57 CDDBA5. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.11$, $p = 0.74$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, and given there was no heterogeneity among the effects, moderator analyses were not possible, and tests of publication bias are not valid.

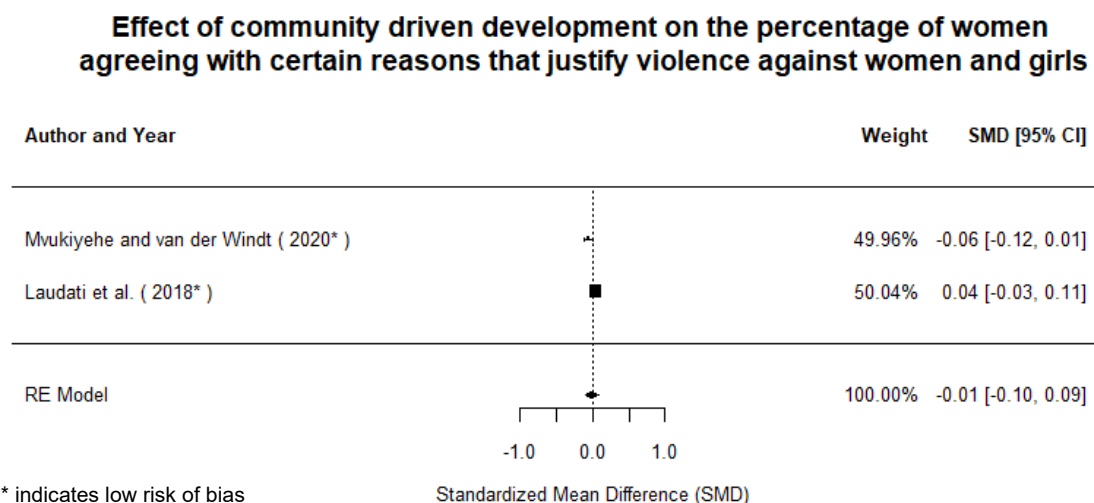
Figure 57: CDDBA5: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of inclusive community driven development on women agreeing with certain reasons that justify violence against women and girls

For women agreeing with certain reasons that justify violence against women and girls, two studies reported disaggregated data, thus we included $k = 2$ studies in the analysis. We assessed both studies as having low risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.01$ (95% CI: -0.10 to 0.08). Therefore, the average outcome did not differ significantly from zero ($z = -0.1512$, $p = 0.88$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 58 CDDBA6. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was some amount of heterogeneity in the true outcomes ($Q(1) = 3.99$, $p = 0.05$, $\hat{\tau}^2 = 0.06$, $I^2 = 74.95\%$). However, with only two studies moderator analyses were not possible, and tests of publication bias are not valid.

Figure 58: CDDBA6: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of inclusive community driven development on women being equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them
Beath and colleagues' (2013a) experimental study in Afghanistan was the only study evaluating the impact of inclusive community driven development on women being equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them. Their report included two effects that fell into this outcome category (days girl attended school in past week and girl correctly completed basic calculation). The effects were very small and positive ($g = 0.07$, [95% CI: 0.01 to 0.12], and $g = 0.10$, [95% CI: -0.01 to 0.21], respectively). We assessed the study as having low risk of bias.

Effects of inclusive community driven development on participation in decision making by women at the household or community level, including during crisis response
Beath and colleagues' (2013a) experimental study in Afghanistan was the only study evaluating the impact of inclusive community driven development on increased participation in decision making by women at the household or community level, including during crisis response. Their report included four effects that fell into this outcome category (e.g., exerts control over income earned). The effects ranged from very small, negative effects ($g = -0.05$, 95% CI [-0.11, 0.01]) to very small, positive effects ($g = 0.05$, [95% CI: -0.01 to 0.11]). We assessed the study as having low risk of bias.

Effects of inclusive community driven development on women's participation in their community
van der Windt and colleagues' (2018) experimental study in Democratic Republic of Congo was the only study evaluating the impact of inclusive community driven development on women participating more in their community. There was a very small, but not statistically significant, effect ($g = -0.05$, 95% CI [-0.43, 0.33]), and we assessed the study as having low risk of bias.

Effects of inclusive community driven development on women forming their own organisations
Beath and colleagues 's (2013b) experimental study in Afghanistan was the only study evaluating the impact of inclusive community driven development on women forming

their own organisations. Their report included three effects that fell into this outcome category (e.g., there exists a village or pan-village women's council). The effects ranged from medium, positive effects ($g = 0.33$, 95% CI [0.14, 0.52]) to large, positive effects ($g = 0.88$, 95% CI [0.69, 1.08]). We assessed the study as having some risk of bias concerns.

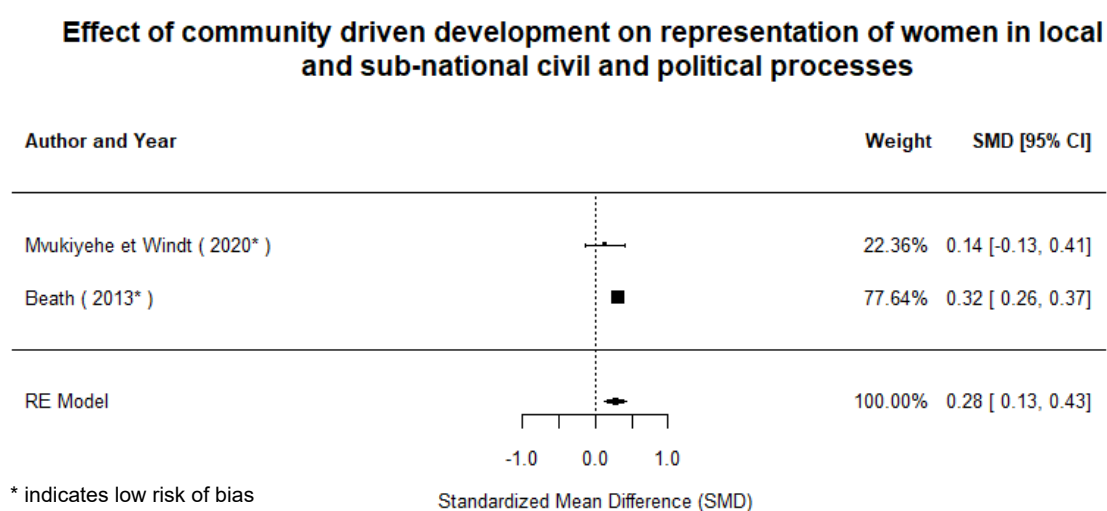
Effects of inclusive community driven development on power holders' awareness and responsiveness to the demands, claims, rights and inputs of women

Beath and colleagues' (2013a) experimental study in Afghanistan was the only study evaluating the impact of inclusive community driven development on power holders' awareness and responsiveness to the demands, claims, rights and inputs of women. There was a very small, but not statistically significant, effect ($g = 0.03$, 95% CI [-0.01, 0.08]), and we assessed the study as having low risk of bias.

Effects of inclusive community driven development on representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration

Two studies reported disaggregated data for representation of women in local and sub-national civil and political processes, including during peace building and post conflict restoration, thus we included a total of $k = 2$ studies in the analysis. We assessed both studies as having low risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.28$ (95% CI: 0.13 to 0.43). Therefore, the average outcome differed significantly from zero ($z = 3.67$, $p < 0.001$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 59 CDDCA2. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 1.66$, $p = 0.20$, $\hat{\tau}^2 = 0.01$, $I^2 = 39.79\%$). With only two studies moderator analyses were not possible and tests of publication bias are not valid.

Figure 59: CDDCA2: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of inclusive community driven development on institutional mechanisms and structures to monitor human rights violations and ensure the security, safety and health of women and girls

Beath and colleagues' (2013a) experimental study in Afghanistan was the only study evaluating the impact of inclusive community driven development on institutional mechanisms and structures monitor human rights violations and ensure the security, safety and health of women and girls. There was a very small, but not statistically significant, effect ($g = 0.01$, 95% CI $[-0.03, 0.05]$), and we assessed the study as having low risk of bias.

Effects of inclusive community driven development on attention and focus on the needs and priorities of women and girls, and other vulnerable groups during relief and recovery in conflict and post-conflict settings

Beath and colleagues' (2013a) experimental study in Afghanistan was also the only study evaluating the impact of inclusive community driven development on increased attention and focus on the needs and priorities of women and girls, and other vulnerable groups during relief and recovery in conflict and post-conflict settings. Their report included two effects that fell into this outcome category (women's views considered in aid allocation and males' response on women's views considered in aid allocation). The effects were positive and statistically significant, very small for males' response ($g = 0.08$, 95% CI $[0.02, 0.14]$) and small for women's views being considered for aid allocation ($g = 0.12$, 95% CI $[0.05, 0.19]$). We assessed the study as having low risk of bias.

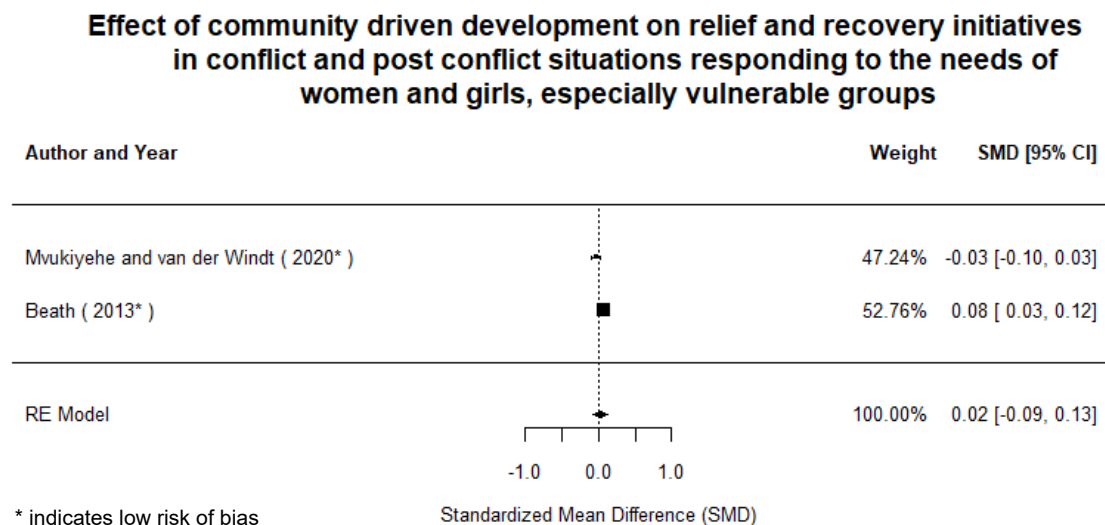
Effects of inclusive community driven development on increased representation of women at all levels as decision-makers in post-conflict countries

Beath and colleagues' (2013a) experimental study in Afghanistan was also the only study evaluating the impact of inclusive community driven development on increased representation of women at all levels as decision-makers in post-conflict countries. Their report included four effects that fell into this outcome category (e.g., women helped mediate most recent dispute). The effects ranged from very small, positive effects ($g = 0.05$, 95% CI $[-0.02, 0.12]$) to small, positive effects ($g = 0.10$, 95% CI $[0.04, 0.16]$), with all of the effects being positive. We assessed the study as having low risk of bias.

Effects of inclusive community driven development on increased community support for women's and children's human, economic and legal rights

Two studies reported disaggregated data for community support for women's and children's human, economic and legal rights, thus we included a total of $k = 2$ studies in the analysis. We assessed both studies as having low risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.02$ (95% CI: -0.09 to 0.13). Therefore, the average outcome did not differ significantly from zero ($z = 0.44$, $p = 0.66$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 60 CDDCA8. According to the Q -test, the true outcomes appear to be heterogeneous ($Q(3) = 7.43$, $p < 0.01$, $\hat{\tau}^2 = 0.01$, $I^2 = 86.53\%$). With only two studies, moderator analyses were not possible.

Figure 60: CDDCA8: Forest plot showing the observed outcomes and the estimate of the random-effects model



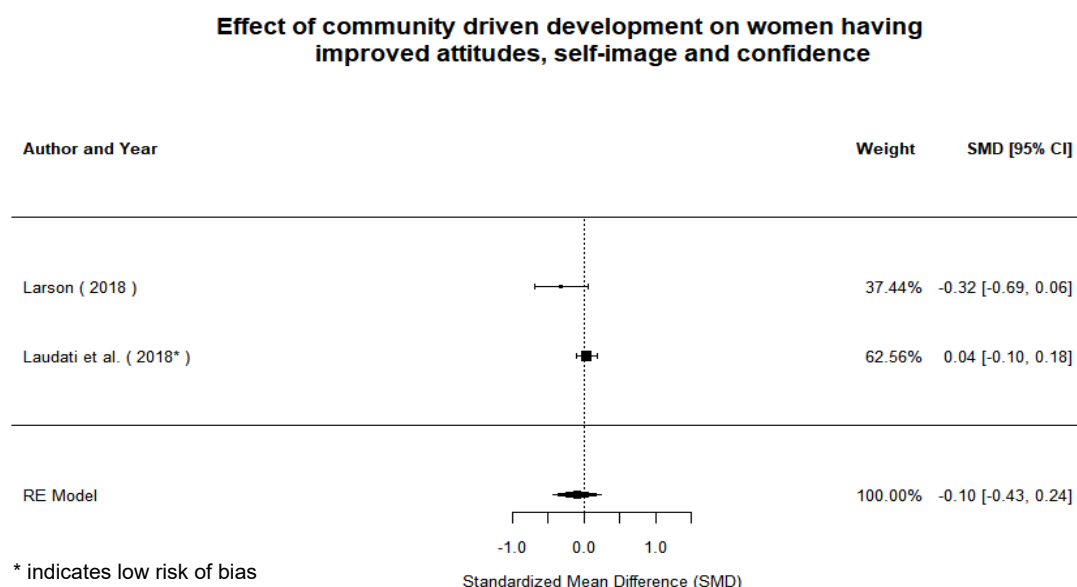
Effects of inclusive community driven development on CSOs being more inclusive to all genders

Beath and colleagues' (2013b) experimental study in Afghanistan was the only study evaluating the impact of inclusive community driven development on CSOs being more inclusive to all genders. Their report included two effects that fell into this outcome category (male and female respondent believes it is appropriate for women to work with NGOs). The effects were very small and positive, but not statistically significant for male respondents ($g = 0.03$, 95% CI [-0.03, 0.09]) whereas statistically significant for female respondents ($g = 0.08$, 95% CI [0.02, 0.14]). We assessed the study as having some risk of bias concerns.

Effects of inclusive community driven development on women having improved attitudes, self-image and confidence

For women having improved attitudes, self-image and confidence, two studies reported disaggregated data, thus we included $k = 2$ studies in the analysis. We assessed one study as having low risk of bias, none as some concerns, and one as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.0958$ (95% CI: -0.4289 to 0.2374). Therefore, the average outcome did not differ significantly from zero ($z = -0.5634$, $p = 0.5732$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 61 CDDCB3. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was some amount of heterogeneity in the true outcomes ($Q(1) = 2.9941$, $p = 0.0836$, $\hat{\tau}^2 = 0.0411$, $I^2 = 66.6009\%$). However, with only two studies moderator analyses were not possible, and tests of publication bias are not valid.

Figure 61: CDDCB3: Forest plot showing the observed outcomes and the estimate of the random-effects model



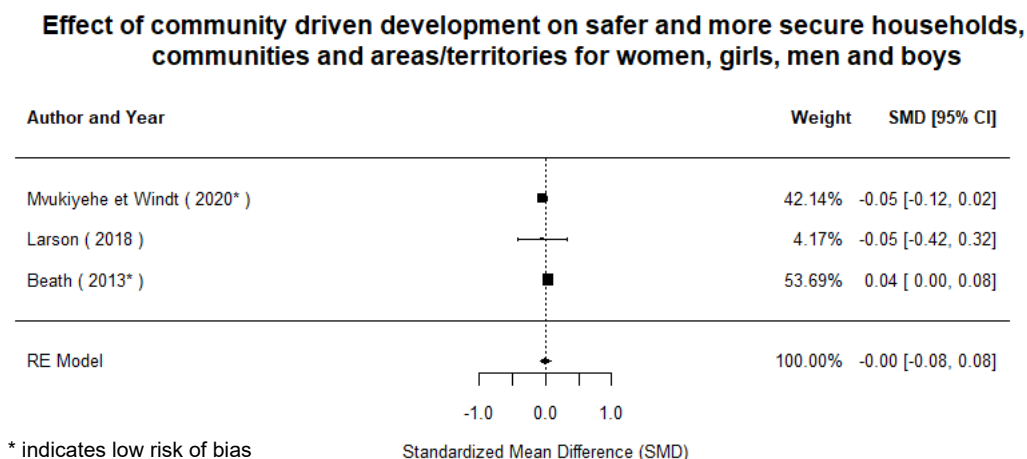
Effects of inclusive community driven development on attitudes and support for women's economic, social and human rights by men, household and family members and community members

Beath and colleagues' (2013) experimental study in Afghanistan was the only study evaluating the impact of inclusive community driven development on attitudes and support for women's economic, social and human rights by men, household and family members and community members. Their report included four effects that fell into this outcome category (e.g., men accept women seeking national office). The effects ranged from very small, negative point estimates ($g = -0.00$, [95% CI: -0.08 to 0.07]) to very small, positive point estimates ($g = 0.03$, [95% CI: -0.04 to 0.10]) We assessed the study as having low risk of bias.

Effects of inclusive community driven development on safer and more secure households, communities and areas/territories for women, girls, men, and boys

We included a total of $k = 3$ studies in the analysis. We assessed two studies as low risk of bias, none as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.05 to 0.04 . The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.0003$ (95% CI: -0.08 to 0.08). Therefore, the average outcome did not differ significantly from zero ($z = -0.01$, $p = 0.99$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 62 CDDCB6.

Figure 62: CDDCB6: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(3) = 5.08$, $p = 0.08$, $\hat{\tau}^2 = 0.00$, $I^2 = 60.62\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.50 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

Effects of inclusive community driven development on quality of relationships between women and their household and community members

Beath and colleagues' (2013) experimental study in Afghanistan and Mvukiyehe et Windt's (2020) experimental study in Democratic Republic of Congo were the only studies evaluating the impact of inclusive community driven development on improved quality of relationships between women and their household and community members. While there are enough studies here, the outcomes do not appear to be similar enough to combine in a meta-analysis.

Beath and colleagues (2013) included three effects that fell into this outcome category (e.g. dispute last year). The effects ranged from small, negative effects ($g = -0.17$, 95% CI $[-0.37, 0.03]$) to very small, positive effects ($g = 0.08$, 95% CI $[-0.06, 0.21]$). We assessed the study as having low risk of bias.

Mvukiyehe et Windt (2020) included two effects that fell into this outcome category (intra-village trust in village member and intra-village trust in chief). Both effects were very small, negative and not statistically significant ($g = -0.01$, 95% CI $[-0.11, 0.10]$), and ($g = -0.06$, 95% CI $[-0.17, 0.05]$), and we assessed the study as having low risk of bias.

Qualitative findings

We conducted a thematic synthesis on the six linked qualitative studies to the included inclusive community-driven development interventions. As indicated above, this thematic synthesis aims to identify themes related to the interplay of intervention design, intervention implementation, target population, and contextual variables with intervention outcomes and effects. In total, we identified 27 descriptive themes, which we configured into six analytical themes (Online Appendix N). These six analytical themes present the synthesis results and are discussed in more detail below.

Theme 1: Socio-economic characteristics affect participation whilst gender-balanced participation in committees, particularly women's representation, can determine the potential benefits derived from CDD projects.

Women's education and literacy level, farmland ownership and residential proximity to the project (e.g., road networks) affected participation and benefits from CDD projects. Involvement and/or representation of women in project implementation committees may not only heighten community members' feeling of ownership, self-value and respect, and reliability on government, but also support changes in cultural and gender norms (e.g., involvement in girls' school development projects brought about positive changes in community members' attitude towards girls' education (Komorowska 2016)).

Successful implementation and sustainability of CDD projects can be facilitated by men and women's cooperation from the ground up. Bottom-up initiatives can be successful even when existing cultural norms are sustained. Mobilisation and awareness raising among community members can enhance ownership of community projects. Women's participation in decision making processes was central to ensuring that women's development needs were captured and protected as a woman reflected: *"In the past, we fetched water from the river but now there is no water in the river, we go in the night to wells far away for water, we take turns and sometimes wait for three or four hours before we can get the water."* (Boesen 2004).

Equal opportunity, access, contribution, participation and representation of men and women were also important for smooth implementation. Humphreys, de la Sierra and Van der Windt (2012) reported men's dominance in access to information, contribution to and use of the projects. Projects and committees that are developed without constructive involvement of the community members and formed just for the sake of the intervention, may have a negative impact on the community members as they may see it as forcefully imposed. Hence, mobilising and involving community members meaningfully and in a timely manner was crucial for CDD projects.

Theme 2: Capacity and legitimacy of the project committee members are crucial to the delivery of the projects.

Members of the project committees can play vital roles throughout the life cycle of inclusive CDD projects. To represent and empower women through these interventions, it was found to be crucial that the composition of the project committees ensure gender parity. In many cases, gender parity was mandated by design. Although Humphreys and colleagues (2012) found a certain level of women's representation in the committees even without such mandated parity, they explained that committees designed with mandated gender parity tend to have a large number of women. Essentially, committee members' knowledge of project design, resource management, operational and logistics management for implementation, and monitoring and evaluation were important for sound implementation of the projects. Barakat and Strand (2006) reported that members of the majority of the project committees that they assessed in their study required training.

Legitimacy of the committee reflected its authority, peoples' trustworthiness and cooperation with relevant government and non-governmental agencies. Boesen (2004) reported that: *"In many districts, the CDC was already perceived as an established and legitimate institution in a wider context than NSP. This was due to two reasons. One*

reason was that the CDC is a structure which is elected by the whole community on a democratic basis without direct deference to traditional local powerholders. Furthermore, the formal authorisation of the CDC by the government through (the) Ministry for Rural Rehabilitation and Development (MRRD), and the official stamp and registration in the provincial MRRD office, was perceived as essential with regard to the legitimacy of the CDC in the communities.”

Theme 3: Timely disbursements of funds and proper human resource management system of government personnel is essential to the successful completion of CDD projects.

Successful completion and sustainability of inclusive CDD projects critically hinge on timely implementation, regular payment and funds release, management of human resources, and maintenance of physical resources. Delayed and faulty implementation and approval processes were often identified as major challenges. Barakat and Strand (2006) noticed delays in payment and fund release which resulted in delay in procurement of critical goods required to implement the projects. Such delays often cause tension between community/committee members and corresponding funding bodies.

Timely implementation of critical project components was also mentioned as a crucial factor for smooth delivery of planned projects. For instance, community mobilisation and awareness raising were reported to be critical for encouraging participation and it was essential that these activities were carried out at the onset of the project. Emphasising on the importance of timing, Barakat and Strand (2006) reported that: *“Most seriously, the delays were blamed for damage caused to the quality of projects; for example, where the building of infrastructure could not be completed before winter.”*

Timing of intervention targeting women was also key to its delivery and success. When interventions coincide with annual harvesting seasons, participation of the community and especially of women were reportedly near impossible (Laudati, Mvukiyehe and van der Windt 2018). These kinds of interventions must be well planned and timed to ensure their effectiveness in reaching the target populations.

Bhatia, Jareer and McIntosh (2018) also pointed out that it is important to ensure that all sections of the community can equally benefit from the project. When project funds are restricted to cover only a segment of the affected community, rivalry and grievances arise that may challenge the delivery and success of the projects.

Proper human resource management, defined responsibilities and workload distribution were reported to be essential for smooth operation. Overburdening staff with workload and engaging project staff in out-of-project activities because of their expertise can hinder implementation. Adequate allocation and maintenance of physical resources were also important for smooth implementation and sustainability of the projects.

Theme 4: Distrust, power dynamics and frictions, corruption and nepotism can undermine communities’ involvement and support for the project.

Communities demonstrated overwhelming preference for funded community projects to be managed by external bodies rather than local authorities or traditional systems. Corruption and nepotism were major factors fuelling peoples’ lack of trust. For instance,

a participant complained that: *“We now have less qualified teachers because those who were qualified were chased by the current schoolmaster; instead, he hired the ones according to his affinity/kinship relations with them and some students left the school for other schools that are far from here. As a result, the overall quality of the teaching has reduced.”* Another participant reported that, *“Rather than this [reluctance to use] being a result of a badly built centre, it was found that villagers had stopped attending the Tuungane 1 facility because the head of the health care centre was reportedly charging patients as much as five times more than neighbouring clinics”* (Laudati et al. 2018).

Against their reputation for corruption and nepotism, the qualitative findings suggested that external management was perceived to better empower the populations, improve efficiency, and reduce corruption. The experience of the community with failed government promises and corrupt practices showed in form of reluctance towards the new programme or intervention. Sometimes, directors of newly provided facilities and services exhibited the same level of corruption and nepotism, which denied beneficiary communities the intended quality, access, use and control of the facilities/services. However, good governance and project management can also improve public trust and reliability towards the government. A staff member expressed that: *“The prestige projects, the ones that are better built, have a huge psychological effect on the community. Having a proper school. A proper health centre. To me it has an impact on how the infrastructure is maintained after the programme leaves, and on how the programme is evaluated by the communities”* (Laudati et al. 2018).

Power dynamics within the community can affect the participation of different stakeholders especially if they affect traditional governance structures. Laudati and colleagues (2018) reported that power dynamics is *“important for the level of overall participation and engagement in the project by different actors, inclusion in the decision-making process regarding project identification, and as discussed here, the ability of different beneficiaries to access and have control over facility services”*. Rivalry among different religious groups (e.g., Protestant and Catholic Churches), as well as inter-village tensions - resulting from their placement in relation to the intervention - can hamper the delivery and success of the intervention (Humphreys, de la Sierra and Van der Windt 2012).

Some qualitative findings highlighted the challenges arising from weak institutional infrastructure linked to rivalry among different government ministries. Bureaucracy and administrative bottlenecks (e.g., the problem of high staff turnover; the existence of too many forms; constant changes to the operational manual or rules of business) and rigid mindset regarding how things should be done also pose serious challenges (Barakat and Strand 2006). Collaboration, cooperation and coordination between relevant public and private agencies are crucial as well.

Programme planners and implementers need to take the issue of strategic involvement of the established community system into critical consideration. It was reported that if inclusive CDD projects potentially affect vested interests of local influential individuals or groups, they may position themselves within the project committees and thereby reorient the projects in their favour (Komorowska 2016). Nevertheless, the study also mentioned that recently a positive change could be noticed where influential elders were giving up powers to the committee or acting as watchdogs to oversee the implementation process.

Financial irregularities, leakage or incidences of larceny such as robbery or theft can also limit success of the projects. For the Tuungane 1 project in DRC, a chief who mentioned: *“the project hasn’t been well executed because of robbery and mismanagement”* (Laudati et al. 2018).

Theme 5: Insecurity and vested influence of powerful individuals can hamper project implementation.

Several external factors pose serious threats to ICDD projects. Security is crucial for the success of development interventions. Security issues range from low-intensity conflicts to civil war, sensitivity around poppy industries (Barakat and Strand 2006), resistance from insurgent groups and violence against women (Bhatia et al. 2018). Where insecurity prevails, intervention staff can be endangered, the community is scared to participate, provided facilities and services are hijacked by the insurgents, and the use and sustainability of the project/facility/services provided is at risk. reported that respondents mentioned: *“CDCs were specifically targeted by insurgent groups as part of their campaign against the government. CDC members reported that they were warned to stop their work by the Taliban, and that they were afraid of being associated with the CDC,”* while another participant said, *“A more commonly mentioned reason in other communities was that, as insecurity rose, government agencies and NGOs withdrew from the area, and there were no projects to implement and therefore no reason for CDCs to continue with meetings”* (Bhatia et al. 2018).

In many cases, the government was unable to tackle these insurgencies. Some participants reported that despite being occupied by insurgents, some communities were able to continue the project by actively collaborating with them. In addition, projects can undermine vested interests of local influential persons who can cause hindrance to implementation. External influence on committee elections can also have negative impacts. Technical issues external to project design and implementation (e.g., seasonality, depth of groundwater level, geographic isolation, drought or other environmental stressors) can affect implementation (Laudati et al. 2018).

Theme 6: Gender and social norms deeply rooted in cultural beliefs and practices are rigid. This applies particularly if changes in such norms and beliefs undermine the authority of influential individuals or groups and hinder implementation.

Even when women are afforded behaviour change communication, safe spaces, and income-generating initiatives, their level of participation and therefore their potential benefit may be constrained by such norms. In Afghanistan, qualitative findings from the National Solidarity Programme (NSP) identified conservative religious beliefs and core values or norms about women’s mobility and participation in community affairs as key obstacles limiting women’s voting rights, membership in committees, and holding decision making positions (Bhatia et al. 2018). Furthermore, programme evaluators acknowledge that positive changes in norms are usually protracted. One NSP woman complained that: *“Men stop you if you start talking about women’s problems, they say that women should be quiet”* (Boesen 2004).

NSP activities typically involved the participation of women in income generating activities of home-based products (for instance, tailoring, handicrafts, carpet weaving). Due to the low marketing potential of these products and women’s limited access to markets due to cultural restrictions, women could not make enough profit from such

activities (Boesen 2004). The involvement of women and girls in projects aimed to improve women's or girls' access to education or health services and could activate changes in traditional norms and beliefs (Komorowska 2016). In the context of Afghanistan's NSP, in communities facing the shared post-war realities, men tend to recognise the importance of women to community reconstruction and development (Boesen 2004).

The involvement of women in community activities rooted within ICDD interventions may pose a threat to established customary governance systems, vested interests and even religious strongholds. Such activities include the building of important communal structures, water and power supply infrastructure. In particular, leadership opportunities for women in the Tuungane 1 for community-related development. In the Tuungane 1 programme, traditional norms and beliefs remained an impediment to this possibility in some instances as one woman responded: *"only men have positions of power"*. Another woman felt that: *"Tuungane did well [for women] as it created a space for a woman to be involved; Tuungane is democratic, it doesn't discriminate and it called for men to understand that women are equally able to make positive change in society."*

In response to rebuttals on such customs and vested interests, it might not be unlikely that influential individuals or groups would try to infiltrate the committees and influence the voting of their elected officials to ensure they retain control (Bhatia et al. 2018). A Tuungane 1 community member advised that: *"Tuungane has to monitor the work on the ground constantly to avoid corruption"* while another expressed uncertainty on the status of a building: *"The CDC project was good but the one we have needed an engineer who let us down; right now, we don't know whether the money has been used or not or why the building is not yet finished."* (Humphreys et al., 2012). Furthermore, communities tend to fear that programme implementers are pretentiously introducing Westernised cultures, intending to erode traditional culture (Boesen 2004).

Discussion

Quantitative meta-analysis revealed that ICDD interventions had positive significant effects on representation of women in local and sub-national civil and political processes. However, effects on community and household level outcomes such as safer and more secure households and communities were non-significant. The intervention group was covered by three programmes and the associated GRADE analysis showed high confidence in the finding related representation of women in civil and political processes and low confidence in the community and household findings.

A complementary qualitative synthesis was carried out to put forward a set of recommendations that may help design more effective ICDD interventions that can support the translation of political representation effects on community and household level outcomes targeting broader empowerment objectives.

- From the qualitative evidence, a number of implementation themes emerged which shed some light on the challenges of translating representation outcomes into more effective project government and implementation. Even when interventions succeeded in enhancing the representation of women on project committees and other decision-making structures, capacity gaps across governance structures impeded project implementation and delivery.

- In addition, the composition of committees and representatives on decision-makers structures mattered largely for the legitimacies of these bodies within the communities targeted by the intervention. While interventions reported progress in ensuring a greater representation of women on these bodies, qualitative evidence referenced challenges in ensuring background diversity of these women. That is, socio-economic characteristics such as class, education level, profession emerged as important for the perceived legitimacy of governance bodies and decision-makers within the communities. Projects often struggled to ensure this additional level of diversity.
- Taken together, governance bodies' relative lack of capacity and legitimacy emerged as two possible explanations as to why ICDD interventions achieved less success at the household and community level.
- Furthermore, qualitative evidence pointed to a complex navigation of competing governance systems within communities as an additional explanatory lens why intervention effects were not observed equally at household and community level. That is, in most target communities, governance structures introduced by ICDD coexist with traditional governance authorities (e.g. local chiefs) as well as local government representatives (e.g. elected district officials). Navigating this co-existence required a careful balancing of politics and power relations, which limited the space for project implementation and reach in some instances.
- All of the themes above, i.e., governance legitimacy, capacity, and co-existence, presented entry points for powerful elites and vested interests to undermine project implementation and impact; with an often-observed alignment and intersection of these actors' agenda with restrictive social norms and preserving existing power imbalances between men and women.
- In terms of intervention design and implementation recommendation, the above qualitative findings and explanatory lenses suggest that investment in 'good governance' for ICDD projects is important, not overlooking more granular project management capacities. Likewise, intervention design should anticipate conflicts to emerge where multiple governance systems coexist and to outline proactively mitigation strategies and a space and contribution of the new ICDD-introduced structures within the communities.

Summary of findings and discussion

We included three programmes in eight countries in Latin America, East Asia and Pacific, South Asia and Sub-Saharan Africa that evaluated the effect of inclusive community driven development in L&MICs. We were able to examine effects on the following: women have increased freedom of movement and association, increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration, women have improved attitudes, self-image and confidence, safer and more secure household, communities and areas/territories for women, girls, men, and boys. Our included studies report against all three secondary outcomes (Resources, Agency and Achievement) and six of the nine immediate outcomes. Overall, the GRADE assessments (Table 27) generally indicate a range of very low to high certainty in this body of evidence.

Six qualitative analytical themes buttressed the quantitative findings by illuminating the importance of localised capacity building, recognising power dynamics and ensuring

particular benefits from projects by balancing representing in committees and potential buy-in by powerholders. With positive and significant impact on women's representation outcomes, ICDD interventions may be able to promote a greater role for women in decision-making as well as deciding key outcomes for their communities. While other outcomes relating to secure households were not significant, the emphasis on women's participation shows the importance of planning ICDD interventions that allow women to have their voices and opinions heard.

Table 27: GRADE summary of findings and certainty of evidence on inclusive community driven development

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(AA1) Women have access to rights, services and opportunities										
1	RCT	not serious	serious ^a	not serious	not serious	none	4307	One negative and four positive effect estimates with a 95% CI range of -0.09 to 0.19.	⊕⊕⊕○ MODERATE	Limited Importance
(AB2) Women have increased access to and ownership of assets, credit and income										
1	RCT	not serious	serious ^a	not serious	not serious	none	8032	Two very small positive and two very small negative effect sizes with a 95% CI range of -0.07 to 0.20.	⊕⊕⊕○ MODERATE	Limited Importance
(BA2) Women have more and better control over their bodies and sexual health										
1	RCT	not serious	serious ^a	not serious	not serious	none	8032	One negative and one zero effect estimates with a 95% CI range of -0.06 to 0.06.	⊕⊕⊕○ MODERATE	Critical
(BA3) Women have increased freedom of movement and association										
2	RCT - 2	not serious	not serious	not serious	serious ^b	none	9134	SMD 0.00 SD (-0.03 lower to 0.04 higher)	⊕⊕⊕○ MODERATE	Limited Importance
(BA5) women have more positive attitude towards taking action to claim their rights										
2	RCT - 2	not serious	not serious	not serious	serious ^c	none	900	SMD 0.03 SD higher (0.1 lower to 0.17 higher)	⊕⊕⊕○ MODERATE	Limited Importance
(BA6) Reduced percentage of women agreeing with certain reasons that justify violence against women and girls										
2	RCT - 2	not serious	serious ^d	not serious	serious ^e	none	6751	SMD 0.01 SD lower (0.1 lower to 0.09 higher)	⊕⊕○○ LOW	Critical
(BA7) Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them										
1	RCT	not serious	serious ^a	not serious	serious ^f	none	4931	One negative and two positive effect estimates with a 95% CI range of -0.06 to 0.20	⊕⊕○○ LOW	Important, but not critical
(BB1) Women participate more in their community										
1	RCT	not serious	serious ^a	not serious	not serious	none	916	Six negative and two positive effect estimates with a 95% CI range of -0.11 to 0.11	⊕⊕⊕○ MODERATE	Limited Importance
(BB2) Women participate more in their community										

Certainty assessment							Sample size	Effect	Certainty	Importance
N ₂ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
1	RCT	not serious	serious ^a	not serious	not serious	none	916	Two negative effect estimates of -0.05 with a 95% CI range of -0.43 to 0.33	⊕⊕⊕○ MODERATE	Limited Importance
(BB4) Women increasingly form their own organisations										
1	RCT	serious ^g	serious ^a	not serious	serious ^h	strong association	424	Three positive (two large) effect estimates with a 95% CI range of 0.13 to 1.08	⊕⊕○○ LOW	Limited Importance
(CA1) Power holders have improved awareness and responsiveness to the demands, claims, rights and inputs of women										
1	RCT	not serious	serious ^a	not serious	not serious	none	8021	SMD 0.03 SD higher (0.01 lower to 0.08 higher)	⊕⊕⊕○ MODERATE	Limited Importance
(CA2) Representation of women in local and sub-national civil and political processes (assessed with: self-reports)										
2	RCT	not serious	not serious	not serious	not serious	strong association	2319	SMD 0.28 SD higher (0.13 higher to 0.43 higher)	⊕⊕⊕⊕ HIGH	Limited Importance
(CA3) Institutional mechanisms and structures monitor human rights violations and ensure the security, safety and health of women and girls										
1	RCT	not serious	serious ^a	not serious	not serious	none	8032	SMD 0.01 SD higher (0.03 lower to 0.05 higher)	⊕⊕⊕○ MODERATE	Critical
(CA4) Increased attention and focus on the needs and priorities of women and girls, and other vulnerable groups during relief and recovery in conflict and post-conflict settings										
1	RCT	not serious	serious ^a	not serious	not serious	none	4095	Two positive effect estimates with a 95% CI range of 0.02 to 0.19	⊕⊕⊕○ MODERATE	Limited Importance
(CA5) Increased representation of women at all levels as decision-makers in post-conflict countries										
1	RCT	not serious	serious ^a	not serious	not serious	none	4308	Four positive effect estimates with a 95% CI range of -0.02 to 0.16	⊕⊕⊕○ MODERATE	Limited Importance
(CA8) Community support for women's and children's human, economic and legal rights										
4	RCT - 4	not serious	not serious	not serious	not serious	none	15674	SMD 0.01 SD lower (0.1 lower to 0.08 higher)	⊕⊕⊕⊕ HIGH	Important, but not critical
(CA11) CSOs are more inclusive to all genders										
1	RCT	serious ⁱ	serious ^a	not serious	not serious	none	4643	Two positive effect estimates with a 95% CI range of 0.03 to 0.14	⊕⊕○○ LOW	Limited Importance
(CB3) Women have improved attitudes, self-image and confidence										
2	RCT - 1 QED - 1	very serious ^j	serious ^k	not serious	serious ^l	none	891	SMD 0.1 SD lower (0.43 lower to 0.24 higher)	⊕○○○ VERY LOW	Limited Importance
(CB4) Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members										
1	RCT	not serious	serious ^a	not serious	not serious	none	8968	One null and three positive effect estimates with a 95%	⊕⊕⊕○ MODERATE	Limited Importance

Certainty assessment							Sample size	Effect	Certainty	Importance
N _e of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
								CI range of -0.08 to 0.10		
(CB6) Safer and more secure household, communities and areas/territories for women, girls, men, and boys										
3	RCT -2 QED - 1	serious ^m	serious ⁿ	not serious	not serious	none	13020	SMD -0.00 SD higher (0.08 lower to 0.08 higher)	⊕⊕○○ LOW	Critical
(CB8) Improved quality of relationships between women and their households and communities										
1	RCT	not serious	serious ^a	serious	not serious	none	916	Two negative effect estimates with a 95% CI range of -0.33 to 0.21	⊕⊕⊕○ MODERATE	Limited Importance

CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

a. All single studies downgraded once for inconsistency.

b. a) Rate down if confidence intervals of RE Model cross the threshold: Yes b) Rate down if one study: No, three studies., c) Sample size: Large, not rated down

c. a) Rate down if confidence intervals of RE Model cross the threshold: Yes b) Rate down if one study: No, two studies, c) Sample size: Large enough, not rated down.

d. a) Wide variance of point estimates across studies b) Minimal overlap of confidence intervals (CI), which suggests variation may be more than what one would expect by chance alone c) Statistical criteria, including tests of heterogeneity which test the null hypothesis that all studies have the same underlying magnitude of effect, have a low p-value ($p < 0.05$), indicating to reject the null hypothesis.

e. a) Rate down if confidence intervals of RE Model cross the threshold: Yes b) Rate down if one study: No, two studies c) Sample size: Large, not rated down.

f. Downgraded once because of a wide range of CIs that cross both sides of the threshold of interest.

g. Downgraded once due to uncertainty related to selection bias, and some risk related to other criteria.

h. While all effect estimates are quite large, there is a discrepancy in one estimate against the other two of more than 100%.

i. Downgraded due to issues related to deviation from intended intervention and uncertainty with respect to selection bias.

j. Downgraded twice because 37.44% of weight (or half of the studies) is coming from a high RoB QED, and when the high RoB study is removed, there is only one study remaining and the effect estimate would change considerably.

k. Downgraded because the point estimates are on either side of the threshold. While the CIs do overlap, there is a considerable difference

between the actual point estimates.

l. Downgraded once on imprecision to reflect the large CI on the RE model point estimate.

m. This analysis has been downgraded because one of the four studies presents a high risk of bias due to reporting bias, outcome measurement bias, and selection bias. All other studies are low risk of bias.

n. Downgraded because some studies present positive, while others negative effect sizes. $I^2 = 41.62\%$.

Effects of community-based services on women's access and ownership to assets, credit and income

Nandi and colleagues' (2020) experimental study in India was the only study evaluating the impact of community-based services on women's access and ownership to assets, credit and income. Their report included two effects that fell into this outcome category (paid in cash and any time on paid work). Both point estimates were very small, positive and not statistically significant ($g = 0.02$, [95% CI: -0.05 to 0.09]). We assessed the study as having some risk of bias concerns.

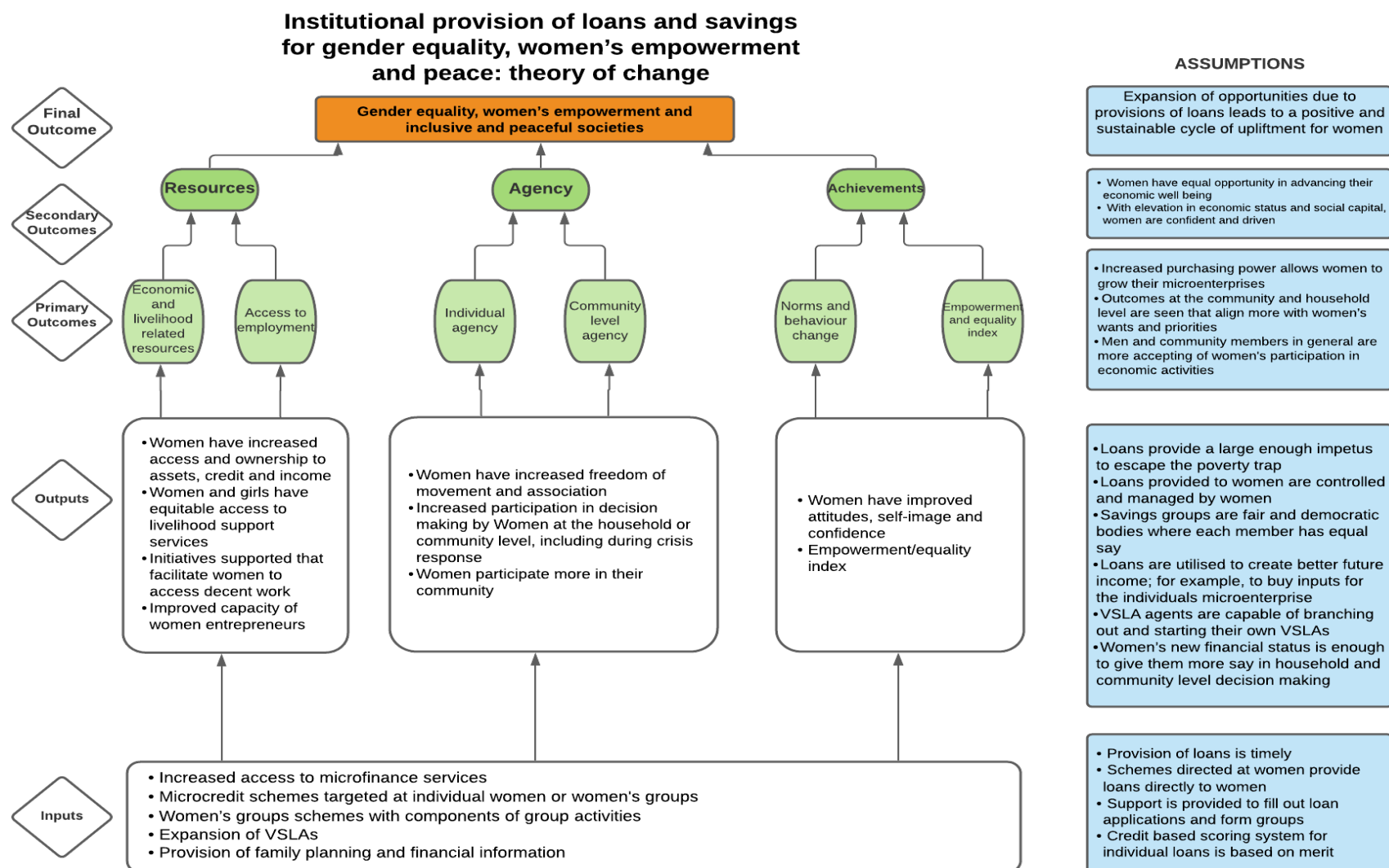
4.3.5 Institutional provision of loans and savings

Institutional provision of loans and savings includes the provision of loans, credit and savings by a formal financial institution or microcredit programme (World Bank 2021).

How do institutional provision of loans and savings affect gender equality, women's empowerment and peace outcomes?

Figure 63 below maps out the causal chain of how the institutional provision of loans and savings may improve gender equality, women's empowerment and peace outcomes. Loans and savings schemes (primarily microcredit schemes) improve women's access to assets, credit and income; hence, expanding the economic resources available to women. At the same time, many schemes offer financial or business training that increase their capacity to interact with these resources. Consequently, women's economic status is improved which simultaneously leads to increased agency, more decision-making ability, improved autonomy and higher empowerment. Moreover, schemes that employ group models give women a *de facto* network of peers thereby increasing their social capital and hence, their empowerment. Therefore, there is a shift towards a more equitable and inclusive society.

Figure 63: Institutional provision of loans and savings for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included seven studies from nine different papers that evaluated the effect of three programmes.

Population

Studies included focussed on low-income women. Some studies extended loans to men as well; however, there was a particular emphasis on enhancing poor women's access to cash.

The included studies evaluated two different programmes and trials in South Asia and Sub-Saharan Africa in five countries: Ethiopia, India, Liberia, Nigeria and Pakistan.

Intervention, inputs and activities

The included studies evaluated a range of different loans and savings activities and inputs including:

- Business loans (n = 1): Offers microfinance loans to businesswomen with higher loans to those at higher loan cycles.
- Microfinance lending (n = 1): Lends microcredit to women entrepreneurs.
- Microfinance scheme for individuals (n = 2): Provide microcredit loans directly to individuals based on some merit-based credit scoring system.
- Group-based lending (n = 3): Microfinance scheme that is contingent on women forming groups that share the liability of the loans.
- Training and activities (n = 2): Provision of financial training alongside loans. Additionally, some programmes provide information on family planning and some conduct activities within loan groups

Table 28: Institutional ision of loans and savings activities features of included studies

Study	Activity/input	Length of treatment	Intervention frequency
Banerjee - 2015	Spandana is a microfinance product targeted towards low-income women that form or participate in loan groups	18 months	Weekly
Duflo - 2013	Spandana is a microfinance product targeted towards low-income women that form or participate in loan groups	12 months	-
Ifelunini - 2012	Increased access to microfinance lending	-	-
Ikenwilo - 2016	Microcredit scheme targeting women with no requirement of group formation/participation	4 months	-
Johnson - 2018	Microcredit loan offered by BRAC to women who participate in groups.	Varies among participants.	-

Study	Activity/input	Length of treatment	Intervention frequency
	Group activities also conducted by BRAC	Average engagement being 10 years	
Tarozzi - 2013	Provision of information on family planning and financial self-help groups	36 months	-
Tarozzi – 2015	Provision of information on family planning and financial self-help groups	36 months	-
Weber - 2014	Increased access to microfinance for women	-	-

Comparison

All our included studies compared treated groups to comparison groups receiving no intervention. One study included multiple treatment arms.

Outcomes

The included studies reported on a number of relevant outcomes, including:

- Women have increased access and ownership to assets, credit and income (n = 4): Women are able to apply for, receive and manage assets/credit and income and have support to manage, claim and execute their assets without pressure or influence from external actors, including male family members, husbands and cultural leaders.
- Improved capacity of women entrepreneurs (n = 2): Capacity can include both technical abilities and resources to achieve a goal. Through skills training, microgrants, and linkage of services, women entrepreneurs are able to pursue their businesses and other pursuits.
- Empowerment/Equality Index (n = 2): Indices using multiple outcomes of the list to aggregate them into a gender equality and/or women's empowerment score

The division of the immediate and secondary outcomes is reported in table 29 below:

Table 29: Institutional provision of loans: secondary and immediate outcomes

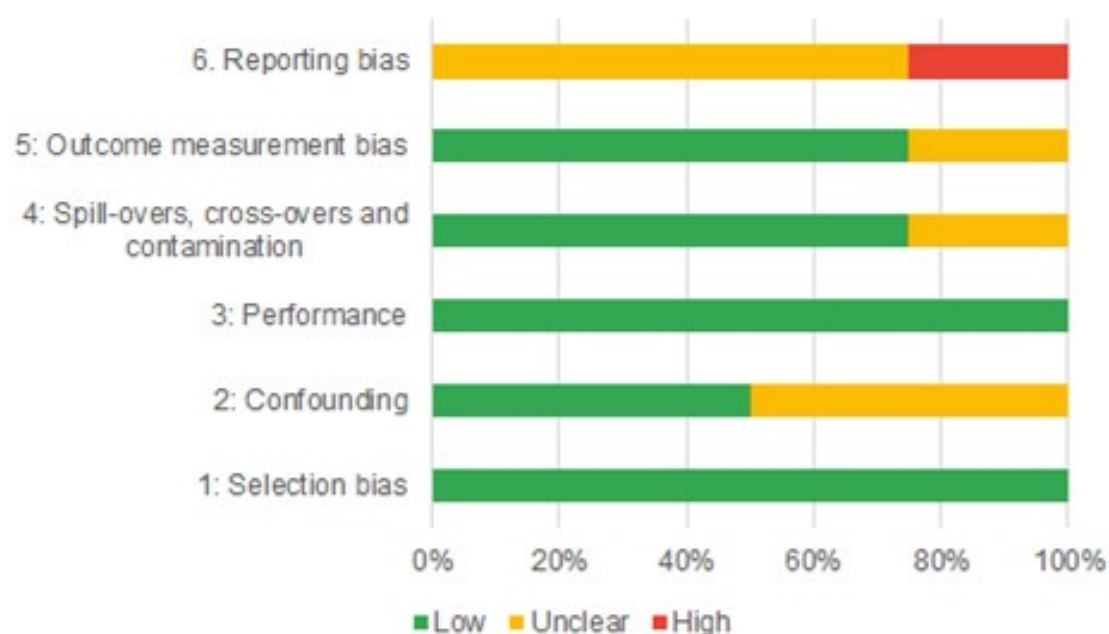
Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	0
	Economic and livelihood related resources	4
	Access to employment	2
Agency	Individual agency	2

Secondary outcome category	Immediate outcome	Number of studies
ability to define one's goals and act upon them and operationalised decision-making;	Community level agency	2
	Institutions supporting agency	0
Achievement ways of being and doing which can be realised by different individuals	Improved systems	1
	Norms and behaviour change	1
	Empowerment index	2

Study design

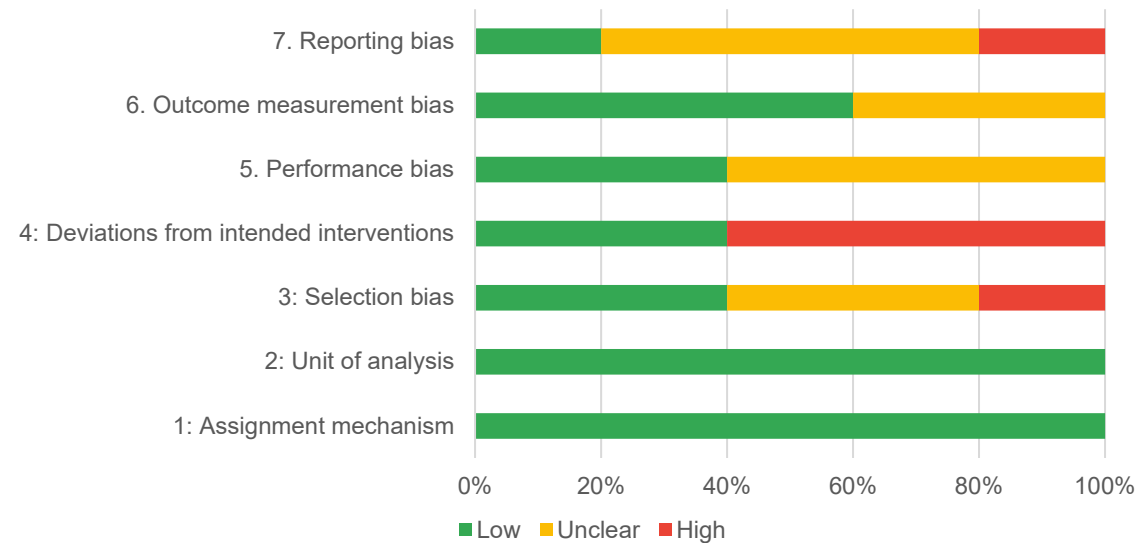
Four institutional provision of loans and savings studies used a QED (Ifelunini 2012; Ikenwilo 2016; Johnson 2018, and Weber 2014). Only one of these studies was assessed as having a high risk of bias, the rest had some concerns related to risk of bias. As detailed in Figure 64 below, the domain in which the high risk was identified was reporting bias, this was because the authors did not provide enough information on the matching process and did not report the results for one of the outcomes because of not finding a significant impact (Weber 2014). A common issue to the rest of QED studies evaluating institutional provision of loans and savings interventions was that authors did not report robustness checks, for which we assessed their reporting bias as unclear. Further, we identified the insufficiency of details provided on characteristics and cluster controls as being unclear whether confounding was an issue or not for two studies (Johnson 2018 and Weber 2014). We did not identify any limitation related to performance or selection bias.

Figure 64: Institutional loans and savings studies QED RoB assessment



Five of the studies evaluating provision of loans and savings used an experimental design (Duflo 2013; Tarozzi 2013). One of the studies had a high risk of bias, and the other presented some concerns related to risk of bias. As detailed in Figure 65 below, the high risk of bias was identified in the deviations from intended interventions domain. The issue in this regard was a potential contamination from other microfinance institutions starting operations in treatment and control groups (Duflo 2013). It was also unclear whether both studies were free from reporting bias since authors do not provide a pre-analysis plan and not enough information was provided to verify that all intended analysis and outcome variables were reported. Finally, Tarozzi and colleagues (2013) did not provide enough information regarding attrition, so we assessed selection bias as unclear. We did not identify any other limitation related to assignment mechanism, unit of analysis, performance, and outcome measurement.

Figure 65: Institutional loans and savings RCT RoB assessment



Qualitative studies, process evaluations and project documents

We identified no linked qualitative studies and were unable to conduct a qualitative evidence synthesis related to savings and loans interventions.

Synthesis of findings

The following subsection presents the results of effectiveness of institutional loans and savings on gender equality, women’s empowerment and peace outcomes.

Quantitative findings

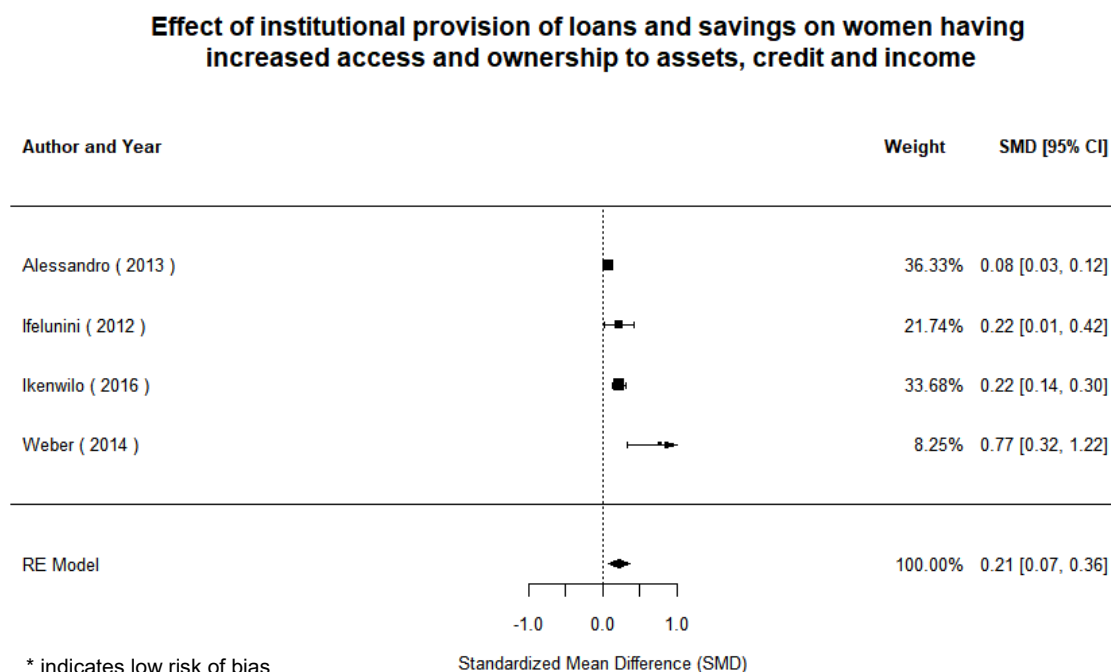
Effects of institutional provision of loans and savings on increased capacity of women to understand and use financial, banking, and business services effectively

Ikenwilo's (2016) quasi-experimental study in Nigeria was the only study evaluating the impact of institutional provision of loans and savings on increased capacity of women to understand and use financial, banking, and business services effectively. There was a large positive and statistically significant effect ($g = 0.49$, [95% CI: 0.41 to 0.57]). We assessed the study as having some risk of bias concerns.

Effects of institutional provision of loans and savings on women having increased access and ownership to assets, credit and income

We included a total of $k = 4$ studies in the analysis. We assessed none of the studies as low risk of bias, three as some concerns, and one as high risk of bias. The observed outcomes ranged from 0.08 to 0.77. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.21$ (95% CI: 0.07 to 0.36). Therefore, the average outcome differed significantly from zero ($z = 2.84$, $p < 0.01$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 66 LSAB2.

Figure 66: LSAB2: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(3) = 17.45$, $p < 0.01$, $\hat{\tau}^2 = 0.01$, $I^2 = 82.80\%$).

An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.50 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

We attempted to test all potential moderators, but several could not be tested because all but one study belonged to a single moderator group (i.e., risk of bias, study design, post-intervention vs. change from baseline). Of the remaining moderators (publication year, whether the model was adjusted for covariates, exposure to intervention, evaluation period), none were significant predictors of heterogeneity.

Effects of institutional provision of loans and savings on women's engagement in other micro, small, and medium-sized enterprises

Banerjee (2015), an experimental study in India was the only study evaluating the impact of institutional provision of loans and savings on women's engagement in other micro,

small, and medium-sized enterprises. The effect was very small and but positive and statistically significant ($g = 0.05$, [95% CI: 0.02 to 0.08]). We assessed the study as having high risk of bias concerns.

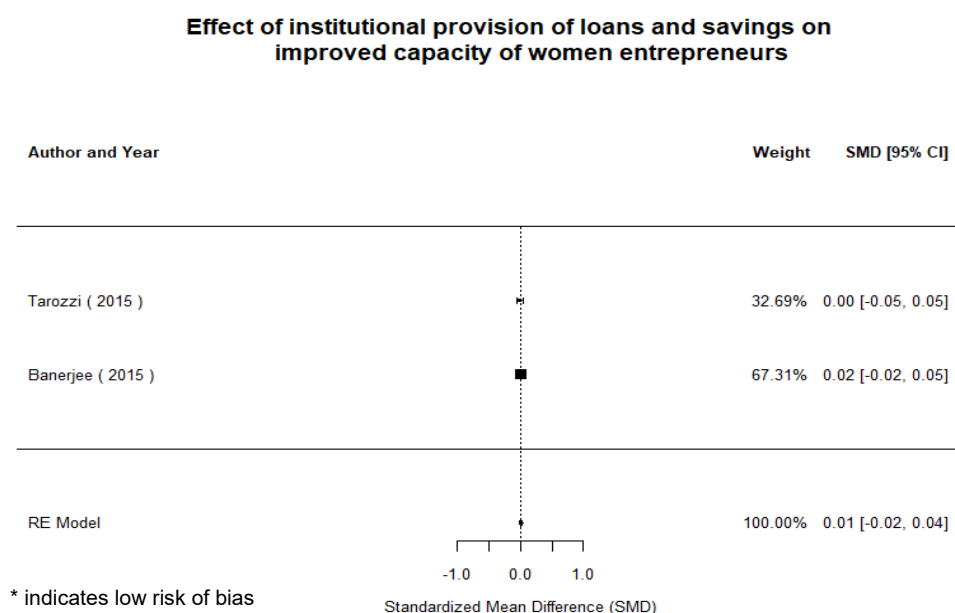
Effects of institutional provision of loans and savings on initiatives supported that facilitate women to access decent work (formal and informal employment)

Tarozzi's (2013) experimental study in Ethiopia was the only study evaluating the impact of institutional provision of loans and savings on initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities. Their report included two effects that fell into this outcome category (hours of work per week in self-employment and outside employment). The point estimates were both very small and not statistically significant, negative for self-employment ($g = -0.02$, [95% CI: -0.07 to 0.02]) and positive for outside employment ($g = 0.02$, [95% CI: -0.02 to 0.06]). We assessed the study as having some risk of bias concerns.

Effects of institutional provision of loans and savings on improved capacity of women entrepreneurs

For capacity of women entrepreneurs, two studies reported disaggregated data, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as some concerns of risk of bias and the other as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.01$ (95% CI: -0.02 to 0.04). Therefore, the average outcome did not differ significantly from zero ($z = 0.77$, $p = 0.44$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 67 LSAC3. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.29$, $p = 0.59$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator analyses were not possible, and tests of publication bias are not valid.

Figure 67: LSAC3: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of institutional provision of loans and savings on women's freedom of movement and association

Weber's (2014) quasi-experimental study in Pakistan was the only study evaluating the impact of institutional provision of loans and savings on women having increased freedom of movement and association. There was a large, and statistically significant, effect ($g = 0.53$, [95% CI: 0.09 to 0.96]), however we assessed the study as having high risk of bias.

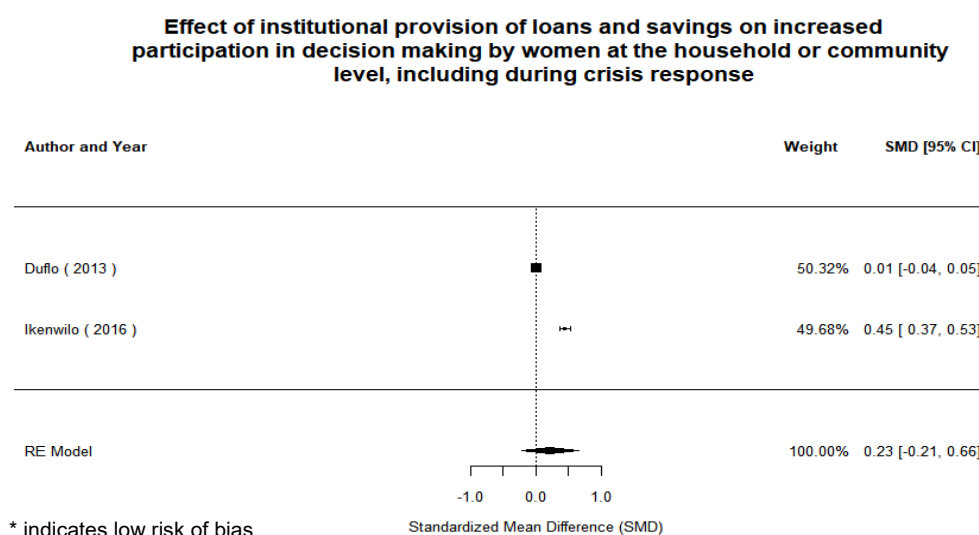
Effects of institutional provision of loans and savings on women's attitude towards taking action to claim their rights

Johnson's (2018) quasi-experimental study in Liberia was the only study evaluating the impact of institutional provision of loans and savings on women having more positive attitude towards taking action to claim their rights. Their report included effects for both individual and group lending technologies. The point estimates were medium, negative but not statistically significant, for individual lending technology ($g = -0.35$, [95% CI: -0.85 to 0.16]), and for group lending technology ($g = -0.32$, [95% CI: -0.72 to 0.09]), and we assessed the study as having some risk of bias concerns.

Effects of institutional provision of loans and savings on increased participation in decision making by women at the household or community level, including during crisis response

Two studies reported disaggregated data for increased participation in decision making by women at the household or community level, including during crisis response, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as some concerns of risk of bias and the other as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.23$ (95% CI: -0.21 to 0.66). Therefore, the average outcome did not differ significantly from zero ($z = 1.02$, $p = 0.31$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 68 LSBB1. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was significant amount of heterogeneity in the true outcomes ($Q(1) = 81.05$, $p < 0.0001$, $\hat{\tau}^2 = 0.10$, $I^2 = 98.77\%$). With only two studies, moderator analyses were not possible, and tests of publication bias are not valid.

Figure 68: LSBB1: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of institutional provision of loans and savings on representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration

Johnson's (2018) quasi-experimental study in Liberia was the only study evaluating the impact of institutional provision of loans and savings on representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration. Their report included effects for both individual and group lending technologies. The effect of individual lending technology intervention was large, negative and statistically significant ($g = -0.70$, [95% CI: -1.21 to -0.20]), and the point estimate of group lending technology was also negative but medium, and not statistically significant ($g = -0.36$, [95% CI: -0.77 to 0.04]). We assessed the study as having some risk of bias concerns.

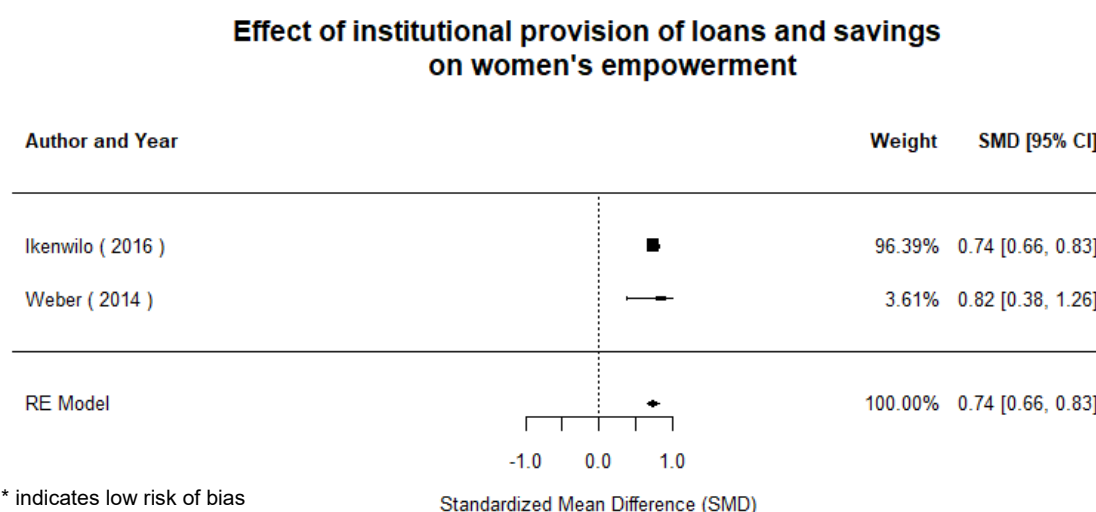
Effects of institutional provision of loans and savings on women's attitudes, self-image and confidence

Ikenwilo's (2016) quasi-experimental study in Nigeria was the only study evaluating the impact of institutional provision of loans and savings on women having improved attitudes, self-image and confidence. There was a large, and statistically significant, effect ($g = 0.58$, 95% CI [0.50, 0.67]), and we assessed the study as having some risk of bias concerns.

Effects of institutional provision of loans and savings on women's empowerment index

Two studies reported disaggregated data for women's empowerment index, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as some concerns of risk of bias and the other as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.74$ (95% CI: 0.66 to 0.83). Therefore, the average outcome differed significantly from zero ($z = 17.40$, $p < 0.001$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 64 LSCC1. According to the Q -test, the true outcomes appear to be homogeneous ($Q(1) = 0.11$, $p = 0.73$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator analyses were not possible, and tests of publication bias are not valid.

Figure 64: LSCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model



Summary of findings and discussion

We included six studies across five countries in South Asia and Sub-Saharan Africa that evaluated the effect of institutional provision of loans and savings in L&MICs. We found positive effects on women having increased access and ownership to assets, credit and income as well as on women's empowerment. However, the number of studies is small, so the results must be viewed with caution. We were also able to examine improved capacity of women entrepreneurs and increased participation in decision making by women at the household or community level, including during crisis response, but there were no significant impacts. Our included studies reported against all secondary outcomes (Resources, Agency and Achievement) and seven of the nine immediate outcomes. Overall, the GRADE assessments generally indicate a very low to low (with one moderate) certainty in this body of evidence.

We identified no linked qualitative studies and were unable to conduct a qualitative evidence synthesis related to savings and loans. The summary of quantitative evidence along with the GRADE certainty ratings are presented in Table 30.

Table 30: GRADE summary of findings and certainty of evidence on institutional provision of loans and savings

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(AB1) Increased capacity of women to understand and use financial, banking, and business services effectively.										
1	RCT – 1 QED - 1	serious ^a	not serious	not serious	serious ^b	none	15250	One large, positive effect size of 0.49 with a 95% CI of 0.41 to 0.57	⊕⊕○○ LOW	Important, but not critical
(AB2) Women have increased access and ownership to assets, credit and income										
4	RCT – 1 QED - 3	very serious ^c	serious ^d	not serious	not serious	publication bias strongly suspected ^e	15314	SMD 0.21 SD higher (0.07 higher to 0.36 higher)	⊕○○○ VERY LOW	Important, but not critical
(AC1) More women engaged in micro, small and medium sized enterprises.										
1	RCT	Very serious ^w	serious ^g	not serious	not serious	none		One positive effect estimate of 0.05 with a 95% CI range of 0.02 to 0.08	⊕○○○ VERY LOW	Limited Importance
(AC3) Improved capacity of women entrepreneurs.										
2	RCT - 2	very serious ^h	not serious	not serious	not serious	none	12687	SMD 0.01 SD higher (0.02 lower to 0.04 higher)	⊕⊕○○ LOW	Important, but not critical
(BA3) Women have increased freedom of movement and association										
1	QED	very serious ⁱ	serious ^g	not serious	serious ^j	strong association	81	SMD 0.53 SD higher (0.09 higher to 0.96 higher)	⊕○○○ VERY LOW	Limited Importance
(BA5) Women have more positive attitude towards taking action to claim their rights										
1	QED	serious ^k	serious ^g	not serious	very serious ^l	none	127	Two negative and one positive effect estimate with a 95% CI range of -0.72 to 0.49	⊕○○○ VERY LOW	Limited Importance
(BB1) Increased participation in decision making by Women at the household or community level, including during crisis response										
2	RCT – 1 QED – 1	very serious ^m	very serious ⁿ	not serious	very serious ^o	none	9039	SMD 0.23 SD higher (0.21 lower to 0.66 higher)	⊕○○○ VERY LOW	Limited Importance
(CA2) Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration										
1	QED	serious ^q	serious ^g	not serious	very serious ^r	none	127	Two negative and one positive effect estimate with a 95% CI range of -0.77 to 0.49.	⊕○○○ VERY LOW	Limited Importance
(CB3) Women have improved attitudes, self-image and confidence										
1	QED	serious ^s	serious ^g	not serious	not serious	strong association	2184	SMD 0.58 SD higher (0.5 higher to 0.67 higher)	⊕⊕⊕○ MODERATE	Limited Importance
(CC1) Empowerment/Equality index										
2	QED – 2	very serious ^t	serious ^u	not serious	serious ^v	strong association	15323	SMD 0.74 SD higher (0.66 lower to 0.83 higher)	⊕○○○ VERY LOW	Important, but not critical

Note. CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

- a. Downgraded once because both studies are of some concern for bias (though neither is high risk). In particular, there is uncertainty around reporting bias and performance bias.
- b. The range of the CI's on the RE model estimate are very wide, ranging from nearly zero to a very large estimate of 0.63.
- c. Downgraded because three of four of the studies are of some concern for bias, and the other one is of high risk. Additionally, three of four of the studies are QEDs. There is uncertainty around the reporting bias ratings. When the high risk of bias study is removed, the effect estimates decreases slightly.
- d. Downgraded because the Weber study is very inconsistent with the others in terms of magnitude of effect. That said, the studies all report positive effect sizes.
- e. The regression test indicated funnel plot asymmetry ($p < 0.01$) but not the rank correlation test ($p = 0.33$).
- f. Downgraded because the reporting bias and the selection bias criteria are unclear.
- g. All single studies downgraded once for inconsistency.
- h. Downgraded twice because both studies present a high risk of bias with significant likelihood of deviations from the intended intervention.
- i. Downgraded twice due to uncertainty surrounding several criteria and a high risk of reporting bias.
- j. Downgraded due to very wide confidence intervals.
- k. Downgraded due to uncertainty with respect to selection and reporting biases.
- l. Downgraded twice because point estimates vary widely and cross both sides of the threshold, including to nearly large effect levels.
- m. Downgraded twice because: 1) Both studies present at least concern for bias, and one is high risk of bias, and 2) When the high risk of bias study is removed, we are left with only a QED, and uncertainty and risk remains.
- n. The effect estimates are very different, including CI's that do not overlap and who cross into both sides of the threshold despite one study indicating a very large effect estimate.
- o. The RE model effect estimate has very large CI's which reach far across both sides of the threshold spanning a range of -0.21 to 0.66.
- p. Downgraded due to uncertainty with respect to performance and reporting bias.
- q. Downgraded due to uncertainty with respect to selection and reporting biases.
- r. Downgraded twice because point estimates vary widely and cross both sides of the threshold, including to nearly large effect levels.
- s. Downgraded due to uncertainty with respect to reporting bias.
- t. Downgraded twice to reflect the three studies' medium and high risk of bias. Particularly there is considerable risk and uncertainty surrounding reporting bias.
- u. Downgraded because the Karlan study is very inconsistent with the others in the group, and that study provides 35.39% of the weight of the meta-analysis.
- v. Downgraded twice because the CI's are very large, crossing into both sides of the threshold despite a very large and positive effect estimate.
- w. This single study was rated as having high risk of bias related to imbalanced attrition.

4.3.6 Quotas

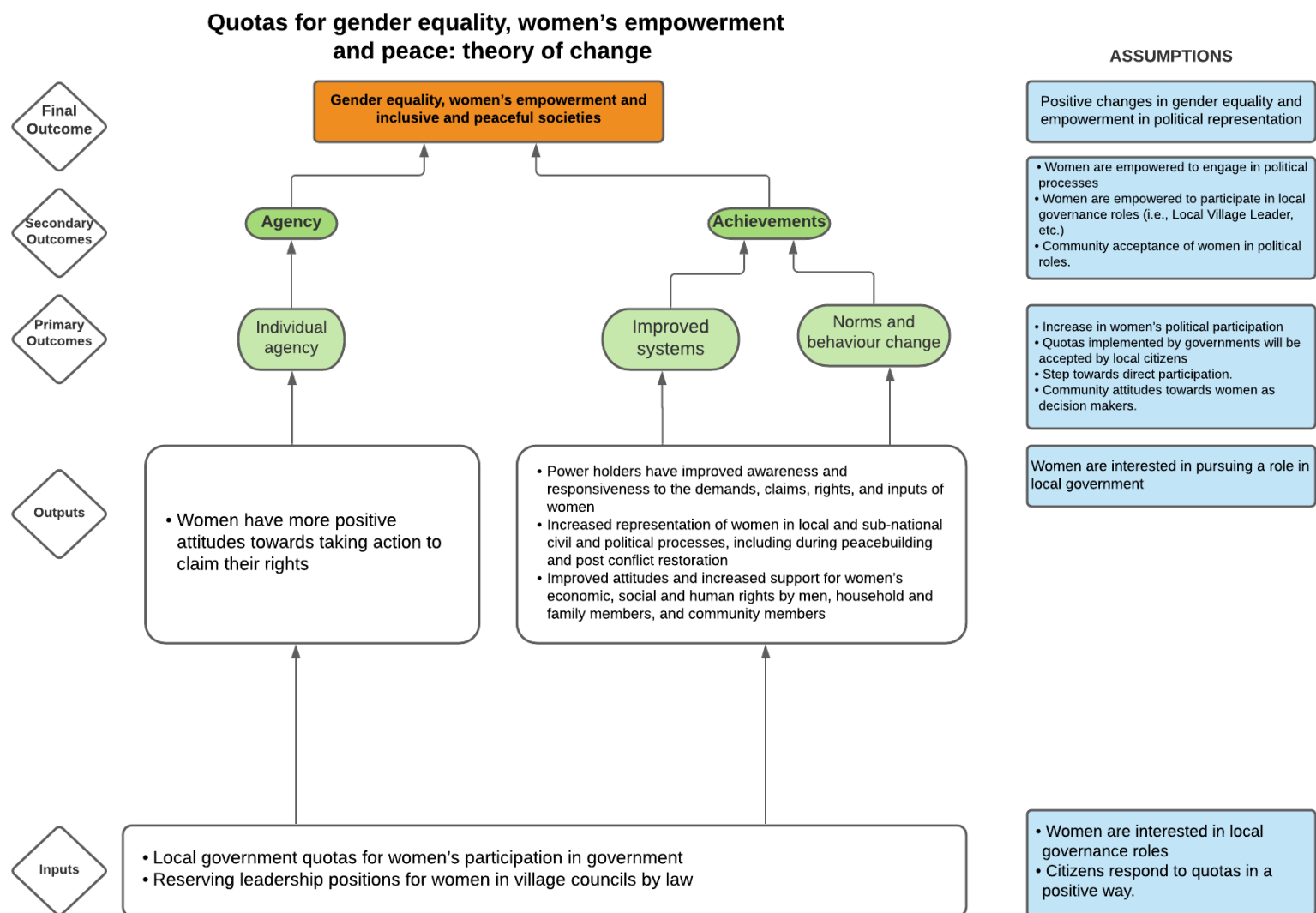
We define quotas as the allocation of positions of power on decision-making bodies in government, organisations, or community groups to a particular demographic (often oppressed or minority groups) (IDEA, 2021).

How do quotas affect gender equality, women's empowerment and peace outcomes?

Figure 69 below maps out the causal chain of how quotas may improve gender equality, women's empowerment and peace outcomes.

As illustrated in the figure below, quotas that are inclusive (such as local government quotas for women's participation in government, or legally reserving leadership positions for women in village councils) can help societies achieve high levels of agency and achievements, which are essential to the overall empowerment of women and inclusive and peaceful society. Agency and achievements rest on individual agency, improved systems, and norms and behaviour change, and are fostered by immediate outcomes that are generated from quota programmes. Immediate outcomes of such programmes include women have more positive attitudes towards taking action to claim their rights, power holders have improved awareness and responsiveness to the demands, claims, rights, and inputs of women, increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration, and improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members. Additionally, women may have different preferences for public goods than men, and thus may invest in different public goods, such as education and healthcare.

Figure 69: Quotas for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included two studies reported in two different papers that evaluated the effect of quotas.

Population

The included studies focused at the community level, with special attention given in many cases to women. The included studies evaluated programmes in both South Asia and Sub-Saharan Africa in two countries including India and Lesotho.

Intervention, inputs and activities

The included studies (Table 31) evaluated a range of different quota-related activities and inputs including:

- Reserving leadership positions for women in village councils (n = 1): A 1993 law reserved leadership positions for women to participate in local village councils.
- Local government act enacted quotas for women's participation in local government positions (n = 1): The 2005 Local Government Elections Act required that 30 per cent of all single-member electoral divisions (EDs) (distributed evenly across the newly created councils) be reserved for only women councillors. Women still competed with other women in these EDs, but men were not allowed to compete.

Table 31: Quota studies design features

Study	Activity/input	Length of treatment	Intervention frequency
Beaman (2012)	Reserving leadership positions for women in village councils	N/A	N/A
Clayton (2015)	Local Government Act enacted quotas for women's participation in local government position	77 months	N/A

Comparison

All our included studies compared treated groups to comparison groups receiving no intervention. No study included multiple treatment arms.

Outcomes

The included studies reported on a number of relevant outcomes, including the following:

- Women have a more positive attitude towards taking action to claim their rights (n = 2): Women feel entitled to be engaged and given the leadership capacity and knowledge to claim their rights and take action on relevant issues. Increased self-efficacy and autonomy. Increased opportunities for women to claim their rights including as a result of education and sensitisation. Women are engaged and given the leadership capacity and knowledge to claim their rights and take action on relevant issues.

- Power holders have improved awareness and responsiveness to the demands, claims, rights and inputs of women (n = 1): Power holders are able to recognise and support women/girls/marginalised groups through their services. They both have knowledge and understanding of these specific rights and how to act accordingly. More opportunities for dialogue and feedback mechanisms are put in place, particularly relating to gender issues. The specific needs of marginalised groups are taken into account in policies and interventions. Power holders take action towards supporting those groups.
- Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration (n = 1): Women are represented in the institutions, organisations and other decision-making processes and have capacities to influence the decision and direction of these processes. Women's specific needs and requests are taken into account and are reflected in the final decision, giving a full role to women in the civil and political processes.
- Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members (n = 1): These positive attitudes can be shaped by sensitisation and education to shift social norms, particularly relating to marginalised groups. As a result, they are aware of the specific needs of these groups and take an active part in reducing inequalities.

We report the distinction between the immediate and secondary outcomes in Table 32 below:

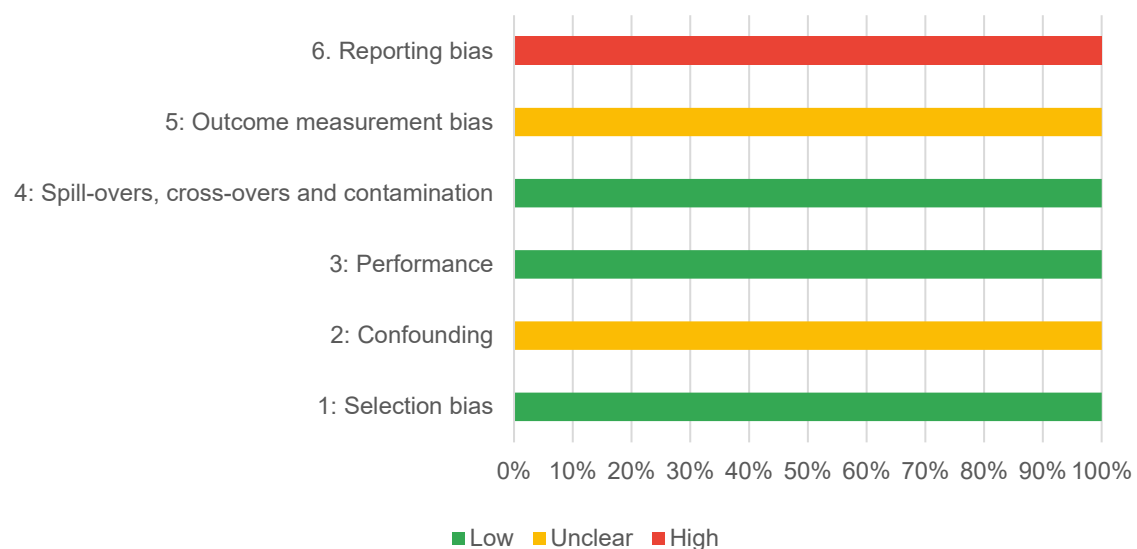
Table 32: Summary of secondary and immediate outcomes for quota studies

Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	0
	Economic and livelihood related resources	0
	Access to employment	0
Agency ability to define one's goals and act upon them and operationalised decision-making	Individual agency	2
	Community level agency	0
	Institutions supporting agency	0
Achievement ways of being and doing which can be realised by different individuals	Improved systems	1
	Norms and behaviour change	1
	Empowerment index	0

Study design

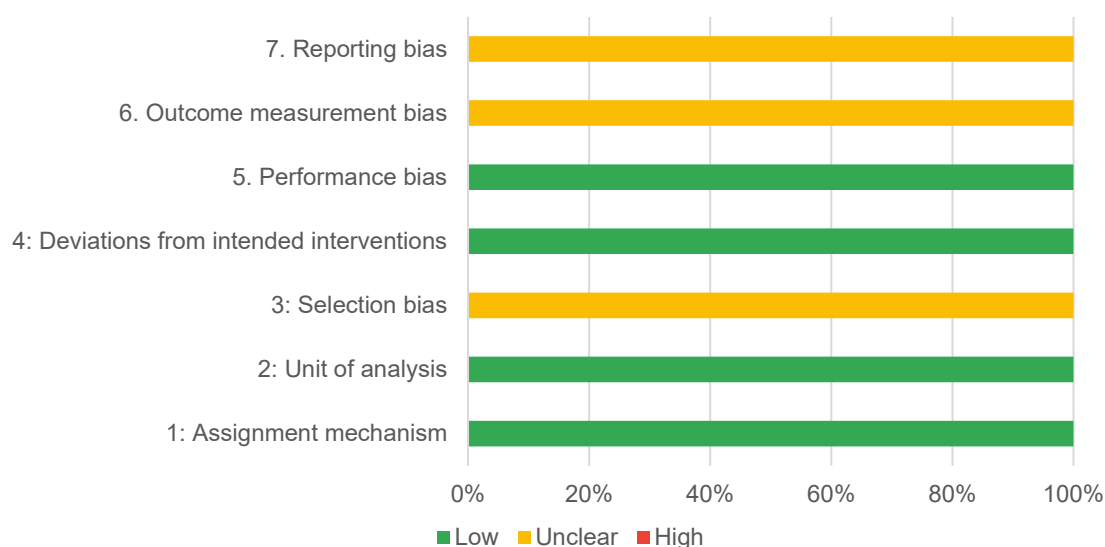
One of our quota-focused studies implemented a QED. The only study was Clayton (2015), assessed as having a high risk of bias because matching methods are not clearly reported and there was no information provided on sensitivity analysis (e.g., using different matching methods or other robustness tests). We assessed outcome measurement as unclear since authors used administrative data and did not provide enough information about data collection processes and outcome measurement methods. We also assessed confounding as unclear because the authors did not report tests for hidden bias nor results from different matching methods varying sample sizes. We did not identify any limitations related to spill-overs, crossovers, contamination, survey effects, performance or selection bias. Figure 66 below presents the RoB assessment for the QED study:

Figure 70: QED quotas RoB assessment



The other quota study used an experimental design (Beaman 2012). This study was assessed as having some concerns related to risk of bias because there was not enough information provided to assess limitations. With regards to selection bias, authors used cross-sectional data collected post-intervention, while baseline balance was tested using census data, yet it was unclear whether results may suffer from sampling issues. We also assessed as unclear whether the study suffered from outcome measurement bias because the authors did not discuss blinding procedures or potential limitations with recall or self-reported outcomes. Finally, it was also unclear whether there are reporting issues because there is not enough information to determine that there is an analysis missing. Figure 71 below presents the RoB assessment for the RCT study:

Figure 71: RCT quotas RoB assessment



Qualitative studies, process evaluations and project documents

We identified no linked qualitative studies and were unable to conduct a qualitative evidence synthesis related to quotas.

Synthesis of findings

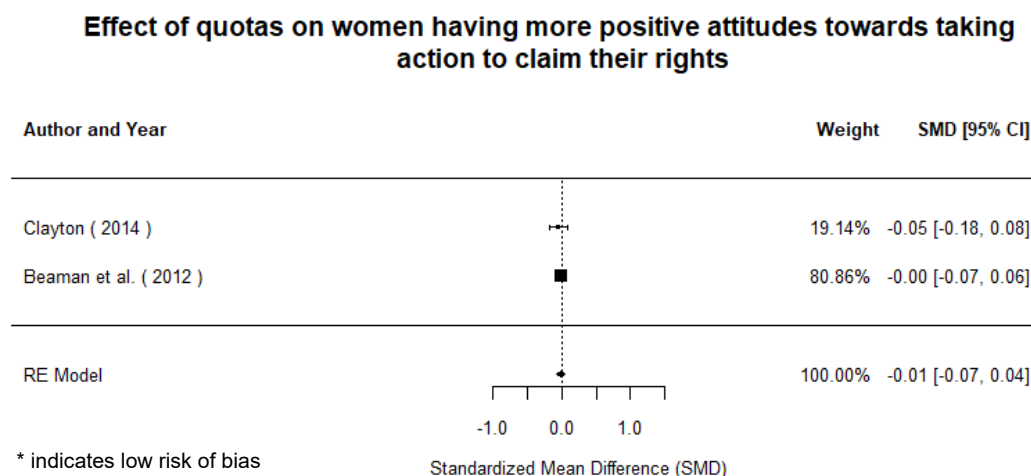
The following subsection presents the results of effectiveness of quotas on gender equality, women's empowerment and peace outcomes.

Quantitative findings

Effects of quotas on women having more positive attitudes towards taking action to claim their rights

Only two studies using a quota intervention reported outcomes related to women having more positive attitudes towards taking action to claim their rights, thus we included $k = 2$ studies in the analysis. We assessed one study as having some concerns of risk of bias and the other as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.01$ (95% CI: -0.07 to 0.04). Therefore, the average outcome did not differ significantly from zero ($z = -0.46$, $p = 0.64$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 72 QuoAttitudes. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.36$, $p = 0.55$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid.

Figure 72: QuoAttitudes. Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of quotas on reduced instances of child or forced marriage.

Beaman and colleagues' (2012) experimental study in India was the only study evaluating the impact of quota on reduced instances of child or forced marriage. There was a very small, but not statistically significant, point estimate ($g = 0.06$, [95% CI: -0.01 to 0.12]), and we assessed the study as having some risk of bias concerns.

Effects of quotas on power holders having improved awareness and responsiveness to the demands, claims, rights and inputs of women

Clayton's (2014) quasi-experimental study in Lesotho was the only study evaluating the impact of quota on power holders having improved awareness and responsiveness to the demands, claims, rights and inputs of women. Their report included two effects that fell into this outcome category (whether respondent can make councillor listen and whether her councillors try to listen). The point estimates were small, negative and not statistically significant ($g = -0.15$, [95% CI: -0.28 to -0.01] and $g = -0.13$, [95% CI: -0.27 to 0.00]). We assessed the study as having high risk of bias.

Effects of quotas on representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration

Clayton's (2014) quasi-experimental study in Lesotho was the only study evaluating the impact of quota on increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration. Their report included two effects that fell into this outcome category (interest in politics and discuss politics). The effect was negative, medium and statistically significant for interest in politics ($g = -0.26$, [95% CI: -0.40 to -0.13]) and the point estimate was very small, not significant for whether the respondent discusses politics ($g = -0.04$, [95% CI: -0.18 to 0.09]). We assessed the study as having high risk of bias.

Effects of quotas on attitudes and support for women's economic, social and human rights by men, household and family members and community members

Beaman and colleagues' (2012) experimental study in India was the only study evaluating the impact of quota on improved attitudes and increased support for women's economic, social and human rights by men, household and family members and

community members. Their report included 12 effects that fell into this outcome category (e.g., gap in parents' aspirations for girls vs. boys: the parent does not wish child to be housewife or whatever in-laws prefer). The effects ranged from very small, negative point estimates ($g = -0.03$, [95% CI: -0.08 to 0.02]) to very small, positive effects ($g = 0.09$, 95% CI [0.02 to 0.15]). We assessed the study as having some risk of bias concerns.

Summary of findings and discussion

We included two studies in two countries in South Asia and Sub-Saharan Africa that evaluated the effect of quotas. We were able to quantitatively examine effects on women having more positive attitude towards taking action to claim their rights, however, there was no significant impact. Our included studies report against two of the three secondary outcomes (Resources, Agency and Achievement) and three of the nine immediate outcomes for our review, the GRADE assessments generally indicate a low certainty in this body of evidence.

We identified no linked qualitative studies and were unable to conduct a qualitative evidence synthesis related to quotas. Table 33 presents the GRADE review of our findings:

Table 33: GRADE summary of findings and certainty of evidence on quotas

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(BA5) Women have more positive attitude towards taking action to claim their rights.										
2	RCT – 1 QED – 1	very serious ^a	not serious	not serious	not serious	none	4550	SMD 0.01 SD lower (0.07 lower to 0.04 higher)	⊕⊕○○ LOW	Limited Importance
(BA8) Reduced instances of child or forced marriage										
1	RCT	serious ^b	serious ^c	not serious	not serious	none	3680	SMD 0.06 SD higher (0.01 lower to 0.12 higher)	⊕⊕○○ LOW	Critical
(CA1) Power holders have improved awareness and responsiveness to the demands, claims, rights and inputs of women.										
1	QED	serious ^d	serious ^c	not serious	not serious	none	870	Four negative effect estimates with a 95% CI range of -0.28 to 0.07	⊕⊕○○ LOW	Limited Importance
(CA2) Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration										
1	QED	serious ^e	serious ^c	not serious	not serious	none	870	Two negative effect estimates with a 95% CI range of -0.20 to 0.08	⊕⊕○○ LOW	Limited Importance
(CB4) Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members										
1	RCT	serious ^f	serious ^c	not serious	not serious	none	3680	One positive and one negative effect estimate (both near null) with a 95% CI range of -0.07 to 0.14	⊕⊕○○ LOW	Limited Importance

Note. CI: Confidence interval; **SMD:** Standardised mean difference

Explanations:

- a. Downgraded because half of the evidence base is coming from a high risk of bias QED, and the other half from a some concerns related to risk of bias.
- b. Downgraded due to uncertainty and some risk of bias across numerous criteria.
- c. All single studies downgraded once for inconsistency.
- d. Downgraded due to bias across a number of criteria, most notably outcome measurement bias.
- e. Downgraded due to bias across a number of criteria, most notably outcome measurement bias.
- f. Downgraded due to uncertainty and some risk of bias across numerous criteria.

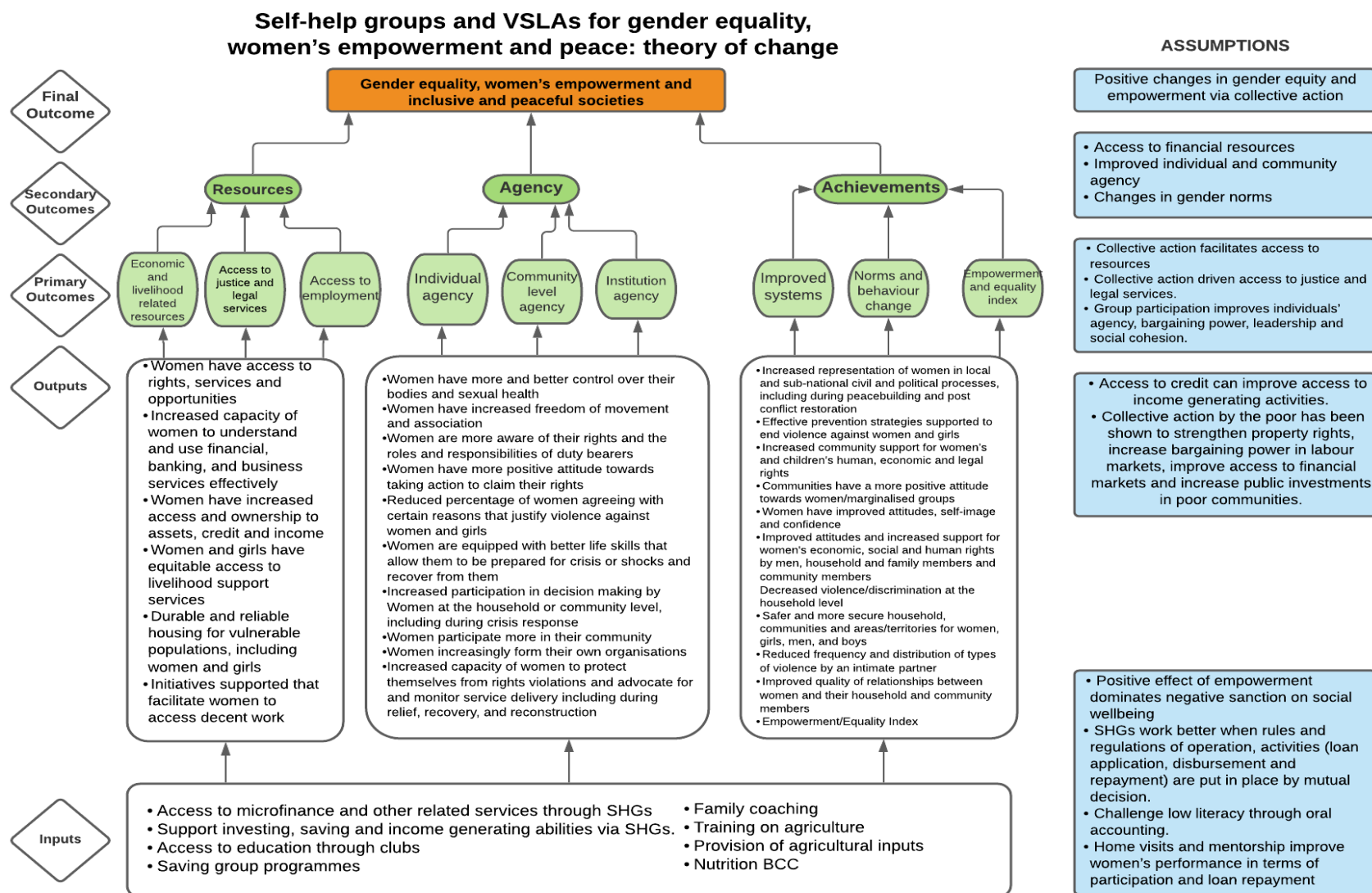
4.3.7 Self-help groups and village savings and loan associations

Self-help groups (SHGs) and village savings and loan associations (VSLAs) are bringing people together informally or through associations to work collectively towards mutual goals, often related to savings and the provision of small loans (Khasnabis et al. 2010). This category includes what is commonly referred to in existing literature as VSLAs, which is a group of people who meet regularly to save together and take small loans from those savings (VSL Associates 2021).

How do self-help groups and VSLAs affect gender equality, women's empowerment and peace outcomes?

Figure 73 below maps out the causal chain of how self-help groups and VSLAs may improve gender equality, women's empowerment and peace outcomes. SHGs and VSLA interventions particularly bring together women (and men) with common needs into self-operating groups. The theory of change implies that poor women, who lack agency and opportunity to access financial and livelihood resources, can do so via collective action. By forming groups and setting rules to govern the groups, women can collectively save and borrow money. Groups can organise and deliver skills to access livelihood activities, provide access to formal financial institutions and markets, and develop soft life-skills. It is expected that such collective actions can improve women's financial sovereignty and families' well-being, boost their self-esteem and confidence, enhance their decision-making power and widen their social network and support system. This builds on the direct benefits to credit and assets that participating in these groups often brings.

Figure 73: Self-help groups and VSLAs for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included 12 studies reported in 11 different papers that evaluated the effect of six different programmes.

Population

Self-help groups and VSLAs primarily operate at the community level. Individual women or more than one member from the same household can participate in a group. Out of the 11 papers evaluated in this category, 10 of them exclusively looked at groups for women and girls. Only one study evaluated SHGs that also allowed men to participate.

The included studies evaluated seven different programmes and trials in South Asia and Sub-Saharan Africa and seven countries: Burkina Faso, DRC, Ethiopia, Ghana, India, Malawi, Mali, Uganda and Zambia.

Intervention, inputs and activities

The included studies evaluated a range of different SHG and VSLA activities and inputs including:

- Access to microfinance for women (n = 1): Implemented by Rojiroti Microfinance, a microfinance institution which serves people who are poorer and more marginalised than those served by other microfinance programmes. The Rojiroti approach consisted of the creation of women's self-help groups, rotated loans from savings, and subsequent credit from CPSI, a Bihar based NGO.
- Formation of SHGs - once SHGs are formed, then access to credit and other services is facilitated (n = 1): Sponsored by the Indian government, the Swarnajayanti Gram Swarozgar Yojana (SGSY) was a credit scheme where self-help groups are setup for members to mobilise, save credit and facilitate loans.
- Girls' and women's clubs for supporting girls and women's empowerment through access to education (n = 1): Implemented by OXFAM, this programme consisted of setting up mothers' associations to conduct training and sensitisation activities, setting up girls' club to deliver sensitisation training, setting up training, identifying role-model women to organise exchanges on topics.
- Savings group programme for women (n = 1): Implemented by the International Rescue Committee and targeted towards women survivors of sexual violence, VSLAs are economic programmes in which individuals join with people they know to regularly save funds and build capital to generate small loans to other members.
- Self-help groups (n = 2): Implemented by the Government of Ethiopia, this programme aimed to create self-help groups for women working in apple production. SHGs are groups of 10 to 20 women initiated by a development agency, which are usually involved in savings and credit programmes and/or advancing group members' rights.
- Microfinance SHG scheme (n = 1): The Self-Employed Women's Association (SEWA) set up SHGs in the Dungarpur district of Rajasthan. These SHGs set their own savings targets for their members. The funds were put in a savings account with a linked bank and were lent out to members of the SHG.
- Multicomponent intervention involving training and group creation, with aim to increase investing, saving, income generating abilities (n = 1): Following the Gender Action Learning System (GALS) approach, the project established women's groups to promote savings and investment opportunities and be trained

on knowledge for running small enterprises. The project also supported women's groups in gaining access to loans to start new business.

- VLSA model, organising women into groups and providing seed capital for initiatives + family coaching (n = 1): The experiment consisted of 1) a package of economic and livelihood interventions for women caregivers that included formation and training using the Village Savings and Loan Association (VSLA), seed capital grants to jump-start or expand livelihood activities and one-on-one mentoring; and 2) a gender sensitive family coaching component.
- VSLAs (n = 1): Begun by CARE International and implemented by the International Rescue Committee and other INGOs, VSLAs are economic programmes in which individuals join with people they know to regularly save funds and build capital to generate small loans to other members.
- Women's groups for agricultural training, provision of initial inputs and nutrition BCC (n = 1): This Concern Worldwide-supported programme entailed group sessions on agriculture, gender equality and women's empowerment, (Ag-G group), nutrition targeted behavioural change communication interventions and provision of standard government services such as maize-focused agricultural extension targeted predominantly to male farmers, and antenatal care visits and growth monitoring for under five-year-old children.

Table 34: Self-help groups and VSLAs design features of included studies

Study	Activity/input	Length of treatment	Intervention frequency
Alemu - 2018	Self-help groups	36 months	Unclear
Bass - 2016	VSLAs	12 months	Weekly
Beaman - 2014	Saving group programme for women	36 months	Weekly
De Hoop - 2010	Self-help groups	Unclear	Unclear
De Hoop – 2014	Self-help groups	36 months	Unclear
Deininger – 2013	Self-help groups	Unclear	Weekly
Desai - 2013	Microfinance self-help group scheme through SEWA	24 months	Monthly
Ismayilova - 2018	VLSA model, organising women into groups and providing seed capital for initiatives + family coaching	24 months	Weekly
Karlan - 2017	VSLA model with agents whose objective was to informally expand the VSLA model to other communities	30 months in Ghana. 22 months in Malawi and Uganda	Weekly

Study	Activity/input	Length of treatment	Intervention frequency
Kundu - 2011	Formation of SHGs and access to credit and other services was facilitated	48 months	Unclear
Lombardini - 2015	Multicomponent intervention involving training and group creation, with aim to increase investing, saving, income generating abilities	33 months	Monthly
Vigneri - 2017	Girls' and women's clubs for supporting girls and women's empowerment through access to education	50 months	Weekly
Yaron - 2018	Access to microfinance for women	18 months	Unclear

Comparison

All our included studies are comparing treated groups to comparison groups receiving no intervention. Seven studies included multiple treatment arms.

Outcomes

The included studies reported on a number of relevant outcomes, including:

- Women have increased access and ownership to assets, credit and income (n = 8): Women are able to apply for, receive and manage assets/credit and income and have support to manage, claim and execute their assets without pressure or influence from external actors, including male family members, husbands and cultural leaders.
- Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them (n = 4): Women develop skills, have access to resources, and make use of these skills and resources to increase their level of resilience and be equipped to face life challenges (health, education, finance, social relations, work etc.).
- Increased participation in decision making by women at the household or community level, including during crisis response (n = 7): Women take part in all or any step of the decision-making process at the household, community and district level, but also are able to meaningfully take part and have influence on the final decision, including in crisis response.
- Decreased violence and discrimination at the household level (n = 3): The occurrence of these acts decreases over time as a result of interventions focusing on prevention and response.

The division of the immediate and secondary outcomes is reported in Table 35 below:

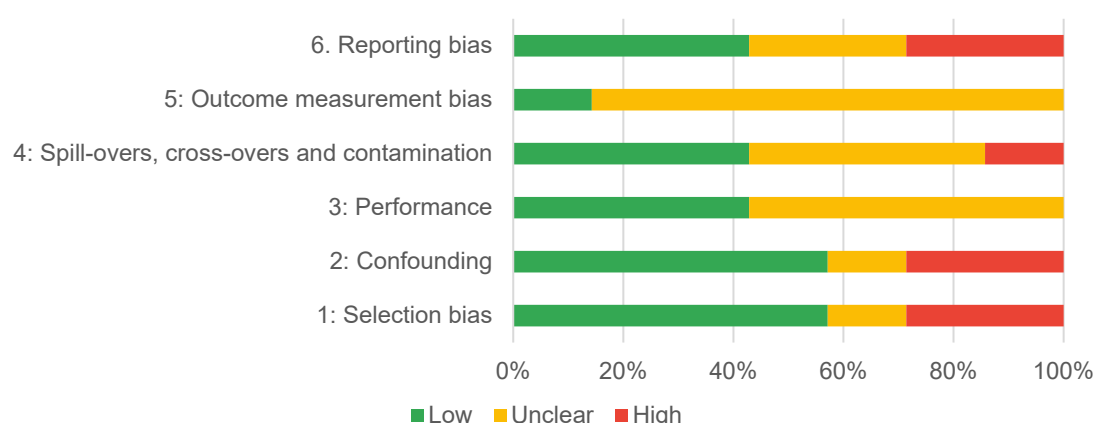
Table 35: Self-help groups and VSLAs summary of secondary and immediate outcomes

Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	1
	Economic and livelihood related resources	9
	Access to employment	2
Agency ability to define one's goals and act upon them and operationalised decision-making;	Individual agency	9
	Community level agency	9
	Institutions supporting agency	1
Achievement ways of being and doing which can be realised by different individuals	Improved systems	4
	Norms and behaviour change	8
	Empowerment index	4

Study design

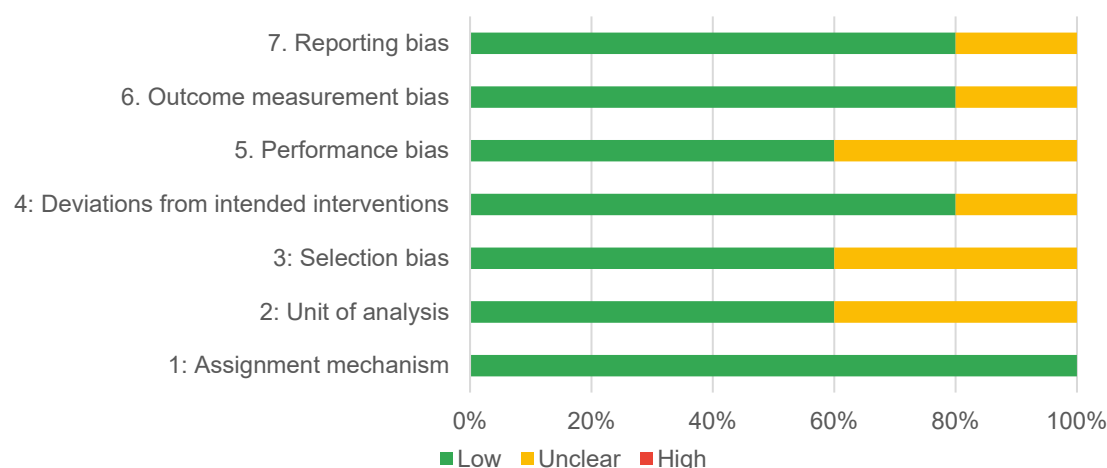
Seven of our SHG and VSLA studies used a QED. This included Alemu (2018), De Hoop (2010), Kundu (2011), Lombardini (2015), Vigneri (2017), Yaron (2018). We assessed half of the QED studies evaluation SHG and VSLA as having high risk of bias and the other half as having some concerns, none was assessed as having low risk or bias. As detailed in Figure 74 below, high risk of bias was identified in four RoB domains (selection bias, confounding bias, spill-overs and reporting bias). Potential issues with selection bias were observed when it was not clear that all relevant characteristics had been accounted for (Vigneri 2017; Kundu 2011). Confounding was a risk when authors did not control for relevant characteristics that may be correlated with both participation and outcomes of interest (Ibid.). Issues with reporting bias included a lack of robustness checks (Lombardini 2015) and uncommon methods used for the analysis such as no adjustment for covariates on a matched sample (Vigneri 2017). Finally, the presence of similar initiatives in the intervention areas was considered a risk of contamination (Ibid.). Outcome measurement bias was assessed as unclear for all but one of the studies because the authors had not provided enough information or discussion about recall or self-reported bias.

Figure 74: SHGs and VSLA studies QED RoB assessment



Five SHG and VSLA studies used an experimental design. This included Bass (2016), Beaman (2014), Desai (2013), Ismayilova (2018), Karlan (2017) and Kumar (2018). All six studies had some concerns related to risk of bias. As detailed in Figure 75 below, high risk of bias was not identified in any RoB domains of the sample of included RCT studies. No major concerns were observed, however one or two studies in each category did not provide enough information to discard issues related to deviations from intended interventions, selection bias, performance bias, reporting bias, assignment mechanism or unit of analysis. The only domain for which all studies were assessed as having low risk of bias was the outcome measurement one.

Figure 75: SHGs and VSLA studies RCT RoB assessment



Qualitative studies, process evaluations and project documents

We identified 10 additional documents related to five programmes covered by the SHG and VSLA group of studies.

- Saving for Change (SfC) (Mali): two qualitative documents, two process evaluations and two descriptive quantitative documents
- SEWA programme (India): one qualitative document
- Village Savings and Loans Associations (VSLAs) (DRC): two descriptive quantitative documents
- Girls Can: Promoting secondary education for girls in West Africa (Mali): one descriptive quantitative document

The highest number of studies were ranked as high quality ($n = 5$), followed by low quality ($n = 4$). Only one study from this intervention group was marked as moderate quality.

Synthesis of findings

Quantitative findings

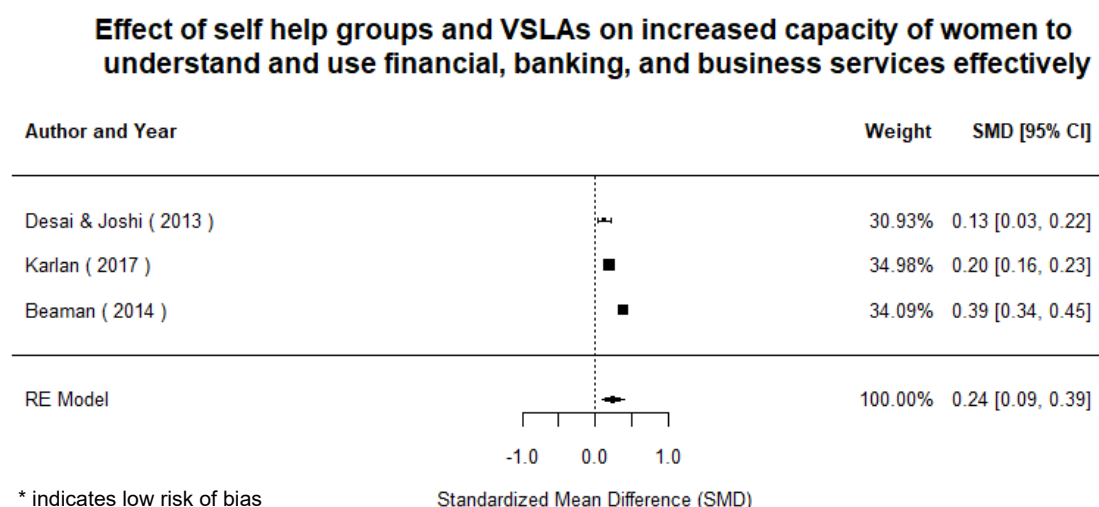
Effect of SHGs on women's access to rights, services and opportunities

Lombardini and Yoshikawa's (2015) quasi-experimental study in Uganda was the only study evaluating the impact of self-help groups and VSLA on women have access to rights, services and opportunities. There was a small, and statistically significant, effect ($g = 0.18$, 95% CI [0.03, 0.32]), however we assessed the study as having high risk of bias.

Effect of SHGs on increased capacity of women to understand and use financial, banking and business services effectively

We included a total of $k = 3$ studies in the analysis. We assessed all three studies as having some concerns of risk of bias. The observed outcomes ranged from 0.13 to 0.39. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.24$ (95% CI: 0.09 to 0.39). Therefore, the average outcome differed significantly from zero ($z = 3.15$, $p < 0.01$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 76 SHGAB1. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, the true outcomes appear to be heterogeneous ($Q(2) = 42.94$, $p < 0.01$, $\hat{\tau}^2 = 0.02$, $I^2 = 95.34\%$). An examination of the studentized residuals revealed that one study (Beaman 2014) had a value larger than ± 2.39 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. Indeed, sensitivity analyses leaving each study out indicated that removing Beaman (2014) would reduce the overall average effect ($\mu^{\wedge} = 0.18$ (95% CI: 0.12 to 0.24), but the effect would still be positive and non-significant ($z = 5.61$, $p < .001$). With only three studies, moderator analyses were not appropriate.

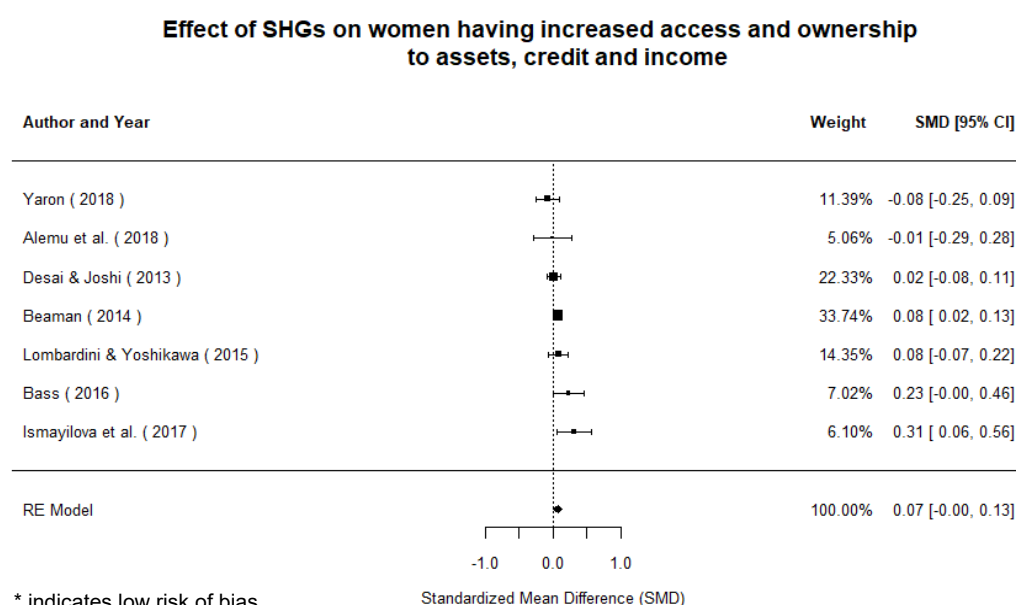
Figure 76: SHGAB1: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effect of SHGs on women having increased access and ownership to assets, credit and income

We included a total of $k = 7$ studies in the analysis. We assessed none of the studies as low risk of bias, six as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.08 to 0.31 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.07$ (95% CI: -0.00 to 0.13). Therefore, the average outcome did not differ significantly from zero ($z = 1.90$, $p = 0.06$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 77 SHGAB2.

Figure 77: SHGAB2: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(6) = 9.69$, $p = 0.14$, $\hat{\tau}^2 = 0.00$, $I^2 = 38.08\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.69 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. We tested all moderators, but there were no significant sources of heterogeneity for this outcome.

Effect of SHGs on women and girls' access to livelihood support services

Alemu and colleagues' (2018) quasi-experimental study in Ethiopia was the only study evaluating the impact of self-help groups and VSLA on women and girls having equitable access to livelihood support services. There was a medium, and statistically significant, effect ($g = 0.31$, [95% CI: 0.03 to 0.60]), and we assessed the study as having some risk of bias concerns.

Effect of SHGs on durable and reliable housing for vulnerable populations, including women and girls

Yaron's (2018) quasi-experimental study in India was the only study evaluating the impact of self-help groups and VSLA on durable and reliable housing for vulnerable populations, including women and girls. There was a large, and statistically significant,

effect ($g = 0.40$, [95% CI: 0.22 to 0.57]), and we assessed the study as having some risk of bias concerns.

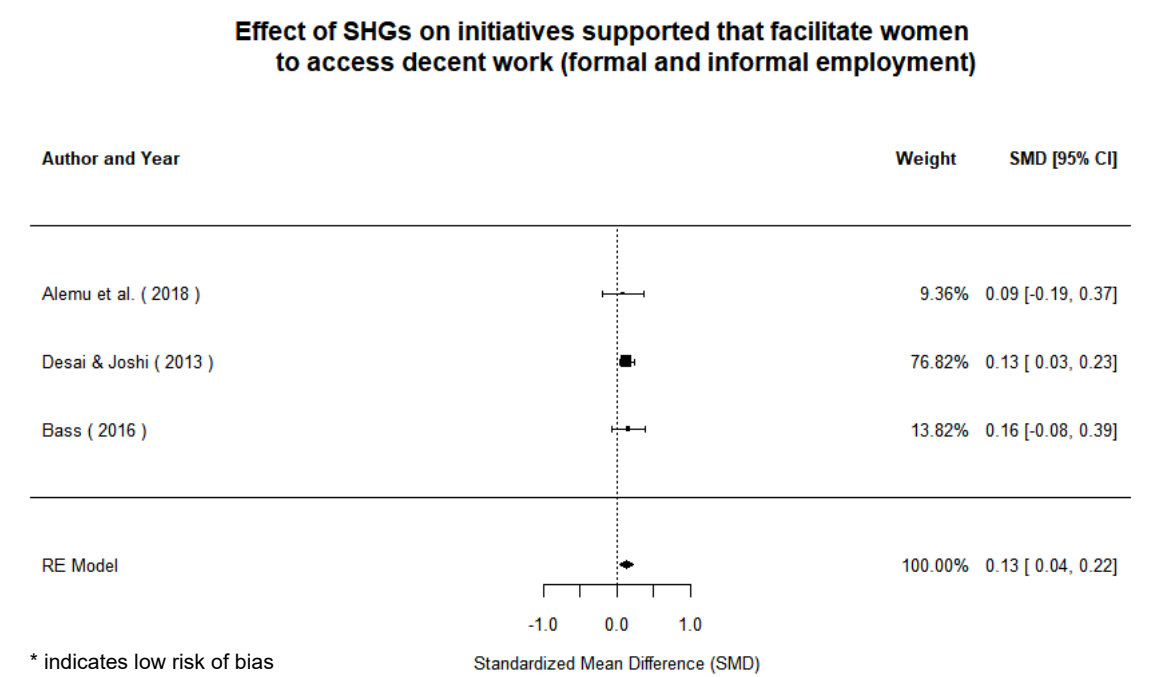
Effect of SHGs on women’s engagement in other micro, small, and medium-sized enterprises

Beaman's (2014) experimental study in Mali was the only study evaluating the impact of self-help groups and VSLA on more women being engaged in other micro, small, and medium-sized enterprises. There was a very small, but not statistically significant, point estimate ($g = 0.03$, [95% CI: -0.01 to 0.07]), and we assessed the study as having some risk of bias concerns.

Effect of SHGs on initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities

We included a total of $k = 3$ studies in the analysis. We assessed all three studies as having some concerns of risk of bias. The observed outcomes ranged from 0.09 to 0.16. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.13$ (95% CI: 0.04 to 0.22). Therefore, the average outcome differed significantly from zero ($z = 2.97$, $p < 0.01$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 78 SHGAC2.

Figure 78: SHGAC2: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(2) = 0.12$, $p = 0.94$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). An examination of the [studentised](#) residuals revealed that none of the studies had a value larger than ± 2.39 and hence there was no indication of outliers in the context of this model. According to the Cook’s distances, none of the studies could be considered to be overly influential. With only three studies and no heterogeneity among the effects, moderator analyses were not possible.

Effect of SHGs on improved capacity of women entrepreneurs

Lombardini and Yoshikawa's (2015) quasi-experimental study in Uganda was the only study evaluating the impact of self-help groups and VSLA on improved capacity of women entrepreneurs. There was a small, and statistically significant, effect ($g = 0.19$, [95% CI: 0.05 to 0.34]), and we assessed the study as having high risk of bias.

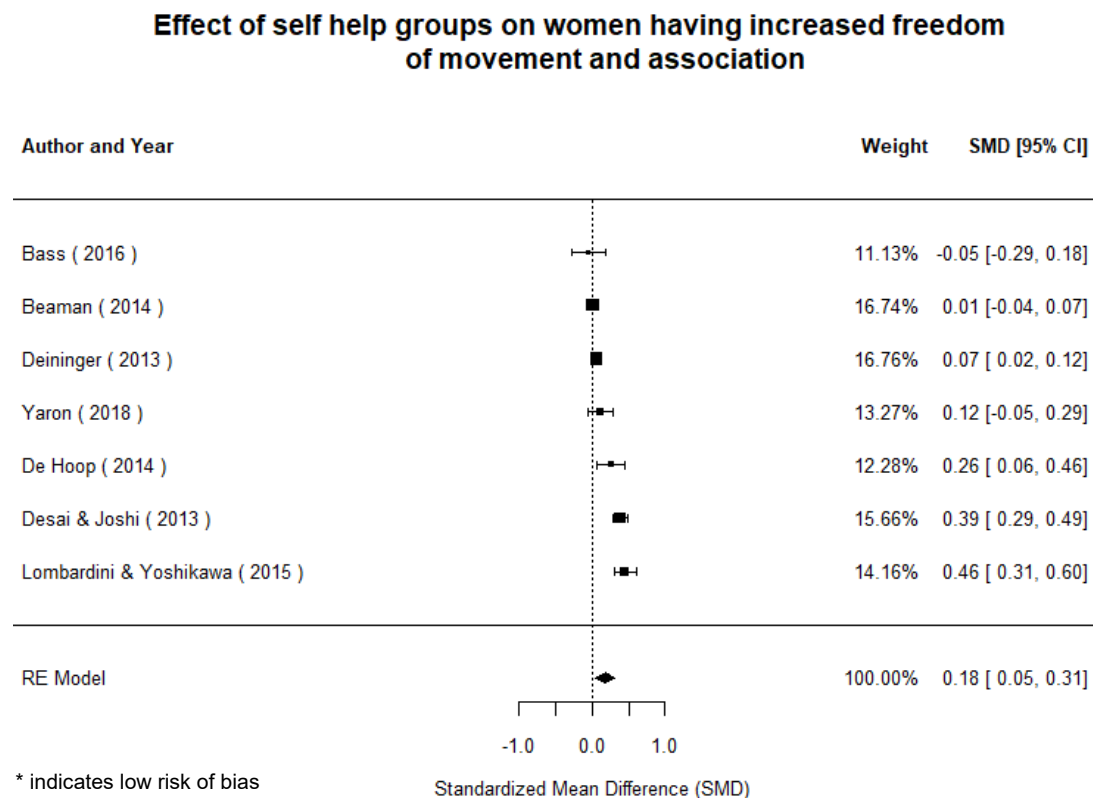
Effect of SHGs on women having more and better control over their bodies and sexual health

Lombardini and Yoshikawa's (2015) quasi-experimental study in Uganda was the only study evaluating the impact of self-help groups and VSLA on women having more and better control over their bodies and sexual health. There was a very small, but not statistically significant, point estimate ($g = -0.01$, [95% CI: -0.16 to 0.13]), and we assessed the study as having high risk of bias.

Effect of SHGs on women having increased freedom of movement and association

We included a total of $k = 7$ studies in the analysis. We assessed none of the studies as low risk of bias, six as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.05 to 0.46 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.18$ (95% CI: 0.05 to 0.31). Therefore, the average outcome did not differ significantly from zero ($z = 2.72$, $p < 0.01$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 79 SHGBA3.

Figure 79: SHGBA3: Forest plot showing the observed outcomes and the estimate of the random-effects model



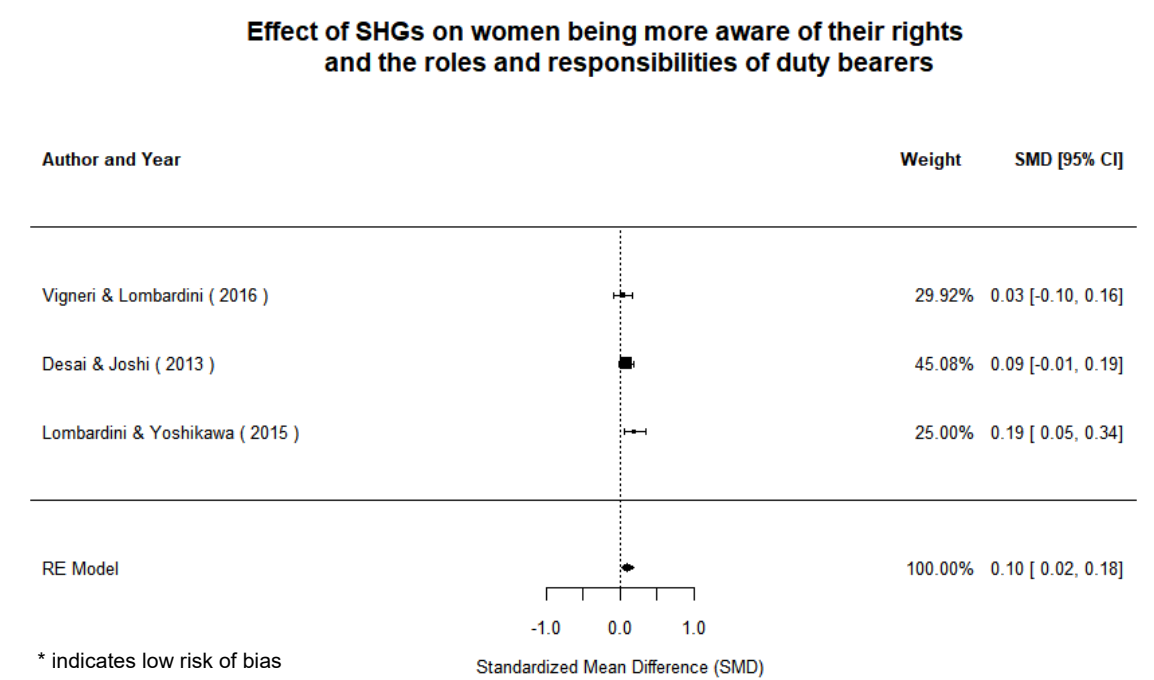
According to the Q -test, the true outcomes appear to be heterogeneous ($Q(6) = 71.20$, $p < 0.01$, $\hat{\tau}^2 = 0.03$, $I^2 = 91.57\%$). An examination of the studentised residuals revealed

that none of the studies had a value larger than ± 2.69 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. We found no significant sources of heterogeneity through our moderator analyses.

Effect of SHGs on women being more aware of their rights and the roles and responsibilities of duty bearers

We included a total of $k = 3$ studies in the analysis. We assessed none of the studies as low risk of bias, one as having some concerns, and two as low risk of bias. The observed outcomes ranged from 0.03 to 0.19. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.10$ (95% CI: 0.02 to 0.18). Therefore, the average outcome differed significantly from zero ($z = 2.39, p = 0.02$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 80 SHGBA4.

Figure 80: SHGBA4: Forest plot showing the observed outcomes and the estimate of the random-effects model



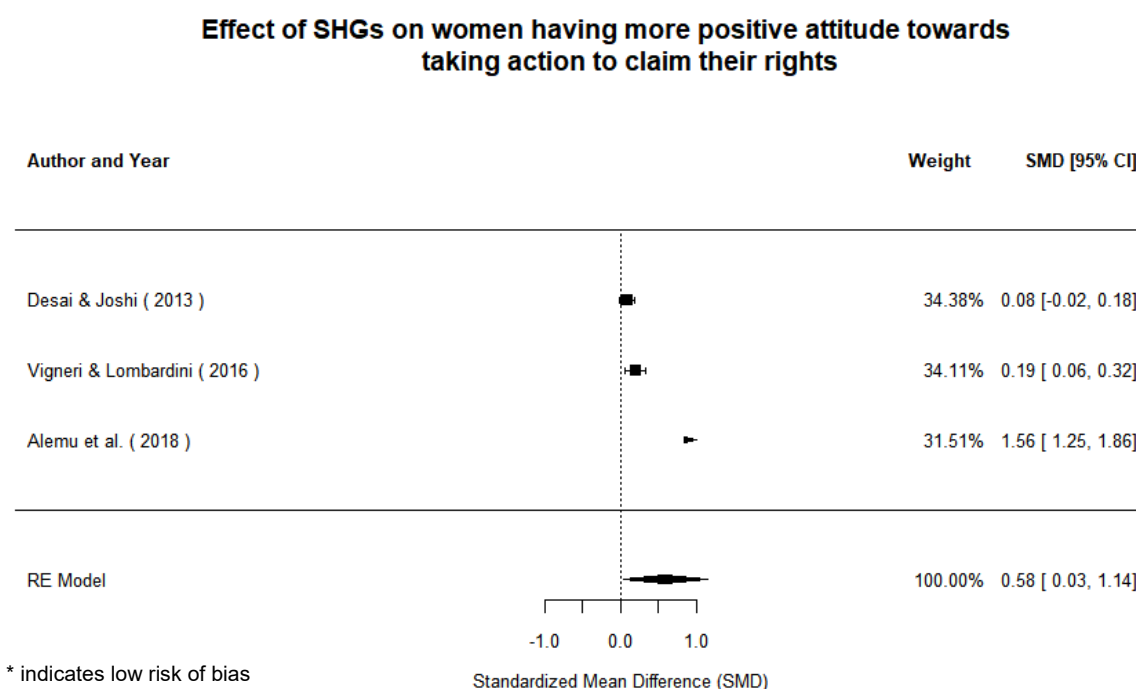
According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(2) = 2.59, p = 0.27, \hat{\tau}^2 = 0.00, I^2 = 22.64\%$). An examination of the [studentised](#) residuals revealed that none of the studies had a value larger than ± 2.39 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

Effect of SHGs on women having more positive attitude towards taking action to claim their rights

We included a total of $k = 3$ studies in the analysis. We assessed none of the studies as low risk of bias, two as having some concerns, and one as low risk of bias. The observed outcomes ranged from 0.08 to 1.56. The estimated average outcome based on the

random-effects model was $\hat{\mu} = 0.58$ (95% CI: 0.03 to 1.14). Therefore, the average outcome differed significantly from zero ($z = 2.06$, $p = 0.04$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 81 SHGBA5.

Figure 81: SHGBA5: Forest plot showing the observed outcomes and the estimate of the random-effects model

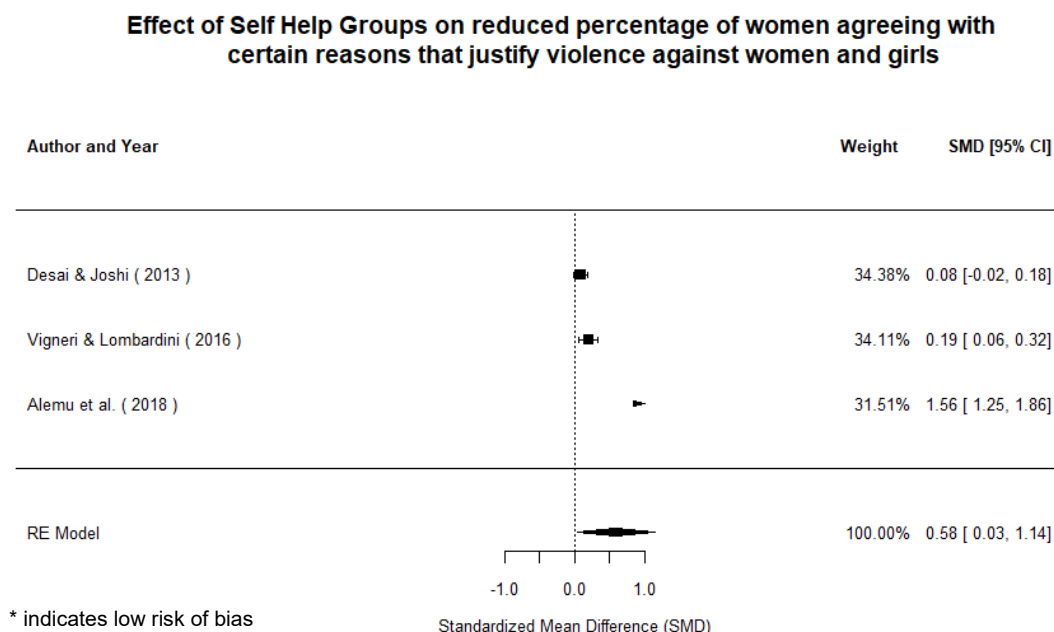


According to the Q -test, the true outcomes appear to be heterogeneous ($Q(2) = 82.66$, $p < 0.01$, $\hat{\tau}^2 = 0.23$, $I^2 = 97.58\%$). An examination of the [studentised](#) residuals revealed that one study (Alemu et al. 2018) had a value larger than ± 2.39 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. Sensitivity analyses leaving each study out indicated that removing Alemu and colleagues (2018) would reduce the overall average effect ($\hat{\mu} = 0.13$; 95% CI: 0.02 to 0.24), but the effect would still be positive and significant ($z = 2.31$, $p = 0.02$). With only three studies, moderator analyses were not appropriate.

Effect of SHGs on reduced percentage of women agreeing with certain reasons that justify violence against women and girls

We included a total of $k = 3$ studies in the analysis. We assessed none of the studies as low risk of bias, one as having some concerns, and two as low risk of bias. The observed outcomes ranged from 0.08 to 1.56. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.58$ (95% CI: 0.03 to 1.14). Therefore, the average outcome differed significantly from zero ($z = 2.06$, $p = 0.04$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 82 SHGBA6.

Figure 82: SHGBA6: Forest plot showing the observed outcomes and the estimate of the random-effects model

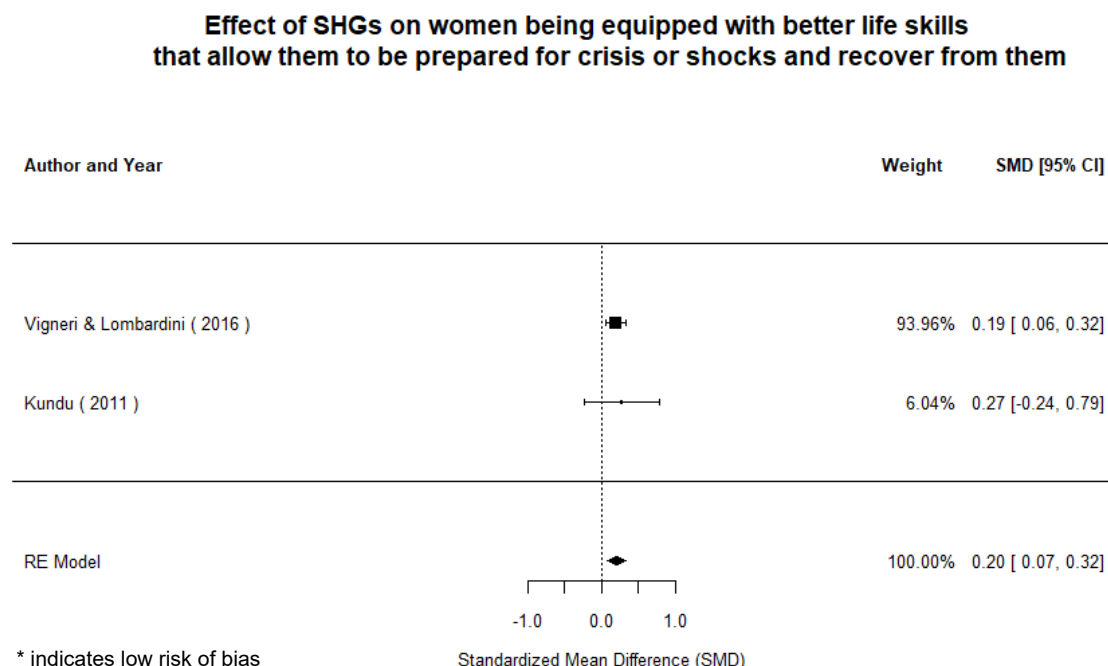


According to the Q -test, the true outcomes appear to be heterogeneous ($Q(2) = 82.66$, $p < 0.01$, $\hat{\tau}^2 = 0.23$, $I^2 = 97.58\%$). An examination of the [studentised](#) residuals revealed that one study (Alemu et al. 2018) had a value larger than ± 2.39 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. sensitivity analyses leaving each study out indicated that removing Alemu (2018) would reduce the overall average effect ($\hat{\mu} = 0.13$ [95% CI: 0.02 to 0.24]), but the effect would still be positive and significant ($z = 2.31$, $p = 0.02$). With only three studies, moderator analyses were not appropriate.

Effect of SHGs on women being equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them

Two studies reported disaggregated data for women being equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them, thus we included $k = 2$ studies in the analysis. We assessed both of the studies as low risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.1987$ (95% CI: 0.07 to 0.33). Therefore, the average outcome differed significantly from zero ($z = 3.0862$, $p = 0.002$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 83 SHG BA7. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.0893$, $p = 0.77$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, and given there was no heterogeneity among the effects, moderator analyses were not possible, and tests of publication bias are not valid.

Figure 83: SHGBA7: Forest plot showing the observed outcomes and the estimate of the random-effects model



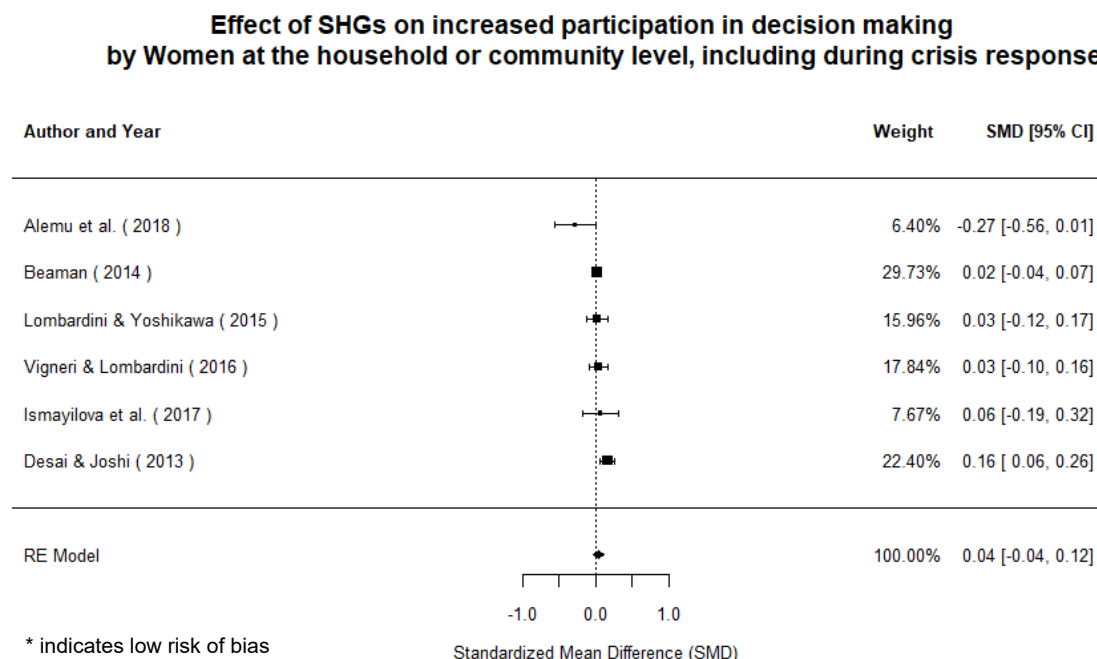
Effect of SHGs on reduced instances of child or forced marriage

Vigneri and Lombardini's (2016) quasi-experimental study in Mali was the only study evaluating the impact of self-help groups and VSLA on reduced instances of child or forced marriage. There was a very small, and not statistically significant, point estimate ($g = 0.03$, [95% CI: -0.10 to 0.16]), and we assessed the study as having high risk of bias.

Effect of SHGs on increased participation in decision making by women at the household or community level, including during crisis response

We included a total of $k = 6$ studies in the analysis. We assessed none of the studies as low risk of bias, four as having some concerns, and two as low risk of bias. The observed outcomes ranged from -0.27 to 0.16 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.04$ (95% CI: -0.04 to 0.12). Therefore, the average outcome did not differ significantly from zero ($z = 0.98$, $p = 0.33$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 84 SHGBB1.

Figure 84: SHGBB1: Forest plot showing the observed outcomes and the estimate of the random-effects model

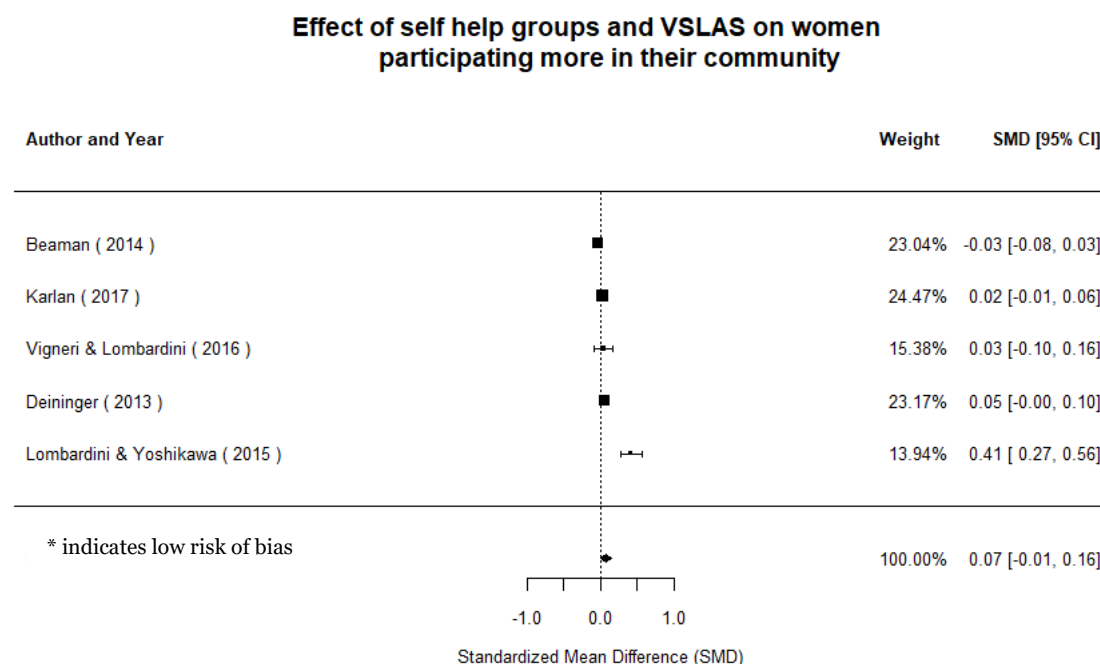


According to the Q -test, the true outcomes appear to be heterogeneous ($Q(5) = 11.43$, $p = 0.04$, $\hat{\tau}^2 = 0.00$, $I^2 = 56.24\%$). An examination of the [studentised](#) residuals revealed that one study (Desai and Joshi 2013) had a value larger than ± 2.64 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. Sensitivity analyses leaving each study out indicated that removing Desai and Joshi (2013) would reduce the overall average effect ($\hat{\mu} = 0.01$ [95% CI: -0.04 to 0.07]), but the effect would still be insignificant ($z = 0.51$, $p = 0.61$). No sources of heterogeneity were found to be significant through our moderator analyses.

Effect of SHGs on women participating more in their community

We included a total of $k = 5$ studies in the analysis. We assessed none of the studies as low risk of bias, three as having some concerns, and two as low risk of bias. The observed outcomes ranged from -0.03 to 0.41 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.07$ (95% CI: -0.01 to 0.16). Therefore, the average outcome did not differ significantly from zero ($z = 1.81$, $p = 0.07$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 85 SHGBB2.

Figure 85: SHGBB2: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(4) = 31.85$, $p < 0.01$, $\hat{\tau}^2 = 0.01$, $I^2 = 87.44\%$). An examination of the [studentised](#) residuals revealed that one study (Lombardini and Yoshikawa 2015) had a value larger than ± 2.58 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. Sensitivity analyses leaving each study out indicated that removing Lombardini and Yoshikawa (2015) would reduce the average effect ($\hat{\mu} = 0.0187$ [95% CI: -0.01 to 0.05]), but the effect would still be insignificant ($z = 1.12$, $p = 0.26$).

We were able to test several moderators in the context of this model. Studies with a high risk of bias were significantly different than studies with some concerns related to risk of bias ($\hat{B} = 0.20$, $p = 0.01$ [95% CI: 0.04 to 0.35]), such that high risk of bias studies had higher effects than studies with some concerns by .20 standard deviation units. Exposure to intervention, publication year and post-intervention versus change from baseline were not significant moderators. There was not sufficient data to test evaluation period or whether the model was adjusted for covariates.

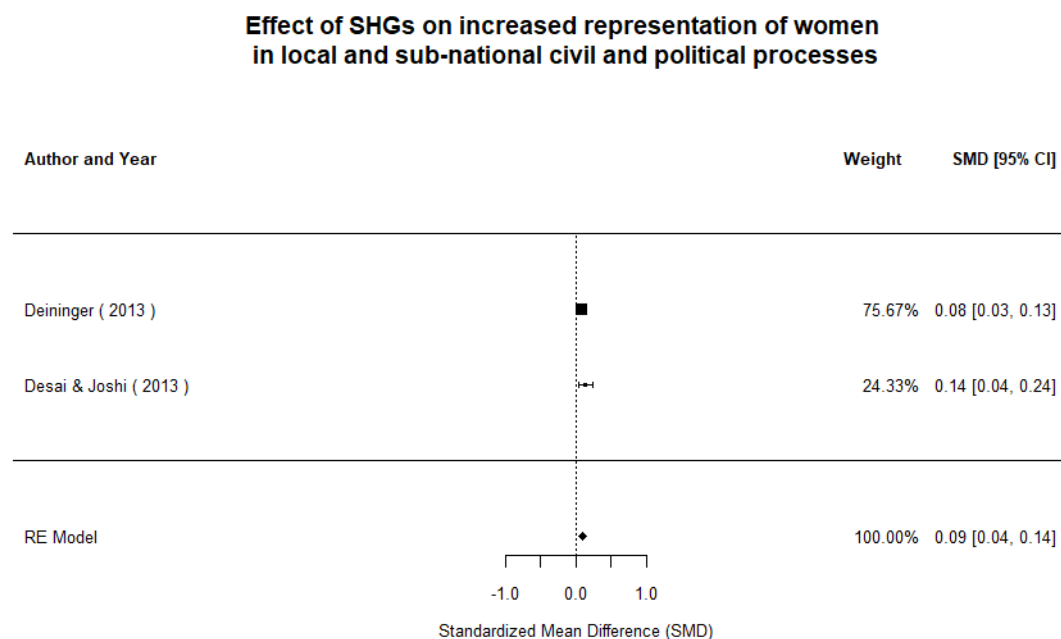
Effect of SHGs on power holders having improved awareness and responsiveness to the demands, claims, rights and inputs of women

Lombardini and Yoshikawa's (2015) quasi-experimental study in Uganda was the only study evaluating the impact of self-help groups and VSLA on power holders having improved awareness and responsiveness to the demands, claims, rights and inputs of women. There was a very small, and not statistically significant, point estimate ($g = 0.09$, [95% CI: -0.06 , 0.23]), and we assessed the study as having high risk of bias.

Effect of SHGs on representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration

Two studies reported disaggregated data for representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration, thus we included $k = 2$ studies in the analysis. We assessed both studies as having some concerns of risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.09$ (95% CI: 0.04 to 0.14). Therefore, the average outcome differed significantly from zero ($z = 3.5987$, $p = 0.0003$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 86 SHGCA2. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 1.1089$, $p = 0.2923$, $\hat{\tau}^2 = 0.00$, $I^2 = 9.82\%$). With only two studies, and given there was no heterogeneity among the effects, moderator analyses were not possible, and tests of publication bias are not valid.

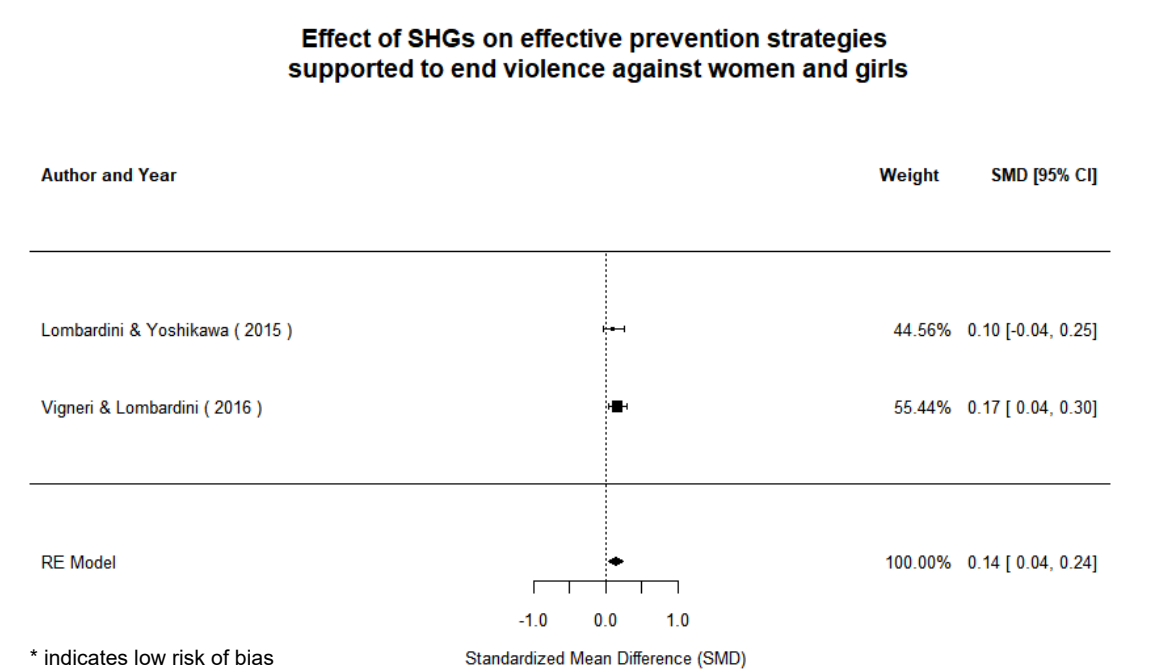
Figure 86: SHGCA2: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effect of SHGs on effective prevention strategies supported to end violence against women and girls

Two studies reported disaggregated data for effective prevention strategies supported to end violence against women and girls, thus we included $k = 2$ studies in the analysis. We assessed both studies as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.1389$ (95% CI: 0.04 to 0.24). Therefore, the average outcome differed significantly from zero ($z = 2.8084$, $p = 0.005$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 87 SHGCA6. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.3792$, $p = 0.538$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, and given there was no heterogeneity among the effects, moderator analyses were not possible, and tests of publication bias are not valid.

Figure 87: SHGCA6: Forest plot showing the observed outcomes and the estimate of the random-effects model



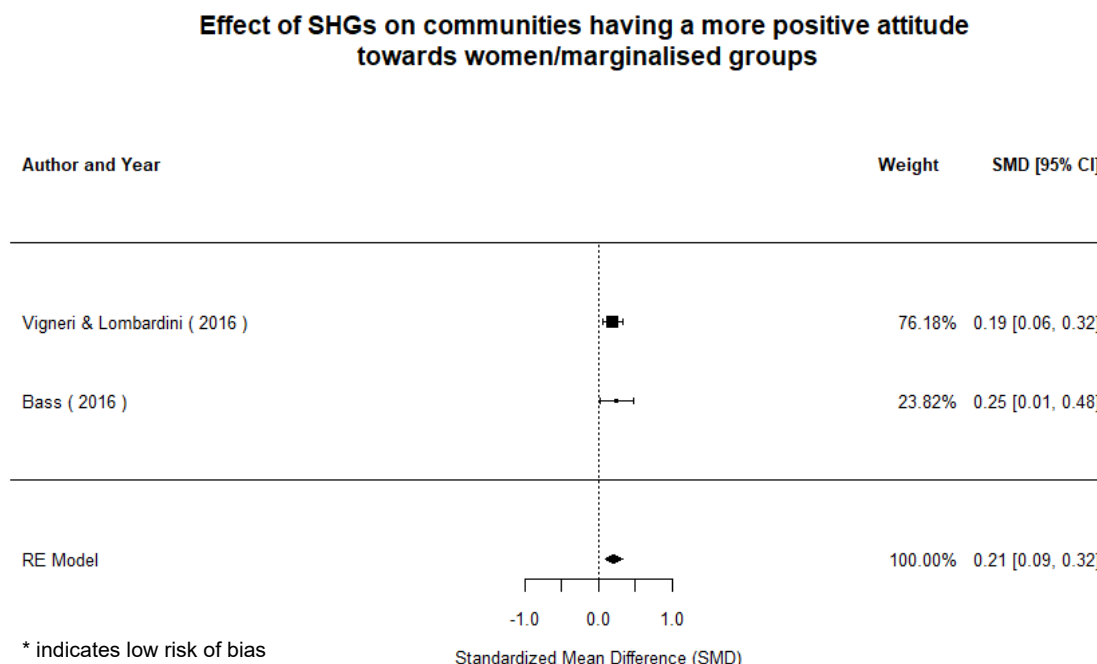
Effect of SHGs on relief and recovery initiatives in conflict and post conflict situations respond to the needs of women and girls, especially vulnerable groups

Vigneri and Lombardini's (2016) quasi-experimental study in Mali was the only study evaluating the impact of self-help groups and VSLA on relief and recovery initiatives in conflict and post conflict situations respond to the needs of women and girls, especially vulnerable groups. There was a small, and statistically significant, effect ($g = 0.19$, [95% CI: 0.06 to 0.32]), and we assessed the study as having high risk of bias.

Effect of SHGs on communities having a more positive attitude towards women/marginalised groups

Two studies reported disaggregated data for having a more positive attitude towards women/marginalised groups, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as having some concerns of risk of bias and the other as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.2066$ (95% CI: 0.09 to 0.32). Therefore, the average outcome differed significantly from zero ($z = 3.5625$, $p = 0.0004$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 88 SHGCB2. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.1534$, $p = 0.6953$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, and given there was no heterogeneity among the effects, moderator analyses were not possible, and tests of publication bias are not valid.

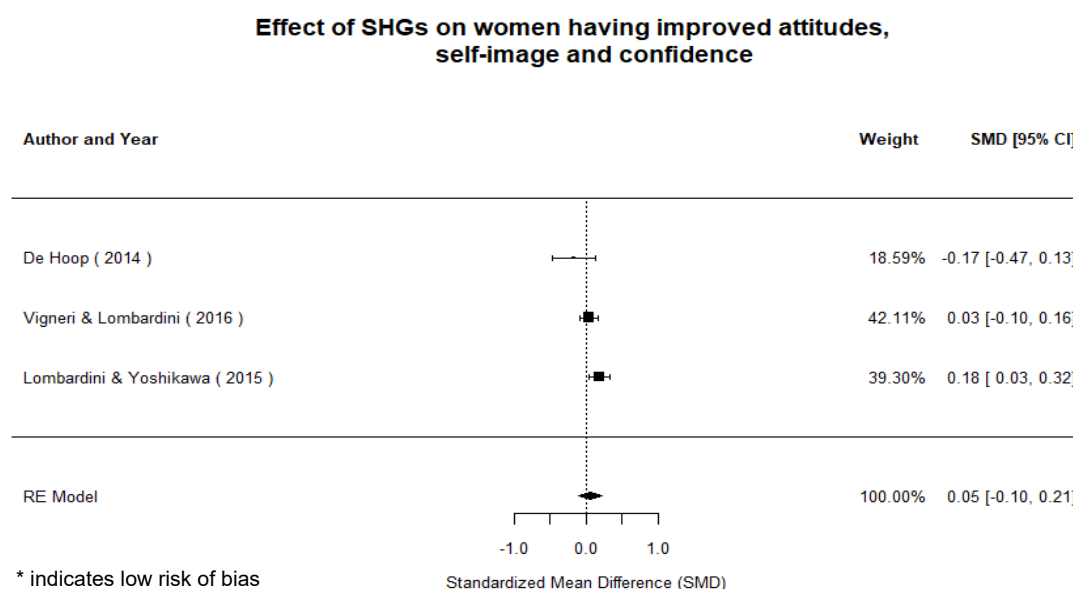
Figure 88: SHGCB2: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effect of SHGs on women having improved attitudes, self-image and confidence

We included a total of $k = 3$ studies in the analysis. We assessed none of the studies as low risk of bias, one as some concerns, and two as high risk of bias. The observed outcomes ranged from -0.17 to 0.18 , with the majority of estimates being positive (67 per cent). The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.05$ (95% CI: -0.10 to 0.21). Therefore, the average outcome did not differ significantly from zero ($z = 0.66$, $p = 0.51$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 89 SHGCB3.

Figure 89: SHGCB3: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(2) = 4.83$, $p = 0.09$, $\hat{\tau}^2 = 0.01$, $I^2 = 58.60\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.39 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

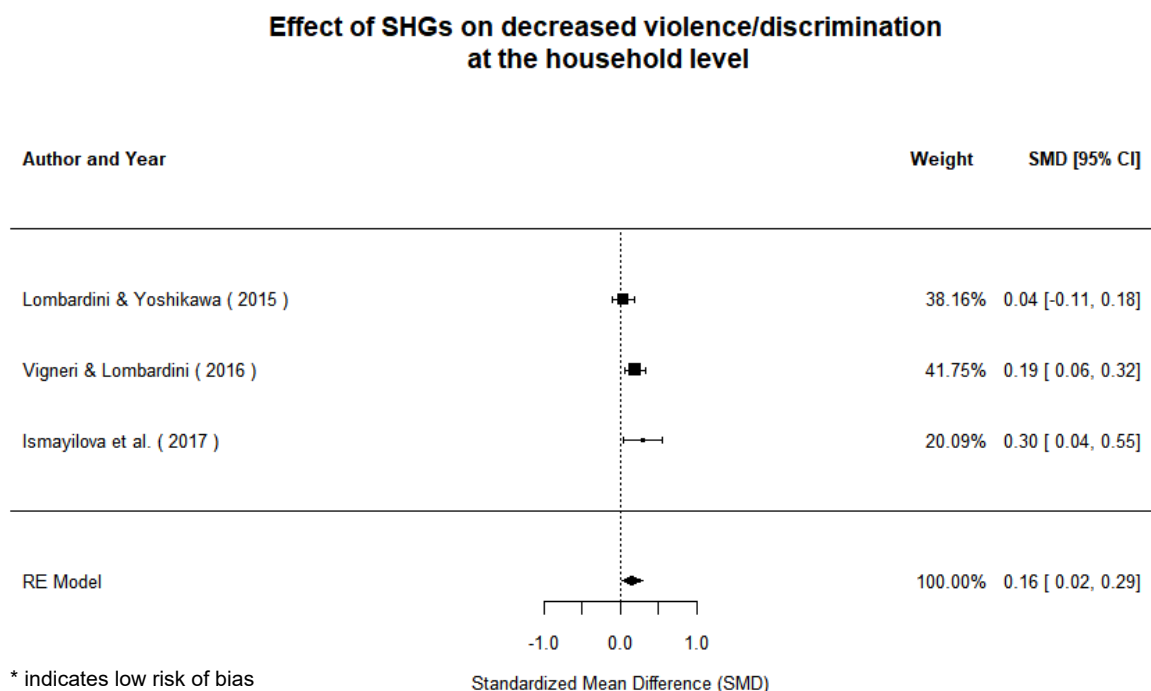
Effect of SHGs on attitudes and support for women's economic, social and human rights by men, household and family members and community members

Alemu and colleagues' (2018) quasi-experimental study in Ethiopia was the only study evaluating the impact of self-help groups and VSLA on improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members. Their report included 12 effects that fell into this outcome category (e.g., women should be more involved in politics). The effects ranged from medium, negative point estimates ($g = -0.25$, [95% CI: -0.53 to 0.04]) to large, positive effects ($g = 2.01$, [95% CI: 1.70 to 2.33]). We assessed the study as having some risk of bias concerns.

Effect of SHGs on decreased violence/discrimination at the household level

We included a total of $k = 3$ studies in the analysis. We assessed none of the studies as low risk of bias, one as some concerns, and two as high risk of bias. The observed outcomes ranged from 0.04 to 0.30. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.16$ (95% CI: 0.02 to 0.29). Therefore, the average outcome differed significantly from zero ($z = 2.25$, $p = 0.02$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 90 SHGCB5.

Figure 90: SHGCB5: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(2) = 3.95$, $p = 0.14$, $\hat{\tau}^2 = 0.01$, $I^2 = 49.36\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.39 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

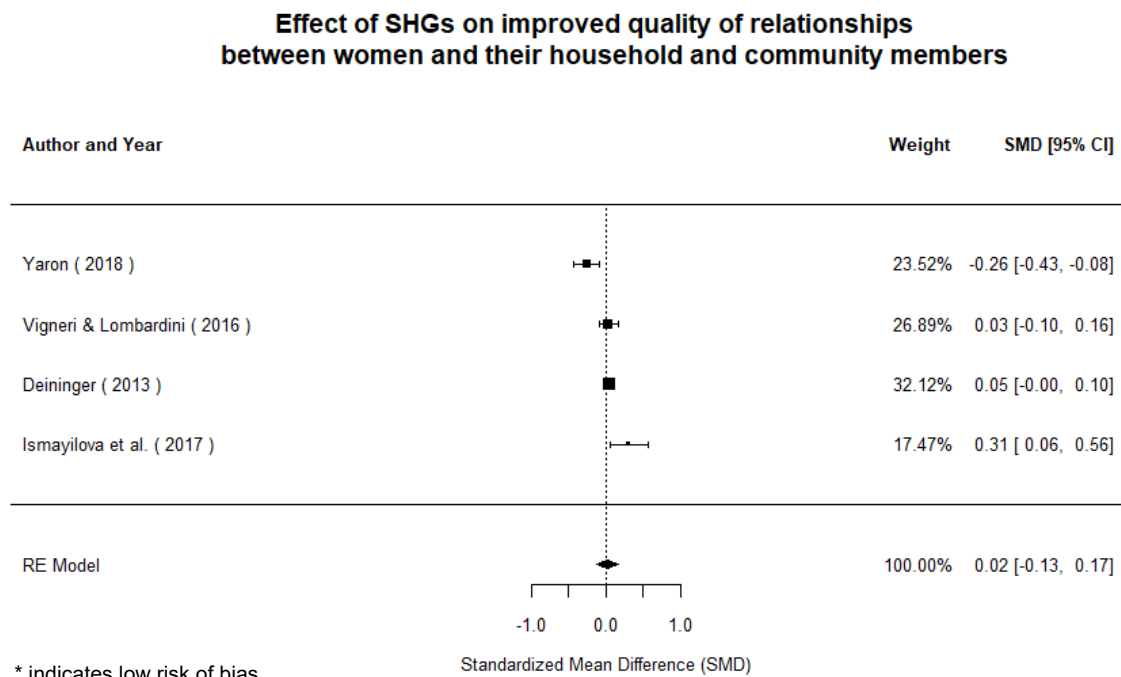
Effect of self-help groups on safer and more secure household, communities and areas/territories for women, girls, men, and boys

Bass's (2016) experimental study in the Democratic Republic of Congo was the only study evaluating the impact of self-help groups and VSLA on safer and more secure household, communities and areas/territories for women, girls, men, and boys. There was a small, but not statistically significant, point estimate ($g = 0.19$, [95% CI: -0.05 to 0.42]), and we assessed the study as having some risk of bias concerns.

Effect of SHGs on improved quality of relationships between women and their household and community members

We included a total of $k = 4$ studies in the analysis. We assessed none of the studies as low risk of bias, three as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.26 to 0.31 (75 per cent). The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.02$ (95% CI: -0.13 to 0.17). Therefore, the average outcome did not differ significantly from zero ($z = 0.24$, $p = 0.81$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 91 SHGCB8.

Figure 91: SHGCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(3) = 15.89$, $p < 0.01$, $\hat{\tau}^2 = 0.02$, $I^2 = 81.12\%$). An examination of the studentised residuals revealed that one study (Yaron 2018) had a value larger than ± 2.50 and may be a potential outlier in the context of this model.. According to the Cook's distances, none of the studies could be considered to be overly influential. Sensitivity analyses leaving each study out indicated that removing Yaron (2015) would increase the average effect ($\hat{\mu} = 0.08$ [95% CI: -0.02 to 0.18], but the effect would still be insignificant ($z = 1.52$, $p = 0.13$).

We were able to examine moderation by exposure to intervention, evaluation period, study design, publication year, adjustment for covariates, GII and FSI, but none were significant predictors of improved quality of relationships between women and their household and community members.

Effect of SHGs on women's empowerment index

We included a total of $k = 5$ studies in the analysis. We assessed none of the studies as low risk of bias, two as some concerns, and three as high risk of bias. The observed outcomes ranged from 0.03 to 0.34. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.09$ (95% CI: 0.03 to 0.15). Therefore, the average outcome differed significantly from zero ($z = 2.95$, $p < 0.01$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 92 SHGCC1.

Figure 92: SHGCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(4) = 10.092$, $p = 0.04$, $\hat{\tau}^2 = 0.01$, $I^2 = 60.35\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than ± 2.58 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

We were able to test several moderators. Studies with a high risk of bias were significantly different than studies with some concerns related to risk of bias ($\hat{B} = 0.14, p = 0.01$ [95% CI: 0.04 to 0.24]), such that high risk of bias studies had higher effects than studies with some concerns by .14 standard deviation units. Exposure to intervention in months ($\hat{B} = 0.004, p = 0.022$ [95% CI: 0.0006 to 0.0079]) was significant, such that each additional month of exposure to the intervention increased the size of the effect by .004 standard deviation units. There were also several moderators that could not be tested in the context of this model. These included study design, evaluation period and whether the model was adjusted for covariates.

Qualitative findings

We conducted a thematic synthesis on the 12 linked qualitative studies to the included SHGs and VSLA interventions. As indicated above, this thematic synthesis aims to identify themes related to the interplay of intervention design, intervention implementation, target population, and contextual variables with intervention outcomes and effects. In total, we identified 17 descriptive themes, which we configured into 4 analytical themes (Online Appendix N). These four analytical themes present the synthesis results and are discussed in more detail below.

Theme 1: Training of leaders and members of self-help groups and VSLAs is crucial to encourage uptake and the support of other groups and agents is valuable for conflict management.

Self-help groups and VSLAs can benefit from training of leadership and members in conjunction with agent and group support. Training activities within these groups was perceived to enhance confidence and trust in the leadership and capability among beneficiaries. For example, in the Saving for Change (SfC) programme in Mali, the participants accepted the programme because replicating agents were considered legitimate; this legitimacy was derived from the formal training and the received certificate d (IPA 2013). The case of Mali's SfC programme clearly shows the importance of training replicating agents in building confidence among members as well as propelling the status and perception by women in the community. As a clear expression of this, one replicating agent in N'Gorosso Peul succeeded in forming a group but faced challenges in forming other groups in other villages because the agent had not received formal training. For people to believe and trust her, she stated that she needed a technical agent as an escort (IPA 2013). According to the replicating agent's comments,

"Everything depends on the TA," she explained. "Women listen to the TA because he comes from outside the village and has more knowledge and has been to school. The women respect me but they do not always have confidence that my solutions to problems are the right ones. When an outsider speaks to them, they listen more."

The replication agent also cited the technical agent's higher level of education and training as the major difference between them, as the community members were more likely to respect the authority of an educated "outsider" more than a local. Another replicating agent from Nekeroro who had received the structured replication training remarked:

"Now that I can show women that I left the village and had training, they understand that I am better able to answer their questions when the TA is not here."

A CAEB coordinator in the SfC programme acknowledged that one of the biggest advantages of being part of the association is the advice from other groups on the handling of internal problems and pooling of resources for large-scale projects. Group support is also seen to promote participation through trust building in the system of self-help groups and VSLAs. In the SfC programme in Mali, there were groups that refused to join the association due to lock-in costs. To garner interest, the agent brought women from different groups to visit and observe association meetings and report back to their respective groups (Bureau of Applied Research in Anthropology and IPA 2008). These traditional modes of raising awareness, namely word-of-mouth and workshops, attracted non-beneficiaries as stories of positive programme outcomes are shared in communities where saving activity improved, and borrowing was used positively (IPA 2013).

Though there were mixed feelings by women from the SfC programme regarding the continued visits of agents, they appreciated the support, stating that they provided valuable help (Bureau of Applied Research in Anthropology and IPA 2008). To ensure sustainability, there is a need to maintain continuous presence in the form of NGO agents, for instance, in areas where the programme is relatively new. These stakeholders can be important for resolving issues related to financial management of funds (Bureau of Applied Research in Anthropology and IPA 2008). For instance, Mali, a president of an SfC group commented, “We need Tonus agents here. Everyone agrees on this point. Sometimes there are important misunderstandings about budget management or loan reimbursement that only they can resolve. An outside eye is needed. If there’s a misunderstanding, or people don’t fulfil their obligations, they need someone to come and explain. Most women rely on collecting and selling shea nuts to get income to pay their contribution or loans. This year there were not enough fruits and people didn’t respect the repayment date. This is when someone from the outside is needed for arbitration.”

Theme 2: Self-help groups and VSLAs are an attractive and convenient option for women who are often excluded from meaningful economic participation but factors related to market access and market conditions can minimise the gains from these programmes.

Due to the inclusive nature of self-help groups and VSLAs, women who could not engage in prior saving and borrowing activities can be granted better opportunities for economic participation. Women from the SfC programme, for instance, reported appreciating the SfC loans as they had less stringent requirements (no need for collateral) than formal credit programmes (IPA 2013). As noted in the Bureau of Applied Research in the Anthropology and IPA (2008) study, poorer households who have fears of taking loans tend to use these programmes as a means of saving. Further evidence for the need to be inclusive to the vulnerable groups is found in this study, which notes that in some villages there are artisanal caste groups that are concentrated in the outlets but are excluded due to distances from the villages. Married women were perceived to be more influential because they enjoy greater social capital and labour access than unmarried women and widows. The former were able to take more risks and/or repay loans, which influenced the nature and extent of borrowing and saving activities (Bureau of Applied Research In Anthropology and IPA 2008).

Access issues and environmental risk were a major concern for the reported success of self-help groups and VSLAs, particularly in the African context. In Mali, for example, villages were constrained by access to markets and centres which increased transportation costs. The prices of produce and goods purchased also varied due to changes in local, national and international market forces. The potential consequence is the negative impact on profits from goods produced and the increase in cost of transporting purchased goods. This prevented women from engaging in some activities, such as diversifying their crops (Bureau of Applied Research in Anthropology and IPA 2008). The absence of formal financial institutions in villages and the lack of trust in such organisations in some cases pushed rural households to have a general preference for investing their savings in animals and commerce rather than banking systems (IPA 2013).

Lack of water for domestic use presents itself as a major observed challenge for women who then spend a substantial amount of time looking for water, especially in the dry season when many sources of water run dry (Bureau of Applied Research in Anthropology and IPA 2008). In many of the villages, droughts severely affected crop yields whilst incidences of malaria and water borne diseases put a serious strain on savings and labour resources (Bureau of Applied Research in Anthropology and IPA 2008; IPA 2013).

Theme 3: Male dominance in household decision making, cultural attitudes towards women's roles and mobility, religion, time constraints exert negative pressure towards gains associated with self-help groups and VSLAs.

Qualitative findings show that because of their greater financial contributions to the household through participating in self-help groups and VSLAs, women can realise increases in household decision-making power and sense of independence (Bermudez and Matuszeski 2010; Deubel and Boyer, 2020; Concern Worldwide 2017; Nair 2012), reduced household-level tension and marital conflicts (Bermudez and Matuszeski 2010; Deubel and Boyer 2020; IPA 2013). In an Indian study by Nair (2012) for example, one respondent highlighted, "After I started working with SEWA, I was consulted for big financial decisions in my family. However, I make the small financial decisions of the family." -Staff F. In Mali, an animator (replication agent) summarised the increased independence trend, noting that "SfC has changed women. They can say yes or no now. Before women would always agree with what men said," (Bermudez and Matuszeski 2010).

Due to gender and cultural norms, men may dominate household decisions, including participation in savings groups, financial decisions and spousal mobility (Deubel and Boyer 2020; Bureau of Applied Research in Anthropology and IPA 2008) which impedes programme uptake. Consequently, women's empowerment is constrained within social bounds. For example, in Mali, some women from Kalifabougou highlighted that they had to ask permission from their husbands to join the savings groups (Bureau of Applied Research in Anthropology and IPA 2008). In Zambia, findings from Concern Worldwide (2017) showed that some men were not supportive of women participating in community groups, which they saw as a waste of time, especially when women came back from training "empty handed" (Concern Worldwide 2017). Results from Nair (2012) capture the influence of in-laws in making financial decisions in rural India, with one woman stating that:

“Every family decision is made by the father-in-law and we have to follow. I save Rs. 500 (\$9 USD) in my account and he asks me, what is the need. Any investment, saving or expenditure, we have to ask him.” – Staff G.”

Lack of support and time constraints resulting from gender skewed domestic workload prevented women from either participating or fully participating in self-help groups and VSLAs (IPA 2013; Nair 2012). In rural Gujarat, India, women’s primary role and responsibilities were household chores and taking care of family members. According to a respondent staff S. from the Nair (2012) study, “Even if a woman is working and earning for the family, she is expected to keep up with the chores, or any small business by a family member. Staff S. continues: “No, my husband nor any male family members do not contribute to domestic chores. It is completely my responsibility. I also have to contribute to his work and my contribution is hardly ever acknowledged.”

Another important factor that deters uptake of self-help groups and VSLAs was religious beliefs and stigma regarding loans. For example, in Mali, strong religious interpretations effectively prevented women and households from taking loans outright (IPA 2013). Traditionally, in some Malian cultures, getting loans from non-relatives is considered shameful, though not forbidden (Bermudez and Matuszeski 2010).

Theme 4: Support from family members and partners can enhance programme participation and association rules improves group management. Domination of elites can stifle development of new leaders in the community whereas legislation can inhibit expansion.

Family and partner support emerged as one of the critical factors in determining participation in self-help groups and VSLAs (Bureau of Applied Research in Anthropology and IPA, 2008; Bermudez and Matuszeski, 2010; Nair, 2012; Concern Worldwide 2017). There is some evidence that, in India, lack of family support is one of the most difficult barriers to overcome in enhancing participation in self-help groups and VSLAs. According to Nair (2012), women pointed out that one should not push their parents or other family members to support them too much: “*It happens slowly, you have to give them time. You have to walk with all of them, you can’t leave them behind. You have to take them all along with you. If we don’t do so, SEWA’s reputation will also be spoilt.*” The support includes being allowed to move freely and conduct daily life activities to balance paid and domestic work. Partner support also becomes key even in the absence of general family support. Sentiments from one woman of the SEWA programme in India adequately captures this barrier: “*I am married into a Muslim family and it is difficult for girls to progress in the community. If you are going out, you have to tell every member of the family. My husband is very supportive. My extended family and relatives dislike that I work.*” - Staff Y. In Mali, husbands of members assisted members with loan repayments in certain situations which had disempowered members from repaying loans, such as illness and business closure. In one household interview in Kalifabougou, an SfC member indicated her doubts about repaying her loan and expected her husband to assist with any outstanding balance to which the husband agreed (Bureau of Applied Research in Anthropology and IPA 2008).

Setting rules of participation by group members provided some comfort, enhanced inclusiveness and minimised chances of default by members. As captured in IPA (2013) regarding the SfC programme, one replicating agent Werekela explained that: “*In our*

group, we set the rules about how much money will be contributed, how loans are taken and what will happen if someone does not repay. Therefore, we understand how to manage the group and there is less fear that if a woman borrows money, she will not understand the terms of how to repay it and will become indebted. This is what can happen when villagers use the caisse d'épargne. There are many rules and a woman who cannot read may also be afraid of participating when the rules are set by people she does not know."

Three contextual factors in Mali's self-help groups and VSLA groups affected the success of the SfC programme. First, the domination of elites. The advancement of existing leaders could potentially block the development of new female leaders in the community. There are cases where women in the villages who held prominent positions and generally wealthier assumed leadership of the SfC programme and exerted undue influence over other women from the village, such as in Bouugola (Bureau of Applied Research in Anthropology and IPA 2008; Bermudez and Matuszeski 2010). To avoid domination by village elites over time, instituting mandatory leadership rotation yearly or biannually would create a more democratic process of leadership change and increase the capacity-building component of the programme in general (Bureau of Applied Research in Anthropology and IPA 2008). Second, group conflicts can emanate from disagreements on the application of rules or general failure to meet obligations. Such conflicts resulted in women exiting the SfC programme (Bermudez and Matuszeski 2010). Lastly, the legislative framework was restrictive. The government of Mali generally forbids informal microfinance programmes though it allowed the SfC programme under the close watch of government officials. This legal grey area can have limited programme expansion and thereby curtailed programme reach (Bureau of Applied Research in Anthropology and IPA 2008).

Discussion

A total of 11 studies were included in the quantitative analysis ($n = 11$). Quantitative meta-analysis showed positive results for self-help groups and VSLAs across a number of outcomes including: 1) women's capacity to understand and use financial, banking and business services effectively, 2) women being more aware of their rights and the roles and responsibilities of duty bearers, 3) women having more positive attitude towards taking action to claim their rights and 4) reduced percentage of women agreeing with certain reasons that justify violence against women and girls.

Some possible pathways from the qualitative evidence base that may explain these quantitative observations are:

- Training self-help group and VSLA members can enhance financial literacy and business skills. Successfully sharing of adequate and relevant knowledge and skills between group members can be beneficial to group replication as well.
- Self-help groups and VSLAs can adopt traditional methods of raising awareness such as allowing prospective members to listen in on meetings, word-of-mouth and workshops to motivate responsible and informed participation and use of financial instruments.
- Social capital gains and the acquisition of skills from self-help groups and VSLAs seem to drive economic participation (including the effective use of loans and saving and labour participation).

So what is so special about self-help groups and SHGs?

Here are some potentially important characteristics and their justification:

- Consolidating results from both our quantitative and qualitative analyses, self-help groups and VSLAs seem to foster more participatory methods of engagement that can increase the confidence of women.
- Activities of self-help groups and VSLAs often challenge existing gender roles and norms by design and include women in activities that they were traditionally excluded from—this can provide a pathway for positive empowerment gains.
- Empowering women to lead can improve their views on self-worth and motivate them to take ownership of their rights.
- Clear rules of participation and communication of programme activities possibly increase adherence (repayments) and willingness to participate for beneficiaries.
- Factors that could limit these positive effects:
 - Legislation and compliance obligations as well as domination by smaller groups.
 - Intergroup conflict and control by a few elite leaders within the beneficiary pool.
 - Partner reluctance to support programme women on their daily activities due to pervasive gender norms and roles.

Summary of findings and discussion

We included 12 studies across seven countries in South Asia and Sub-Saharan Africa that evaluated the effect of SHGs and VSLAs. The story told by the quantitative evidence on the effect of SHGs and VSLAs is an overwhelmingly positive one. We found moderate positive and significant impacts of SHGs and VSLAs on increased capacity of women to understand and use financial, banking and business services effectively. We found additional small but significant positive impacts on women's increased access to and ownership of assets, credit and income, women's access to decent work, women being more aware of their rights and the roles and responsibilities of duty bearers, better life skills for women, effective prevention strategies against violence being supported, decreased violence/discrimination at the household level, and increased women's empowerment. There were also positive but non-significant effects on increased freedom of movement and association, women having more positive attitudes towards taking action to claim their rights, increased participation in decision making by women at the household or community level, women participating more in their communities, communities having more positive attitudes towards women, and women having improved attitudes/self-image/confidence. There were two negative but non-significant summary effects, which were on reduced percentages of women agreeing with reason justifying violence against women, as well as improved quality of relationships between women and their household/community members.

The body of evidence in this intervention category also allowed for us to test several moderators. While there were no ubiquitous patterns among these moderators, we did find lower effects on assets, credit and income in India than we did in other countries. For impacts on freedom of movement and association, there were larger effects for RCTs than there were for QEDs, and smaller effects in more gender inequitable places. Women's decision making was moderated by exposure to intervention (in months) such that longer interventions achieved larger effects. Effects on decision making by women were also smaller in more recent studies. Women's participation in their communities was moderated by study quality such that studies with high risk of bias reported larger

effects than those with some concerns related to risk of bias. Finally, longer interventions also achieved larger effects on women's empowerment, and there were smaller effects reported for women's empowerment in more recent studies. Overall, the GRADE assessments generally indicate a very low to low (with one moderate) certainty in this body of evidence.

Between the 12 quantitative studies and the 10 linked qualitative studies, we have a sufficient evidence base to support the outcomes observed. Through four different analytical themes, several contextual and implementation themes emerged when looking at SHGs and VSLAs. Prevailing gender norms and household structures can impede the ability of women to participate and actively challenge these norms that have prevented them from economic and social mobility. SHGs and VSLAs can exist outside or in conjunction with. The summary of our quantitative findings along with the GRADE certainty of evidence ratings are presented in Table 36.

Table 36: GRADE summary of findings and certainty of evidence on SHGs and VSLAs

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(AA1) Women have access to rights, services and opportunities.										
1	QED	very serious ^a	serious ^b	not serious	not serious	none	728	SMD 0.18 SD higher (0.03 higher to 0.32 higher)	⊕○○○ VERY LOW	Limited Importance
(AB1) Capacity of women to understand and use financial, banking, and business services effectively										
3	RCT – 3	serious ^c	serious ^d	not serious	not serious	none	8642	SMD 0.24 SD higher (0.09 to 0.39 higher)	⊕⊕○○ LOW	Important, but not critical
(AB2) Women have increased access and ownership to assets, credit and income.										
7	RCT – 4 QED – 3	very serious ^e	serious ^f	not serious	not serious	none	12521	SMD 0.07 SD higher (0.00 higher to 0.13 higher)	⊕○○○ VERY LOW	Important, but not critical
(AB3) Women and girls have equitable access to livelihood support services.										
1	QED	serious ^g	serious ^b	not serious	serious ^h	none	192	SMD 0.31 SD higher (0.03 higher to 0.6 higher)	⊕○○○ VERY LOW	Important, but not critical
(AB5) Durable and reliable housing for vulnerable populations, including women and girls										
1	QED	serious ⁱ	serious ^b	not serious	not serious	strong association	1046	SMD 0.39 SD higher (0.28 higher to 0.56 higher)	⊕⊕⊕○ MODERATE	Important, but not critical
(AC1) More women engaged in other micro, small, and medium-sized enterprises.										
1	RCT	serious ^j	serious ^b	not serious	not serious	none	8595	SMD 0.03 SD higher (0.07 lower to 0.07 higher)	⊕⊕○○ LOW	Important, but not critical
(AC2) Initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities										
3	RCT – 2 QED – 1	serious ^k	not serious	not serious	not serious	none	3632	SMD 0.13 SD higher (0.04 higher to 0.22 higher)	⊕⊕⊕○ MODERATE	Important, but not critical
(AC3) Improved capacity of women entrepreneurs										
1	QED	very serious ^l	serious ^b	not serious	not serious	none	728	SMD 0.19 SD higher (0.05 higher to 0.33 higher)	⊕○○○ VERY LOW	Important, but not critical
(BA2) Women have more and better control over their bodies and sexual health.										
1	QED	very serious ^m	serious ^b	not serious	serious ⁿ	none	728	SMD 0.01 SD lower (0.16 lower to 0.13 higher)	⊕○○○ VERY LOW	Critical
(BA3) Women have increased freedom of movement and association.										
7	RCT – 4 QED – 3	very serious ^o	serious ^p	not serious	serious ^q	none	14316	SMD 0.18 SD higher (0.05 lower to 0.31 higher)	⊕○○○ VERY LOW	Limited Importance
(BA4) Women are more aware of their rights and the roles and responsibilities of duty bearers.										
3	RCT – 1 QED – 2	very serious ^r	not serious	not serious	not serious	none	4793	SMD 0.1 SD higher (0.02 higher to 0.18 higher)	⊕⊕○○ LOW	Limited Importance
(BA5) Women have more positive attitude towards taking action to claim their rights.										
3	RCT – 1 QED – 2	very serious ^s	very serious ^t	not serious	not serious	strong association	4257	SMD 0.58 SD higher (0.03 higher to 1.14 higher)	⊕○○○ VERY LOW	Limited Importance

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(BA6) Reduced percentage of women agreeing with certain reasons that justify violence against women and girls										
3	QED – 3	very serious	very serious	not serious	serious	none	1827	SMD 0.23 lower (0.57 lower to 0.12 lower)	⊕○○○ VERY LOW	Critical
(BA7) Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them.										
2	QED – 2	very serious ^u	serious ^w	not serious	not serious	none	5192	SMD 0.16 SD higher (0.08 higher to 0.23 higher)	⊕○○○ VERY LOW	Important, but not critical
(BA8) Reduced instances of child or forced marriage										
1	QED	very serious ^v	serious ^b	not serious	not serious	none	907	SMD 0.03 SD higher (0.1 lower to 0.16 higher)	⊕○○○ VERY LOW	Critical
(BB1) Increased participation in decision making by Women at the household or community level, including during crisis response										
6	RCT – 3 QED -- 3	very serious ^w	very serious ^x	not serious	serious ^y	none	14186	SMD 0.04 SD higher (0.04 lower to 0.12 higher)	⊕○○○ VERY LOW	Limited Importance
(BB2) Women participate more in their community.										
4	RCT – 2 QED – 2	very serious ^z	very serious ^{aa}	not serious	not serious	none	9132	SMD 0.07 SD higher (0.01 lower to 0.16 higher)	⊕○○○ VERY LOW	Limited Importance
(CA1) Power holders have improved awareness and responsiveness to the demands, claims, rights and inputs of women.										
1	QED	very serious ^{ac}	serious ^b	not serious	not serious	none	728	SMD 0.09 SD higher (0.02 lower to 0.21 higher)	⊕○○○ VERY LOW	Limited Importance
(CA2) Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration										
2	RCT – 1 QED -- 1	serious ^{ad}	not serious	not serious	not serious	none	3205	SMD 0.09 SD higher (0.04 higher to 0.14 higher)	⊕⊕⊕○ MODERATE	Limited Importance
(CA6) Effective prevention strategies supported to end violence against women and girls										
2	QED – 2	very serious ^{ae}	not serious	not serious	not serious	none	1635	SMD 0.14 SD higher (0.04 higher to 0.24 higher)	⊕⊕○○ LOW	Critical
(CA8) Increased community support for women's and children's human, economic and legal rights										
1	QED	serious ^{af}	serious ^b	not serious	not serious	none	907	SMD 0.19 SD higher (0.06 higher to 0.32 higher)	⊕⊕○○ LOW	Important, but not critical
(CB2) Communities have a more positive attitude towards women/marginalised groups.										
2	RCT – 1 QED – 1	very serious ^{ag}	not serious	not serious	serious ^{ai}	none	4726	SMD 0.21 SD higher (0.09 lower to 0.32 higher)	⊕○○○ VERY LOW	Limited Importance
(CB3) Women have improved attitudes, self-image and confidence.										
3	QED – 3	very serious ^{aj}	serious ^{ak}	not serious	serious ^{al}	none	1635	SMD 0.1 SD higher (0.04 lower to 0.24 higher)	⊕○○○ VERY LOW	Limited Importance
(CB4) Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members										
1	QED	serious ^{am}	serious ^b	not serious	very serious ^{an}	none	192	Seven positive and five negative effect estimates, which vary widely	⊕○○○ VERY LOW	Limited Importance
(CB5) Decreased violence/discrimination at the household level										
3	RCT – 1	very serious ^{ao}	serious ^{ap}	not serious	not serious	none	1875	SMD 0.16 SD higher	⊕○○○	Critical

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
	QED – 2							(0.02 higher to 0.29 higher)	VERY LOW	
(CB6) Safer and more secure household, communities and areas/territories for women, girls, men, and boys										
1	RCT	serious ^{aq}	serious ^b	not serious	serious ^{ar}	none	283	SMD 0.19 SD higher (0.05 lower to 0.42 higher)	⊕○○○ VERY LOW	Critical
(CB8) Improved quality of relationships between women and their household and community members										
4	RCT – 2 QED – 2	serious ^{as}	very serious ^{at}	not serious	serious ^{au}	none	5729	SMD 0.02 SD (0.13 lower to 0.17 higher)	⊕○○○ VERY LOW	Limited Importance
(CC1) Empowerment/Equality index										
5	RCT – 2 QED – 3	very serious ^{av}	not serious	not serious	not serious	none	2481	SMD 0.09 SD higher (0.01 higher to 0.24 higher)	⊕⊕○○ LOW	Important, but not critical

Note. CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

- a. One study assessed as having a high risk of bias. Reporting bias is likely. The deviations from intended interventions and the outcome measurement bias are unclear.
- b. All single studies downgraded once for inconsistency.
- c. This analysis has been downgraded once due to some concerns related to risk of bias for both studies although there is no major concern.
- d. Downgraded once: 1. Variation of points estimates: no overlaps, 2. overlaps of CI: do not cross no effect line, 3. $I^2 = 95.46\%$, 4. test of stat: According to the Q-test, there was a significant amount of heterogeneity in the true outcomes
- e. This analysis has been downgraded twice due to one study having a high risk of bias and all other studies with some concerns related to risk of bias
- f. Downgraded to reflect widely varying point estimates and CIs that do not overlap.
- g. Downgraded as the deviations from intended interventions and the outcome measurement bias were unclear.
- h. Downgraded once because of very wide confidence intervals.
- i. Assessed as having some concerns related to risk of bias, due to all the criteria being unclear except outcome measurement bias.
- j. Downgraded because of risk of reporting bias.
- k. This analysis has been downgraded once due to some concerns related to risk of bias for both studies although there is no major concern
- l. Assessed as having a high risk of bias. Reporting bias is likely. The deviations from intended interventions and the outcome measurement bias are unclear.
- m. Assessed as having a high risk of bias. Reporting bias is likely. The deviations from intended interventions and the outcome measurement bias are unclear.
- n. Downgraded once because of very wide confidence intervals.
- o. This analysis has been downgraded twice due to two studies having a high risk of bias and all other studies with some concerns related to risk of bias
- p. downgraded once: 1. Variation of points estimates: points estimates on both side of the threshold, 2. overlaps of CI: not systematic overlap of CI.
- q. Downgraded once to reflect the effect estimates wide confidence interval.
- r. Downgraded twice because all studies present risk of bias, and two-thirds of the evidence comes from high risk of bias QEDs.
- s. This analysis has been downgraded twice due to one study having a high risk of bias and all other studies with some concerns related to risk of bias and a mix of RCT and QED
- t. 1. Variation of points estimates: points estimates on the same side of the threshold 2. overlaps of CI: no systematic overlaps of CI 3. $I^2 = 97.58\%$ 4. test of stat: the true outcomes appear to be heterogeneous downgrade twice

- u. Downgraded twice because both studies are high risk of bias QEDs.
- v. Assessed as having a high risk of bias. All the criteria are assessed as being likely biased, except the deviations from intended interventions being likely not biased and the outcome measurement bias being unclear.
- w. Downgraded twice times: 1. the RCTs have some concerns related to risk of bias 2. two studies have a high risk of bias 3. removing either high risk study would have an impact on the meta-analysis
- x. downgrade twice: 1. Variation of points estimates: points estimates on both side of the threshold, 2. overlaps of CI: no systematic overlaps of CI, 3. $I^2 = 53.69\%$, 4. test of stat: the true outcomes appear to be heterogeneous
- y. Downgraded once to reflect a very large p-value and wide-ranging CI of the meta-analytic effect estimate.
- z. Downgraded twice because all studies present risk of bias, and 2/4 of the evidence comes from high risk of bias QEDs.
- aa. Downgraded twice because one study is very inconsistent with the others, including CI's that do not overlap and a point estimate that varies considerably from all others. When this study is removed, the meta-analytic results change.
- ac. Assessed as having a high risk of bias. Reporting bias is likely. The deviations from intended interventions and the outcome measurement bias are unclear.
- ad. Downgraded due to risk of reporting bias.
- ae. Downgraded to very serious because all of the evidence in this group comes from high risk of bias studies.
- af. Downgraded due to risk of reporting bias.
- ag. This group has been downgraded twice because both included studies present at least some concern, and one of two studies present high risk. In particular, there was uncertainty or high risk of bias for all studies with respect to deviation from intended intervention.
- ai. Downgraded because the CI of the RE effect estimate are large, and cross over both sides of the threshold of interest.
- aj. Downgraded twice because: 1) All three studies included are high risk of bias QEDs, and 2) Removing any study would significantly impact the effect estimate and leave the group with one high risk of bias study.
- ak. Downgraded once to reflect very different effect sizes and CIs that cross both sides of the threshold.
- al. Downgraded because the RE effect estimate cross the threshold of interest and has CI that are larger than the effect estimate itself.
- am. Downgraded as the deviations from intended interventions and the outcome measurement bias were unclear.
- an. Downgraded twice due to widely varying point estimates far across both sides of the threshold of interest.
- ao. Downgraded twice because: 1) 79.91% of the weight of the meta-analysis comes from high risk of bias QEDs, and 2) The only non-high risk of bias study is unclear across several domain criteria like performance bias and outcome measurement bias.
- ap. Downgraded because there are point estimates on either side of the threshold and a very wide CI range across the studies.
- aq. Downgraded as having some concerns related to risk of bias. The unit of analysis criteria has a likely bias and the deviations from intended interventions criteria is unclear.
- ar. Downgraded due to very wide confidence intervals.
- as. 1/4 of the studies in this group was high risk of bias, and several others were unclear.
- at. This group includes studies that are both significantly positive and significantly negative, including studies whose CIs are entirely on one side of the other of the threshold of interest.
- au. Downgraded because the RE model effect estimate has wide CIs that reach widely onto both side of the threshold.
- av. Downgraded twice because 3/5 studies in this group comes from high risk of bias QEDs. There is concern or uncertainty with respect to selection bias, outcome measurement, performance bias (and other criteria) across all studies in the group.
- aw. Downgraded once because, despite the two point estimates being quite similar, the very wide CI of one study presents a risk of inconsistency. Given the low number of studies, this wide CI is of particular concern.

4.3.8 Technical and vocational education and training (TVET)

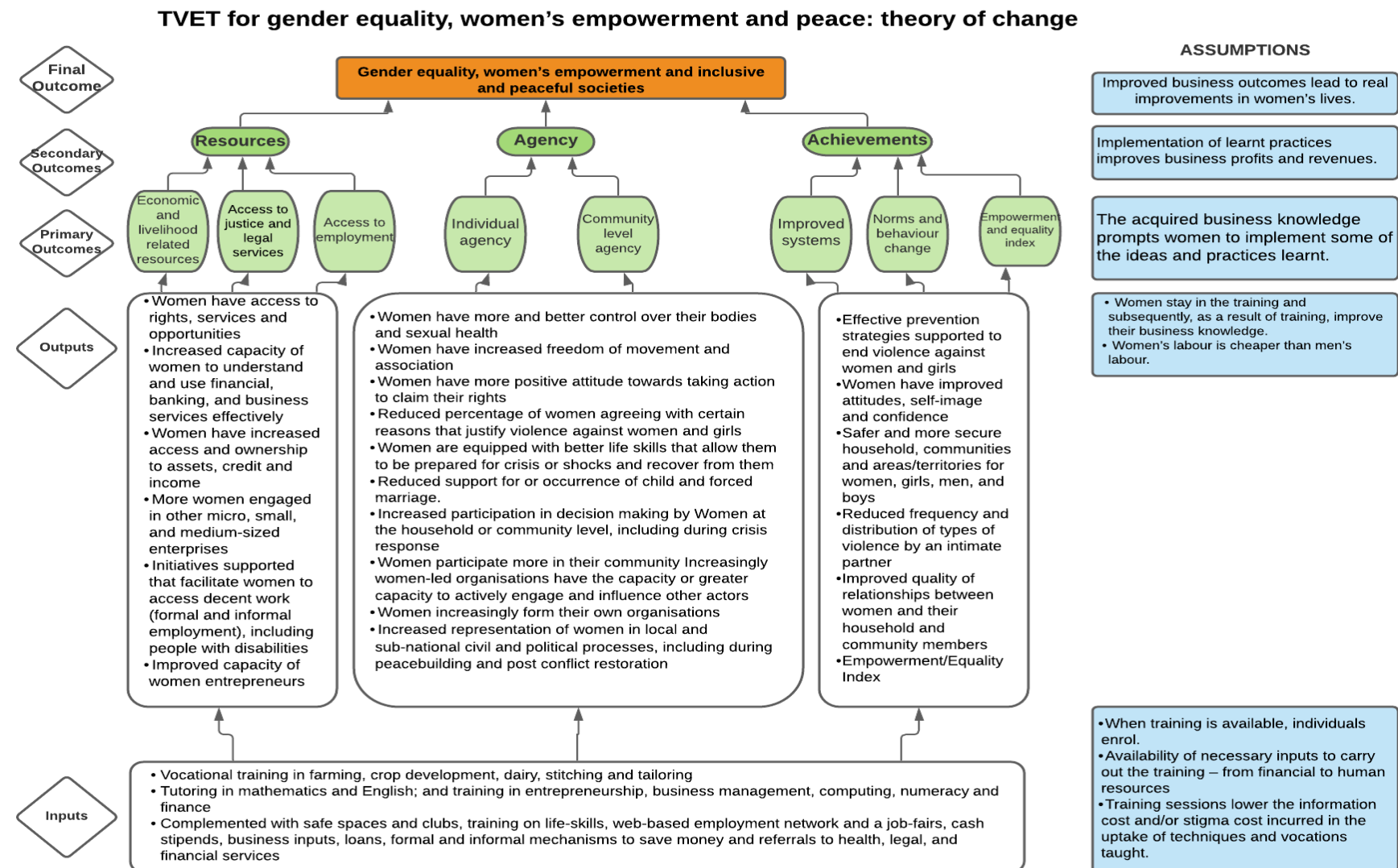
TVET are educational interventions primarily designed to lead participants to acquire practical skills, know-how and understanding necessary for employment in a particular occupation, trade or group of occupations or trades (Pompa 2014).

How does TVET affect gender equality, women's empowerment and peace outcomes?

Figure 93 below maps out the causal chain of how TVET may improve gender equality, women's empowerment and peace outcomes.

Vocational education and training in entrepreneurship, business management, computing, numeracy and finance, is expected to empower women through the development of skills that improve their access to financial services, employment, assets, and resources in general. These new skills help women participate more in decision making processes and in their communities with improved attitude, self-image and confidence to take action and claim their rights. This is based on the assumption that, when available, individuals take up training, stay in it and, subsequently, improve their vocational or business knowledge so that the implementation of learnt practices can increase business profits and revenues.

Figure 93: TVET for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included 10 studies reported in 12 different papers that evaluated the effect of eight identified programmes. We included more than one paper that evaluated the same programme if the author reported different outcomes over several papers.

Population

The interventions evaluated by included studies targeted adolescent girls aged 12 years or older, young women, college graduates and women farmers.

The included studies looked at eight different programmes and trials in South Asia and Sub-Saharan Africa and ten countries including: Afghanistan, Bangladesh, India, Kenya, Liberia, Nigeria, Pakistan, Sierra Leone, Uganda and Zambia.

Intervention, inputs and activities

The included studies evaluated a range of different TVET activities and inputs including the following:

- Vocational training plus resource provision (n = 3): Skill-building in numeracy and chosen vocational skill (such as sewing) together with foundational training in modules that include: the value of women's work, financial literacy, basic business skills and inputs, ways to improve health and well-being, women's rights and prevention of SGBV, strategies to make decisions and negotiate, civic action and advocacy, social networks and safety nets. Resources provision in the form of loans or cash stipend, mechanisms to save money, and referrals to health, legal and financial services.
- Tutoring and training (n = 7): Classroom-based technical and life skills training, with a focus on skills with high market demand such as stitching and tailoring, or farmer skills training and crop development.
- Entrepreneurship skills-building (n = 1): Training on gender and entrepreneurship, business woman and her environment, the business project, and people, organisation and management.
- Community groups training (n = 1): training on the dairy sector plus establishment of collection centres for collecting milk from small farm producers.

Table 37: Activities and design features of included TVET studies

Study	Activity/input	Length of treatment	Intervention frequency
Carney - 2018	Farmer skills training and crop development	NA	Unclear
Gibbs - 2020	Foundational training, skill-building in numeracy and a chosen vocational skill, cash stipend, formal and informal mechanisms to save money, and referrals to health, legal, and financial services, combined with safe spaces	Unclear	Weekly
Gibbs - 2018	Foundational training, skill-building in numeracy and a chosen vocational skill, cash stipend, formal and informal mechanisms to	Unclear	Weekly

Study	Activity/input	Length of treatment	Intervention frequency
	save money, and referrals to health, legal, and financial services, combined with safe spaces		
McKenzie - 2019	Training on gender and entrepreneurship, businesswoman and her environment, the business project, and people, organisation and management.	30	Daily
Lombardini - 2015	Community groups who received training on the dairy sector combined with collection centres and an established enterprise with a formal agreement with a major company in the dairy sector for selling the milk from collected sectors	24	Unclear
Croke - 2017	Classroom-based training, web-based employment network and a job-fair	13	Weekly
Bandiera - 2019	Peer education in life skills, safe spaces, vocational skills training, financial literacy, business inputs, loan	19	Daily
Adoho - 2014	Classroom-based technical and life skills training, with a focus on skills with high market demand, followed by placement and follow-up support in their transition to self or wage employment	12	Unclear
Bandiera - 2018	Safe spaces, life skills training, vocational skills training, recreational activities, financial literacy courses	22	Daily
Maitra - 2017	Stitching and tailoring vocational training	6	Daily
Bandiera - 2012	Safe spaces, life skills training, vocational skills training, recreational activities, financial literacy courses	22	Daily
Amin - 2016	Tutoring in mathematics and English (in-school girls), and computing or financial training (out-of-school girls); life skills training on gender rights and negotiation, critical thinking, and decision-making; and training in computers, entrepreneurship, mobile phone servicing, photography, and basic first aid	18	Weekly

Comparison

All our included studies compared treated groups to comparison groups receiving no intervention. Three studies included multiple treatment arms.

Outcomes

The included studies reported on a number of relevant outcomes, including the following:

- Women have increased access and ownership to assets, credit and income (n = 10): Women are able to apply for, receive and manage assets/credit and income and have support to manage, claim and execute their assets without pressure or influence from external actors, including male family members, husbands and cultural leaders.
- Women can access decent work (formal and informal employment) (n = 6): Women are able to apply for, receive and work in jobs (and have support for the above), without discrimination by sex, gender or other identifying factors, including development of skills for improved access.
- Improved capacity of women entrepreneurs (n = 5): The capacity can both cover technical abilities and resources to achieve a goal. Through skills training, microgrants, and linkage of services, women entrepreneurs are able to pursue their businesses and other pursuits.
- Increased participation in decision making by women at the household or community level, including during crisis response (n = 4): Women take part in all or any step of the decision-making process at the household community and district level, but also are able to meaningfully take part and have influence on the final decision, including in crisis response.
- Women have improved attitudes, self-image and confidence (n = 6): Women feel more entitled to claim their rights and needs in community and change social norms and behaviours. They are aware of the importance of their status in society and are empowered to take this role and make use of the available resources to guarantee their rights.

The division of the immediate and secondary outcomes is reported in Table 38 below:

Table 38: TVET summary of secondary and immediate outcomes

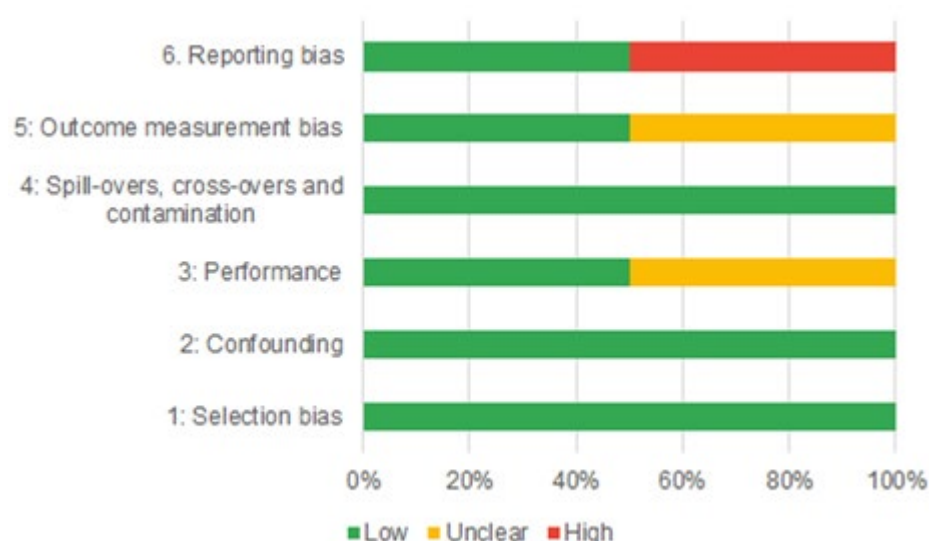
Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	0
	Economic and livelihood related resources	10
	Access to employment	8
Agency ability to define one's goals and act upon them and operationalised decision-making;	Individual agency	9
	Community level agency	6
	Institutions supporting agency	0

Secondary outcome category	Immediate outcome	Number of studies
Achievement ways of being and doing which can be realised by different individuals	Improved systems	3
	Norms and behaviour change	8
	Empowerment index	5

Study design

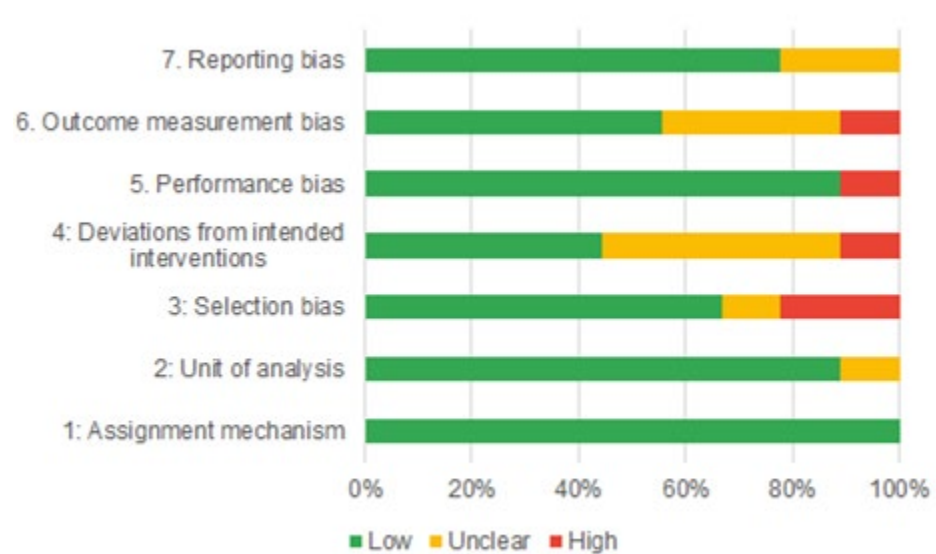
Two of the TVET studies used a QED (Carney 2018; Lombardini 2015). We assessed Carney (2108) as having a low risk of bias and Lombardini (2015) as having high risk of bias. As detailed in Figure 94 below, high risk of bias was identified on the reporting bias domain because the authors do not provide any robustness checks (Lombardini 2015). The same study did not provide enough information to discard issues related to outcome measurement and performance bias. No limitations related to reporting bias, nor issues with confounding or selection bias, were detected.

Figure 94: TVET studies QED RoB assessment



Nine of the TVET papers used an experimental design (Figure 95). This included Adoho (2014), Amin (2016), Bandiera (2012, 2018 and 2019), Gibbs (2018 and 2020), Maitra (2017) and McKenzie (2019). Only one study was assessed as having a low risk of bias and four were assessed as having a high risk of bias, the rest (n = 4) were assessed as having some concerns related to risk of bias. Some of the issues with RCTs evaluating TVET interventions were differences in attrition rates between treatment and control groups that could lead to selection bias (Gibbs 2020; Amin 2016); we also identified high risk of deviations from intended intervention in one study that reported non-compliance of the assignment to treatment (Bandiera 2012), and a risk of performance bias in a study that reported a possible change in behaviour in response to the presence of the program (Adoho 2014). The only category for which all studies were assessed as having low risk of bias was the assignment mechanism, while for unit of analysis and reporting bias, there was one and two studies that did not provide enough information to discard any issues.

Figure 95: TVET studies RCT RoB assessment



Qualitative studies, process evaluations and project documents

We identified five additional documents related to three programmes covered by the TVET group of studies.

- GET Ahead for Women in Enterprise (Kenya): one qualitative study and one descriptive quantitative study
- Economic Empowerment of Adolescent Girls and Young Women (EPAG) (Liberia): one process evaluation and one descriptive quantitative study
- Empowerment and Livelihood for Adolescents (ELA) programme (Sierra Leone): one descriptive quantitative study

All but one ($n = 4$) were ranked as moderate quality, with the remaining study marked as low quality.

Synthesis of findings

The following subsection presents the results of effectiveness of [intervention group] on Gender Equality, Women's Empowerment and Peace outcomes.

Quantitative findings

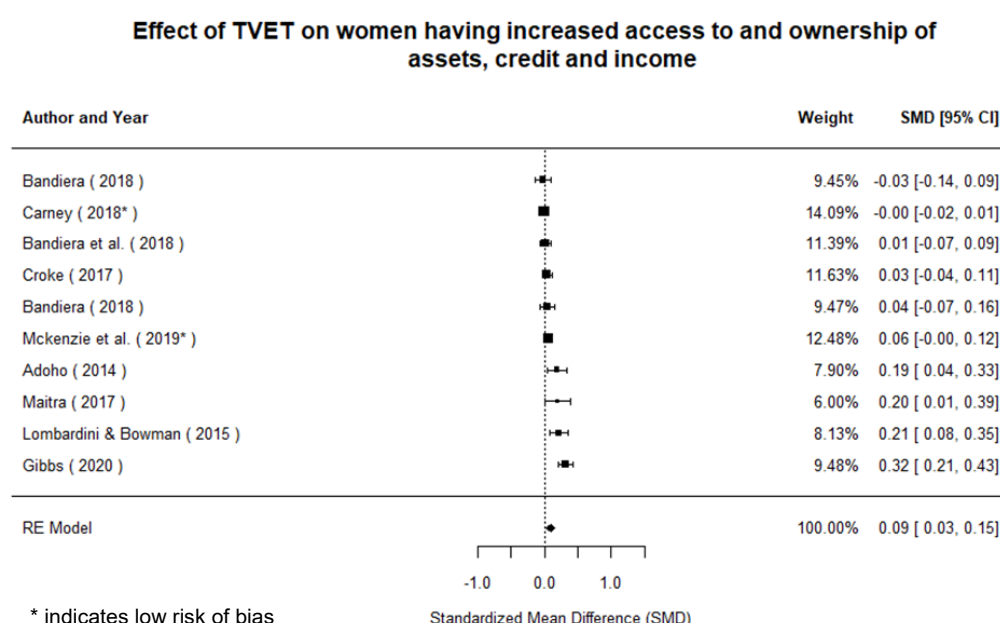
Effect of TVET on women having increased capacity to understand and use financial, banking and business services

Bandiera's (2018) experimental study in Sierra Leone was the only study evaluating the impact of TVET on women having increased capacity to understand and use financial, banking and business services. Their report included one effect for four groups that fell into this outcome category (e.g. financial literacy for 12-17 year olds in high and low disruption villages and 18-25 year olds in high and low disruption villages). The effects ranged from small, negative point estimates ($g = -0.05$, [95% CI: -0.16 to 0.06]) to small, positive point estimates ($g = 0.10$, [95% CI: -0.01 to 0.22]). The effects in the high disruption (Ebola) villages were larger than those in the low disruption villages, however none of the effect sizes were statistically significant at $p < .05$. We assessed the study as having some risk of bias concerns.

Effect of TVET on women having increased access to and ownership of assets, credit, and income

We included a total of $k = 10$ studies in the analysis. We assessed two studies as low risk of bias, three as some concerns, and five as high risk of bias. The observed outcomes ranged from -0.03 to 0.32 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.09$ (95% CI: 0.03 to 0.15). Therefore, the average outcome differed significantly from zero ($z = 2.85$, $p < 0.01$). Two of the Bandiera (2018) effects reflect separate effect for 12-17 year olds ($g = -0.03$) and 18-25 year olds ($g = 0.04$) from Sierra Leone, while Bandiera et al. (2018) is the effect from Uganda. A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 96 TVETAB2.

Figure 96: TVETAB2: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(9) = 52.67$, $p < 0.01$, $\hat{\tau}^2 = 0.01$, $I^2 = 82.91\%$). An examination of the studentised residuals revealed that one study (Gibbs 2020) had a value larger than ± 2.84 and may be a potential outlier in the context of this model. According to the Cook's distances, one study (Gibbs 2020) could be considered to be overly influential. Indeed, sensitivity analyses leaving each study out indicated that removing Gibbs (2020) would reduce the overall average effect ($\mu = 0.04$ (95% CI: 0.004 to 0.08), but the effect would still be positive and significant ($z = 2.15$, $p = .03$).

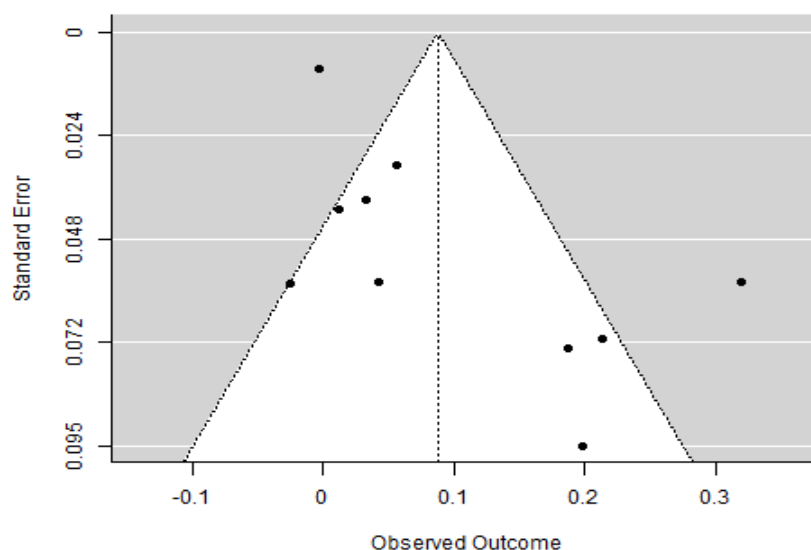
We were able to test several moderators in the context of this model. Exposure to intervention in months ($\hat{\beta} = 0.001$, $p = 0.76$ [95% CI: -0.005 to 0.01]) and evaluation period in months ($\hat{\beta} = 0.002$, $p = 0.64$ [95% CI: -0.01 to 0.01]) were also not significant. Effects from studies using experimental designs were similar to effects found in studies with quasi-experimental designs ($\hat{\beta} = 0.03$, $p = 0.70$ [95% CI: -0.13 to 0.20]). Studies examining change from baseline had similar effects to those examining post-intervention changes ($\hat{\beta} = -0.12$, $p = 0.07$ [95% CI: -0.26 to 0.11]). Effects from models that adjusted for covariates did not differ from models with no covariate adjustments ($\hat{\beta} =$

$-0.13, p = 0.29$ [95% CI: -0.38 to 0.11]). Effects also did not differ by scores on the gender inequality index ($\hat{\beta} = -0.50, p = 0.60$ [95% CI: -2.40 to 1.39]) or the fragile states index ($\hat{\beta} = 0.002, p = 0.50$ [95% CI: -0.004 to 0.01]). Finally, the qualitative evidence suggested that there might be a potential difference between programmes that targeted women only versus those that targeted both men and women. Thus, we tested this as another source of potential heterogeneity specifically for TVET programmes. We found that effects were not different between programmes exclusively targeting women versus mixed target programmes ($\hat{\beta} = 0.02, p = 0.79$ [95% CI: -0.16 to 0.21]).

The only significant moderator was the risk of bias assessment. Studies with a high risk of bias were significantly different than studies with some concerns or low risk of bias ($\hat{\beta} = 0.13, p = 0.03$ [95% CI: 0.01 to 0.26]), such that studies assessed as high risk of bias had larger effects by .13 standard deviation units.

A funnel plot of the estimates is shown in Figure 97. The regression test indicated funnel plot asymmetry ($p = 0.01$) but not the rank correlation test ($p = 0.38$).

Figure 97: TVETAB2b: Funnel plot



Effect of TVET on women being more engaged in micro, small, and medium-sized enterprises

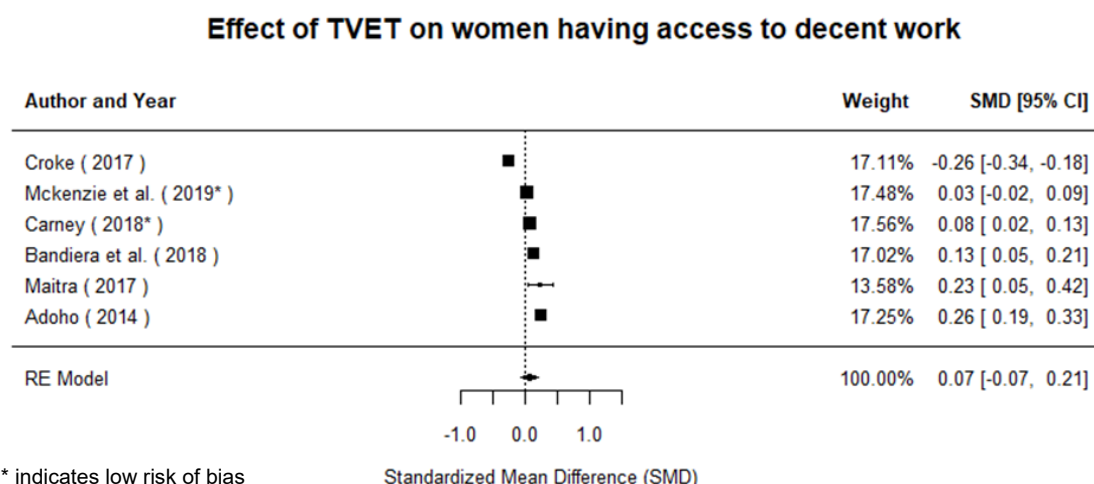
Mckenzie and colleague's (2019) experimental study in Kenya was the only study evaluating the impact of TVET on women being more engaged in micro, small, and medium-sized enterprises. Their report examined how many women were talking about their businesses, providing impact estimates at both the 12- and 36-month follow-up periods. The point estimate was small and positive for the 12-month follow-up ($g = 0.04$, [95% CI: -0.05 to 0.13]) while there was a small and significant positive effect at the 36-month follow-up ($g = 0.10$, [95% CI: 0.01 to 0.19]). We assessed the study as having a low risk of bias.

Effect of TVET on women having access to decent work

We included a total of $k = 6$ studies in the analysis. We assessed two studies as low risk of bias, two as some concerns, and two as high risk of bias. The observed outcomes ranged from -0.26 to 0.26 . The estimated average outcome based on the random-

effects model was $\hat{\mu} = 0.07$ (95% CI: -0.07 to 0.21). Therefore, the average outcome did not differ significantly from zero ($z = 1.03$, $p = 0.30$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 98 TVETAC2.

Figure 98: TVETAC2: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(5) = 106.30$, $p < 0.01$, $\hat{\tau}^2 = 0.03$, $I^2 = 95.30\%$). An examination of the studentised residuals revealed that one study (Croke 2017) had a value larger than ± 2.64 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. Indeed, sensitivity analyses leaving each study out indicated that removing Croke (2017) would increase the overall average effect ($\hat{\mu} = 0.14$ (95% CI: 0.05 to 0.22), and the resulting effect would be positive and significant ($z = 3.05$, $p = 0.002$).

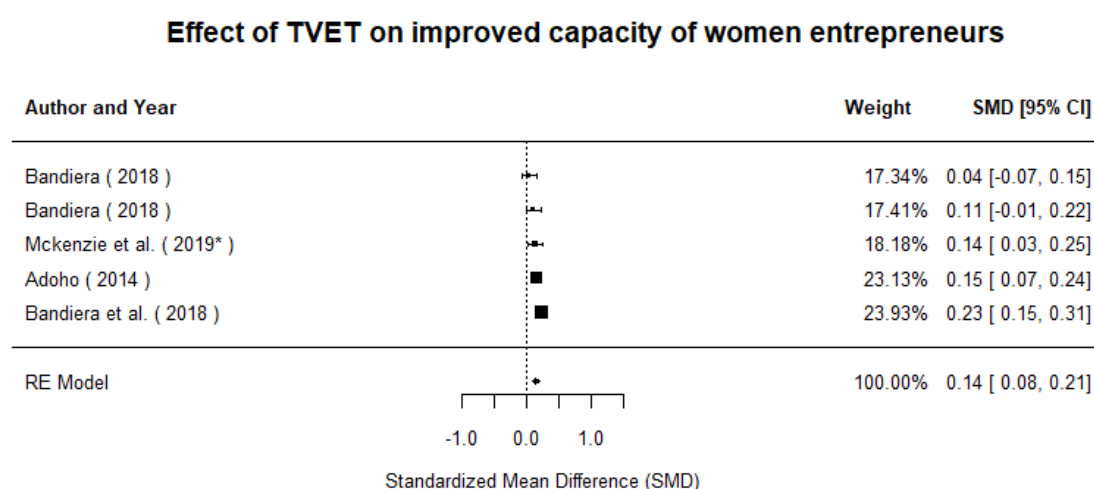
We were able to test several moderators in the context of this model. Studies with a high risk of bias were not significantly different than studies with some concerns or low risk of bias ($\hat{B} = 0.18$, $p = 0.17$ [95% CI: -0.08 to 0.45]). Exposure to intervention in months ($\hat{B} = 0.01$, $p = 0.25$ [95% CI: -0.01 to 0.04]) and evaluation period in months ($\hat{B} = -0.01$, $p = 0.07$ [95% CI: -0.02 to 0.001]). were also not significant. Studies using and experimental design had effects similar to those of studies using a quasi-experimental design ($\hat{B} = -0.14$, $p = 0.38$ [95% CI: -0.18 to 0.46]). Effects also did not differ by scores on the fragile states index ($\hat{B} = -0.01$, $p = 0.20$ [95% CI: -0.03 to 0.01]).

There were several moderators that could not be tested for this outcome group. First, only one study was looking at post-intervention changes, while the rest were examining change from baseline, thus we could not test this as a moderator. We were also unable to the for differences based on the gender equality index score as there was not sufficient variation. Finally, while we wanted to test for a potential difference between programmes that targeted women only versus those that targeted both men and women, only one study in this outcome group used a mixed-target approach, so we could not explore this as moderator.

Effect of TVET on approved capacity of women entrepreneurs

We included a total of $k = 5$ studies in the analysis. We assessed one of the studies as low risk of bias, two as some concerns, and three as high risk of bias. The observed outcomes ranged from 0.04 to 0.23. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.14$ (95% CI: 0.08 to 0.21). Therefore, the average outcome differed significantly from zero ($z = 4.35, p < 0.01$). Two of the Bandiera (2018) effects reflect separate effect for 12-17 year olds ($g = 0.04$) and 18-25 year olds ($g = 0.11$) from Sierra Leone, while Bandiera et al. (2018) is the effect from Uganda. A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 99 TVETAC3.

Figure 99: TVETAC3: Forest plot showing the observed outcomes and the estimate of the random-effects model



The Q -test for heterogeneity was not significant, but some heterogeneity may still be present in the true outcomes ($Q(4) = 8.46, p = 0.08, \hat{\tau}^2 = 0.00, I^2 = 52.74\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.58 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

We were able to test several moderators in the context of this model. Studies with a high risk of bias were significantly different than studies with some concerns or low risk of bias ($\hat{B} = 0.10, p = 0.04$ [95% CI: 0.003 to 0.20]) such that studies with a high risk of bias had larger effects by .10 standard deviation units. Exposure to intervention in months ($\hat{B} = 0.0000, p = 0.99$ [95% CI: -0.01 to 0.01]) and evaluation period in months ($\hat{B} = -0.0001, p = 0.99$ [95% CI: -0.02 to 0.01]) were also not significant. Finally, effects did not differ by scores on the gender inequality index ($\hat{B} = -0.89, p = 0.31$ [95% CI: -2.61 to 0.82]) or the fragile states index ($\hat{B} = 0.01, p = 0.23$ [95% CI: -0.01 to 0.03]).

All but one study was an experimental design, and all examined change from baseline. In addition, all studies only targeted women and all models adjusted for covariates. Thus, none of these could be tested as moderators.

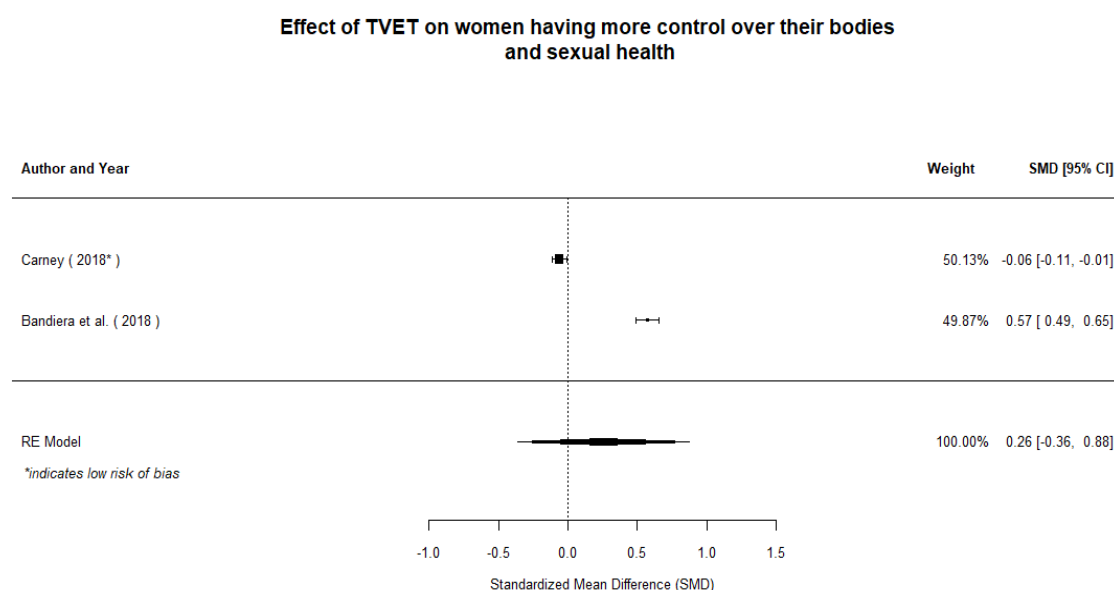
Effect of TVET on women having increased success in the workplace

Adoho's (2014) quasi-experimental study was the only study evaluating the impact of TVET on women having increased success in the workplace. Their report examined women's satisfaction with their job, finding moderate positive and significant impact their programme ($g = 0.23$, 95% CI [0.16, 0.31]). We assessed the study as having a high risk of bias.

Effect of TVET on women having more control over their bodies and sexual health

Only two studies reported disaggregated data for women having more control over their bodies and sexual health, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as low risk of bias and the other as high risk of bias. The observed outcomes ranged from -0.06 to 0.57 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.26$ (95% CI: -0.36 to 0.88). Therefore, the average outcome did not differ significantly from zero ($z = 0.81$, $p = 0.42$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 100 TVETBA2.

Figure 100: TVETBA2: Forest plot showing the observed outcomes and the estimate of the random-effects model

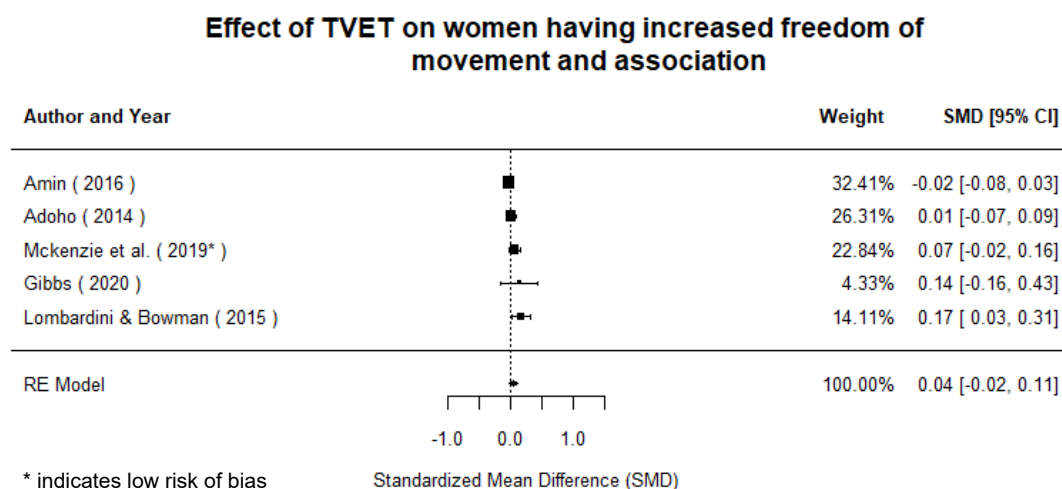


The Q -test for heterogeneity was significant, thus the true outcomes appear to be heterogeneous ($Q(4) = 171.96$, $p < 0.001$, $\hat{\tau}^2 = 0.20$, $I^2 = 99.42\%$). With only two studies, moderator analyses are not appropriate and tests of publication bias are not valid.

Effect of TVET on women having increased freedom of movement and association

We included a total of $k = 5$ studies in the analysis. We assessed one of the studies as low risk of bias, four as some concerns, and none as high risk of bias. The observed outcomes ranged from -0.02 to 0.17 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.04$ (95% CI: -0.02 to 0.11). Therefore, the average outcome did not differ significantly from zero ($z = 1.24$, $p = 0.21$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 101 TVETBA3.

Figure 101: TVETBA3: Forest plot showing the observed outcomes and the estimate of the random-effects model



The Q -test for heterogeneity was not significant, but some heterogeneity may still be present in the true outcomes ($Q(4) = 8.43, p = 0.08, \hat{\tau}^2 = 0.00, I^2 = 52.57\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.58 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

We were able to test several moderators in the context of this model. Studies with a high risk of bias were not significantly different than studies with some concerns or low risk of bias ($\hat{\beta} = 0.18, p = 0.17$ [95% CI: -0.08 to 0.45]). Exposure to intervention in months ($\hat{\beta} = 0.003, p = 0.29$ [95% CI: -0.002 to 0.01]) and evaluation period in months ($\hat{\beta} = 0.004, p = 0.60$ [95% CI: -0.01 to 0.02]) were also not significant. Studies using an experimental design had effects similar to those of studies using a quasi-experimental design ($\hat{\beta} = 0.04, p = 0.58$ [95% CI: -0.11 to 0.20]). Effects also did not differ by scores on the gender inequality index ($\hat{\beta} = -0.40, p = 0.74$ [95% CI: -2.73 to 1.93]), but they did differ by their scores on the fragile states index ($\hat{\beta} = 0.01, p = 0.01$ [95% CI: 0.003 to 0.02]) such that each 1 point increase on the fragile states index (indicating an increase in fragility) increased the effect by .01 standard deviation units. In other words, the more fragile states saw a larger benefit than the less fragile states.

Only one study was examining post-intervention changes, while the rest were exploring change from baseline. There was also only one study targeting both men and women, while the remainder targeted women only. Finally, one study was not adjusted for covariates while all other studies were. These moderators could not be tested.

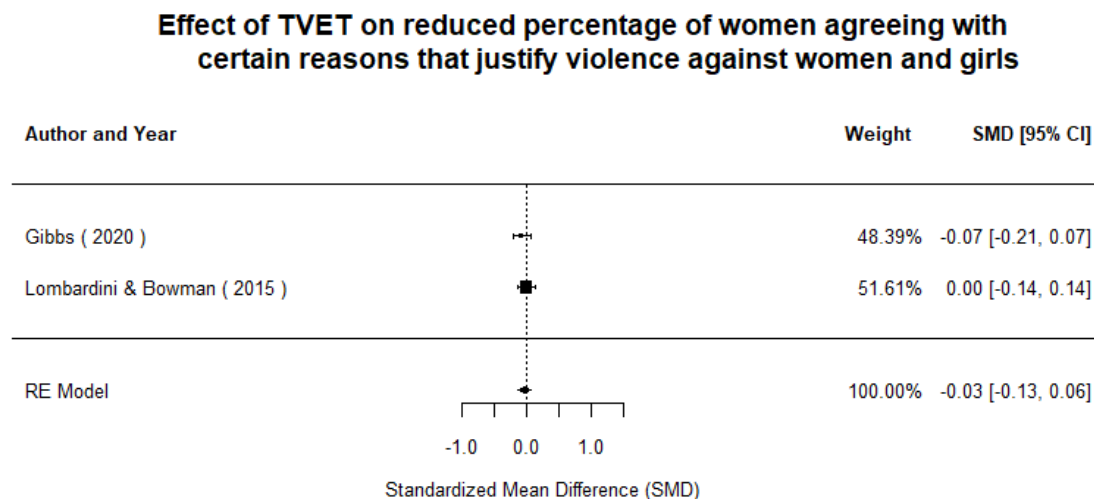
Effect of TVET on women's attitude towards taking action to claim their rights

Gibbs's (2020) experimental study in Afghanistan was the only study evaluating the impact of TVET on women having more positive attitude towards taking action to claim their rights. There was a very small, but not statistically significant, point estimate ($g = 0.03$, [95% CI: -0.08 to 0.14]), and we assessed the study as having high risk of bias.

Effect of TVET on reduced percentage of women agreeing with certain reasons that justify violence against women

Only two studies reported the impact of TVET on a reduced percentage of women agreeing with certain reasons that justify violence against women, thus we included $k = 2$ studies in the analysis. We assessed both of the studies as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.3$ (95% CI: -0.13 to 0.06). This positive effect did not differ significantly from zero ($z = -0.69$, $p = 0.49$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 102 TVETBA6. Given the small number of studies, as well as the fact that both studies were assessed as high risk of bias, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.51$, $p = .49$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator analyses were not possible, and tests of publication bias were not valid.

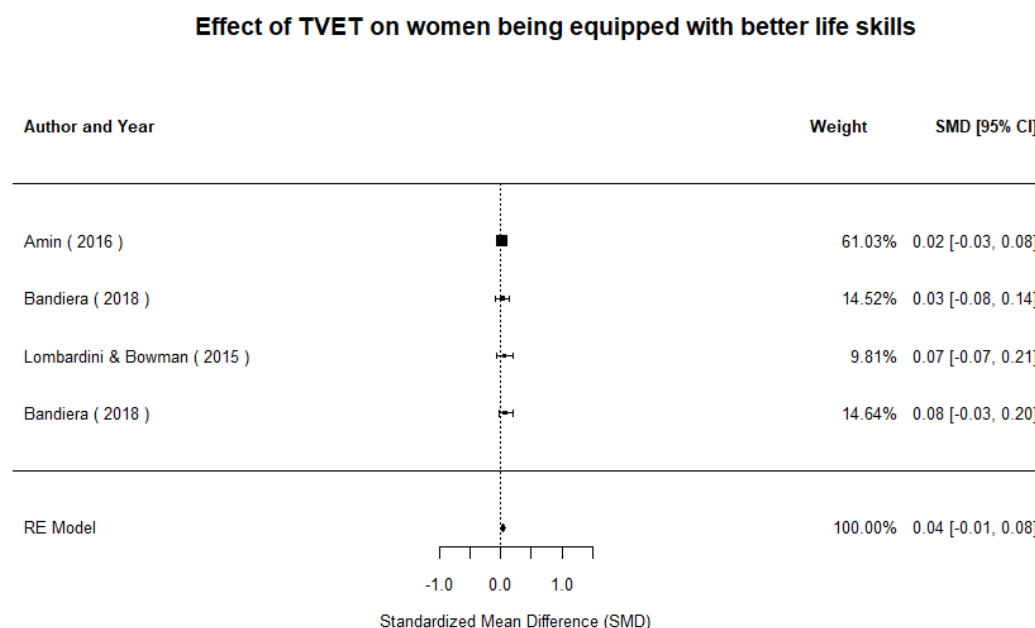
Figure 102: TVETBA6: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effect of TVET on women being equipped with better life skills

We included a total of $k = 4$ studies in the analysis. We assessed none of the studies as low risk of bias, two as some concerns, and two as high risk of bias. The observed outcomes ranged from 0.02 to 0.08. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.04$ (95% CI: -0.01 to 0.08). Therefore, the average outcome did not differ significantly from zero ($z = 1.70$, $p = 0.09$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 103 TVETBA7.

Figure 103: TVETBA7: Forest plot showing the observed outcomes and the estimate of the random-effects model

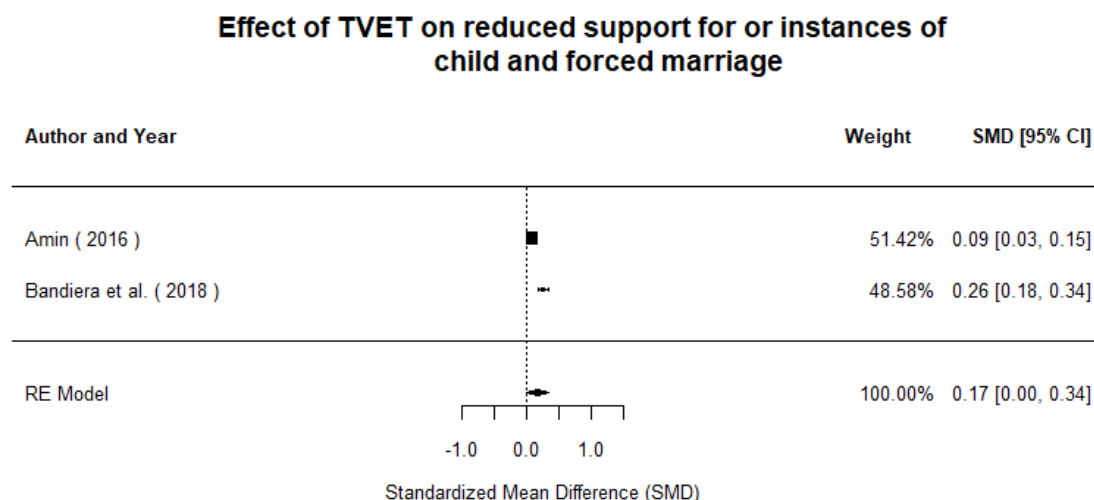


According to the Q-test, there was no significant amount of heterogeneity in the true outcomes ($Q(3) = 1.11$, $p = 0.77$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.50 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With no heterogeneity, moderator analyses were not appropriate.

Effect of TVET on reduced support for or instance of child and forced marriage

Only two studies reported the impact of TVET on a reduced support for or instance of child and forced marriage, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as having some concerns of risk of bias and the other as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.17$ (95% CI: 0.004 to 0.34). This positive effect differed significantly from zero ($z = 2.01$, $p = 0.04$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 104 TVETBA8. Given the small number of studies, as well as the one study was assessed as high risk of bias, and the other was assessed as some concerns, this result should be interpreted with caution. According to the Q-test, there was significant heterogeneity in the true outcomes ($Q(1) = 11.65$, $p < .001$, $\hat{\tau}^2 = 0.01$, $I^2 = 91.41\%$). With only two studies, moderator analyses were not possible, and tests of publication bias are not valid.

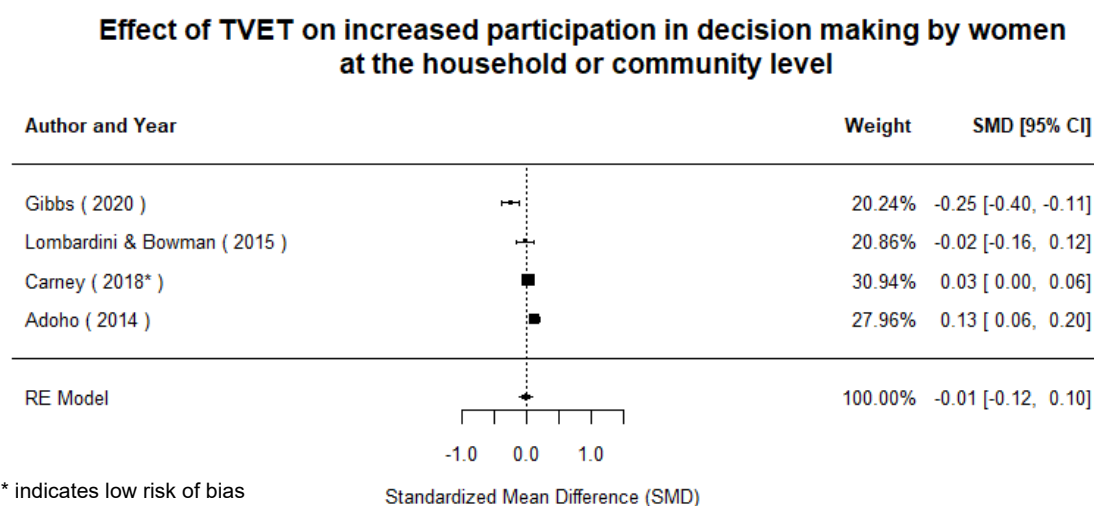
Figure 104: TVETBA8: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effect of TVET on increased participation in decision making by women at the household and community level

We included a total of $k = 4$ studies in the analysis. We assessed one of the studies as low risk of bias, none as some concerns, and three as high risk of bias. The observed outcomes ranged from -0.25 to 0.13 . The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.01$ (95% CI: -0.12 to 0.10). Therefore, the average outcome did not differ significantly from zero ($z = -0.21$, $p = 0.84$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 105 TVETBB1.

Figure 105: TVETBB1: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(3) = 22.60$, $p < 0.01$, $\hat{\tau}^2 = 0.01$, $I^2 = 86.72\%$). An examination of the studentised residuals revealed that one study (Gibbs 2020) had a value larger than ± 2.50 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

Exposure to intervention in months ($\hat{\beta} = 0.001, p = 0.81$ [95% CI: -0.01 to 0.01]) was not a significant moderator, but evaluation period in months was significant ($\hat{\beta} = -0.03, p = 0.02$ [95% CI: -0.05 to 0.01]) such that each additional month between the end of the evaluation and the data collection, the size of the effect decreased by .03 standard deviation units.

Several moderators could not be tested, including whether the model was adjusted for covariates, whether the programmed targeted women or a mix of women and men, whether the study was examining post-intervention changes or changes from baseline, GII score, FSI score and study design. We also could not examine risk of bias because three of the four studies were assessed as having a high risk of bias.

Effect of TVET on women participating more in their communities

Lombardini & Bowman's (2015) quasi-experimental study in Pakistan was the only study evaluating the impact of TVET on women participating more in their communities. There was a significant large positive effect ($g = 1.11$, [95% CI: 0.97 to 1.25]), but we assessed the study as having high risk of bias.

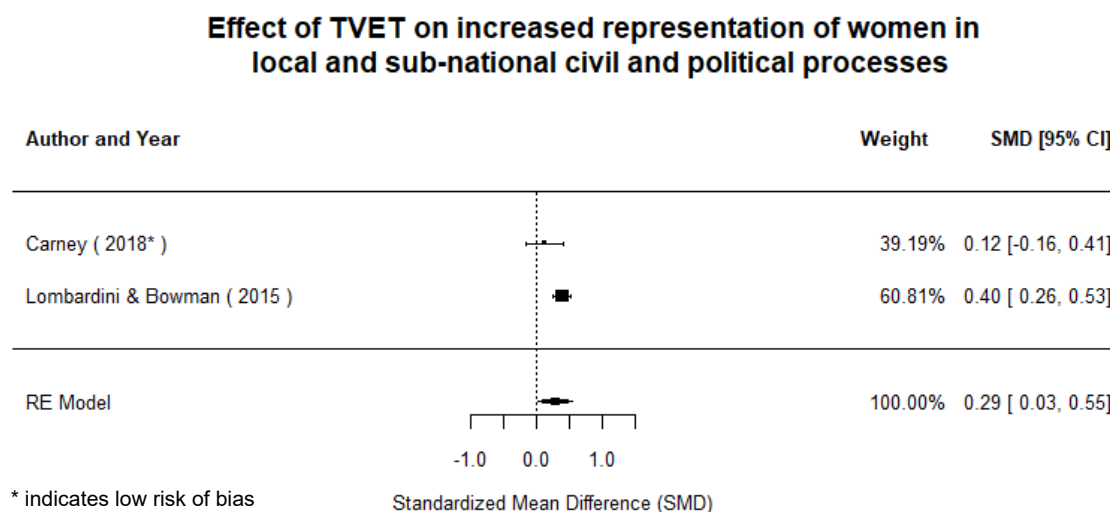
Effect of TVET on women forming their own organisations

Mckenzie and colleagues' (2019) experimental study in Kenya was the only study evaluating the impact of TVET on women forming their own organisations, reporting on women's associations at the 12- and 36-month follow-up. While there was a very small positive but non-significant estimate at the 12-month follow-up ($g = 0.01$, [95% CI: -0.08 to 0.10]), and there was a moderate positive and significant effect noted at the 36-month follow-up ($g = 0.15$, [95% CI: 0.06 to 0.24]). We assessed the study as having high risk of bias.

Effect of TVET on increased representation of women in local and sub-national civil and political processes

Only two studies reported the impact of TVET on increased representation of women in local and subnational civil and political processes, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as low risk of bias and the other as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.249$ (95% CI: 0.03 to 0.56). This positive effect did not differ significantly from zero ($z = 2.16$, $p = 0.03$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 106 TVETCA2. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 2.85$, $p = .09$, $\hat{\tau}^2 = 0.02$, $I^2 = 64.93\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid.

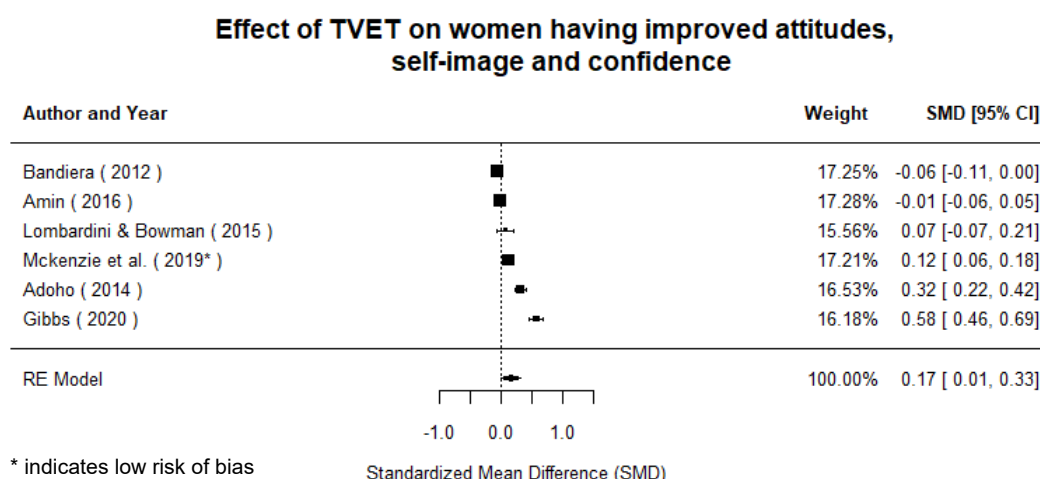
Figure 106: TVETCA2. Effect of TVET on women having improved attitudes, self-image and confidence



Effect of TVET on women having improved attitudes, self-image and confidence

We included a total of $k = 6$ studies in the analysis. We assessed one of the studies as low risk of bias, one as some concerns, and four as high risk of bias. The observed outcomes ranged from -0.06 to 0.58 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.17$ (95% CI: 0.01 to 0.33). Therefore, the average outcome differed significantly from zero ($z = 2.07$, $p = 0.04$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 107 TVETCB3.

Figure 107: TVETCB3: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(5) = 129.22$, $p < 0.01$, $\hat{\tau}^2 = 0.04$, $I^2 = 96.13\%$). An examination of the studentised residuals revealed that one study (Gibbs 2020) had a value larger than ± 2.64 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

We examined several moderators in the context of this model. Effects from high risk of bias studies were not different than effects from low or some concerns related to risk of bias ($\hat{B} = 0.17, p = 0.43$ [95% CI: -0.25 to 0.54]). Exposure to intervention in months ($\hat{B} = -0.01, p = 0.40$ [95% CI: -0.02 to 0.01]) was not a significant moderator, nor was evaluation period in months ($\hat{B} = 0.02, p = 0.20$ [95% CI: -0.01 to 0.06]). Effects from experimental studies were not different than effects from quasi-experimental studies ($\hat{B} = 0.04, p = 0.85$ [95% CI: -0.40 to 0.49]). Effects also did not differ by scores on the gender equality index ($\hat{B} = 2.8, p = 0.08$ [95% CI: -0.29 to $0.5.9$]), or the fragile states index ($\hat{B} = -0.01, p = 0.70$ [95% CI: -0.04 to 0.03]).

Several moderators could not be tested, including whether the model was adjusted for covariates, whether the programmed targeted women or a mix of women and men and whether the study was examining post-intervention changes or changes from baseline.

Effect of TVET on decreased violence and discrimination at the household level
Lombardini & Bowman's (2015) quasi-experimental study in Pakistan was the only study evaluating the impact of TVET on decreased violence and discrimination at the household level. There was a small positive point estimate ($g = 0.09$, [95% CI: -0.05 to 0.22]) that was non-significant, and we assessed the study as having high risk of bias.

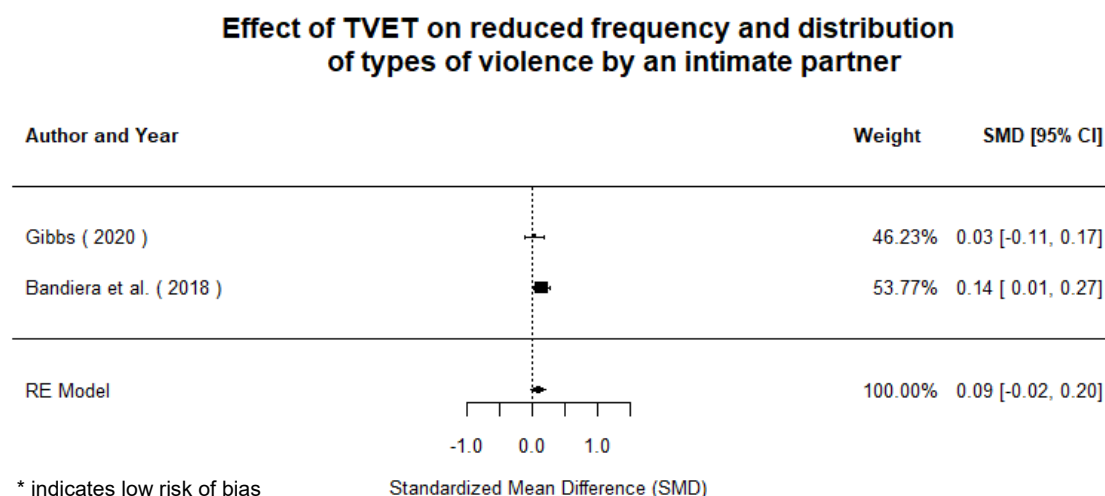
Effect of TVET on safer and more secure households, communities and territories for women and girls

Amin's (2016) experimental study in Bangladesh was the only study evaluating the impact of TVET on safer and more secure households, communities and territories for women and girls. They included nine effects, three for teach treatment arm, that looked at harassment at home, outside the home, or in school/class. The point estimates ranged from very small, negative and non-significant ($g = -0.002$, [95% CI: -0.06 , 0.05]) to small, positive and significant effects ($g = 0.11$, [95% CI: 0.06 to 0.17]). We assessed the study as having some concerns related to risk of bias.

Effect of TVET on reduced frequency of intimate partner violence

Only two studies reported the impact of TVET on the reduced frequency of intimate partner violence, thus we included $k = 2$ studies in the analysis. We assessed both of the studies as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.09$ (95% CI: -0.02 to 0.20). This positive effect did not differ significantly from zero ($z = 1.62, p = 0.10$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 108 TVETCB7. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 1.32, p = .25, \hat{\tau}^2 = 0.002, I^2 = 24.26\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid.

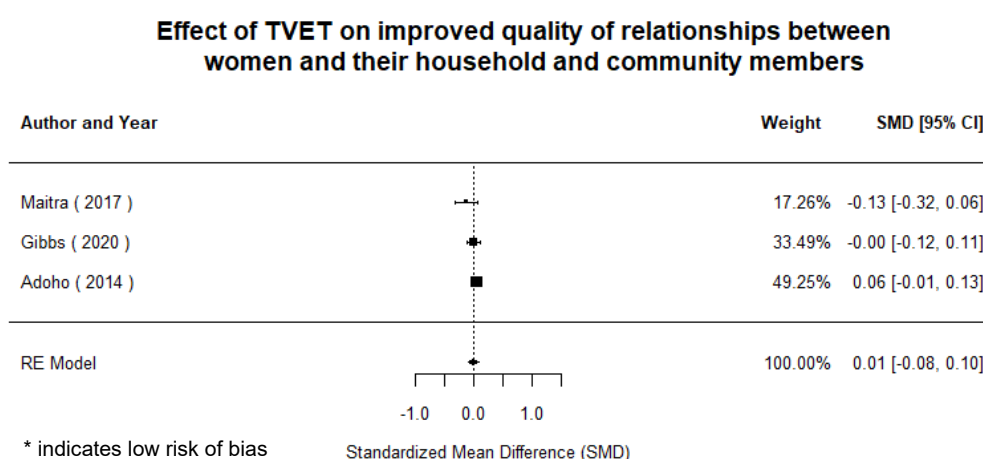
Figure 108: TVETCB7: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effect of TVET on improved quality of relationships between women and their household and community members

We included a total of $k = 3$ studies in the analysis. We assessed none of the studies as low risk of bias, one as some concerns, and two as high risk of bias. The observed outcomes ranged from -0.13 to 0.06 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.01$ (95% CI: -0.08 to 0.10). Therefore, the average outcome did not differ significantly from zero ($z = 0.14$, $p = 0.89$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 109 TVETCB8.

Figure 109: TVETCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model

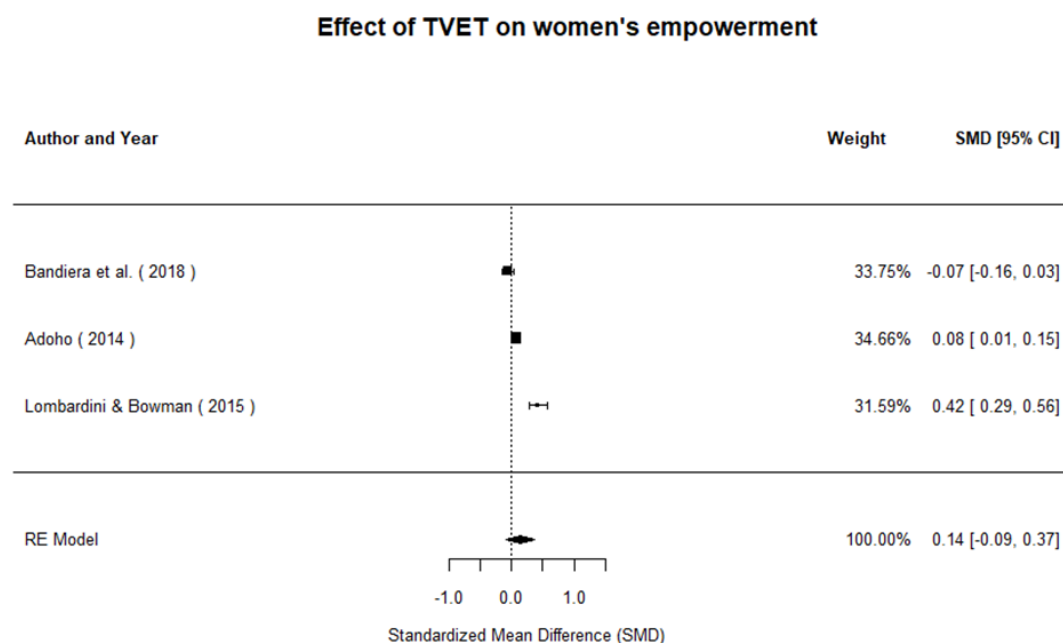


According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(2) = 3.68$, $p = 0.16$, $\hat{\tau}^2 = 0.00$, $I^2 = 45.70\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.39 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

Effect of TVET on women's empowerment index

We included a total of $k = 3$ studies in the analysis. We assessed all of the studies as high risk of bias. The observed outcomes ranged from -0.07 to 0.42 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.14$ (95% CI: -0.09 to 0.37). Therefore, the average outcome did not differ significantly from zero ($z = 1.19$, $p = 0.24$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 110 TVETCC1.

Figure 110: TVETCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(2) = 33.11$, $p < 0.01$, $\hat{\tau}^2 = 0.04$, $I^2 = 93.96\%$). An examination of the studentised residuals revealed that one study (Lombardini and Bowman 2015) had a value larger than ± 2.50 and may be a potential outlier in the context of this model. Indeed sensitivity analyses leaving each study out indicated that removing Lombardini and Bowman (2015) would reduce the overall average effect ($\hat{\mu} = 0.01$ (95% CI: -0.13 to 0.15), but the effect would still be positive and non-significant ($z = 0.11$, $p = .91$). According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

Qualitative findings

We conducted a thematic synthesis on the five linked qualitative studies to the included TVET interventions. As indicated above, this thematic synthesis aims to identify themes related to the interplay of intervention design, intervention implementation, target population, and contextual variables with intervention outcomes and effects. In total, we identified 18 descriptive themes, which we configured into 4 analytical themes (Online Appendix N). These four analytical themes present the synthesis results and are discussed in more detail below.

Theme 1: Women-only TVET programmes are preferred by participants

The qualitative evidence indicated a consistent theme of women preferring single-sex TVET programme approaches (Stangl et al 2015; McKenzie et al 2019; Bandiera et al 2018; Subah-Belleh Associates 2010). All three included interventions—the Economic Empowerment of Adolescent Girls and Young Women (EPAG), the Empowerment and Livelihood for Adolescents (ELA) programme in Sierra Leone, and the GET Ahead for Women in Enterprise in Kenya were designed to only include women participants and men were not eligible for participation. In light of this, all qualitative data is based on women’s experiences in single-sex programmes. Reasons expressed for this preference by women for single-sex TVET programmes included a lack of space for open engagement, safety concerns, and need for peer-to-peer exchanges and collective-building. Demand was expressed for ‘men to see the benefits’ of the programmes, but not for men to be eligible to join them.

Women in the Kenyan GET Ahead for Women in Enterprise programme felt that the inclusion of men would reduce the space for open engagement within the training with participants expressing that *‘I am not of the opinion of including men because some of the women like myself will not be free to share their ideas in the presence of the men. (...) It’s better to be women alone because if we were mixed the men would have sold out what the women converse in the training outside and hence, they would be viewed differently’* (Stangl et al 2015). This fear of negative repercussions against women was reported in the programme evaluation which further recorded incidents of men and community members challenging women about their late return home and status gained from attending the training. Women expressed an explicit fear that should men be included ‘women entrepreneurs would again be forced into domestic responsibilities instead of being able to focus on growing their businesses’ (Stangl et al 2015).

In addition, a further theme in favour of women-only TVET programmes focused on the role of peer-to-peer interaction and the building of a collective. Kenyan entrepreneurs in particular valued the ‘camaraderie, support and mentorship they received from one another’ (in addition to the freedom to express themselves without the censure of men) (Stangl et al 2015). The ELA programme in Sierra Leone focused on a similar pathway to change aiming to create women-only clubs that offer a space where young women can gather and socialise, and thus provide an alternative to spending time with men or other leisure activities (Bandiera et al 2018). The essence of this peer-to-peer focus and emphasis on a safe collective was perhaps best captured in a Kenyan’s entrepreneurs’ expression that “Having the programme with women only helps since we put our thoughts together. For men, the only goal is to get money. As for women, we help one another move to a better status” (Stangl et al 2015).

However, there was also dissenting evidence on this theme. In both Liberia and Kenya, trainers and implementers advocated for a mixed-sex programme approach arguing that the involvement of women’s partners is required to change household dynamics and that the role of men in changing structural gender norms cannot be overlooked. A small minority of women in the Kenyan GET Ahead for Women in Enterprise programme agreed with this rationale but the overall opinion of the included women was summarised in the evaluation as the expression of ‘men will spoil it’ (Stangl et al 2015).

Theme 2: Enhancing social capital and peer-to-peer exchange is a contributing factor, if designed for within TVET

A second set of themes identified in the qualitative evidence base expands on the above reported potential of women's peer-to-peer exchange and collective building. In addition to the benefits of gendered exchanges aiming to build support structures against restrictive social environments and norms, the qualitative evidence also points to the potential of peer-to-peer design elements within TVET programmes more broadly. TVET programme designs that emphasised 'collective-building' and 'peer-to-peer' exchanges such as mentoring and the embedding within social clubs or women's self-help groups reported an enhanced social capital and networks of women. This social capital and networks in return were seen as contributing to creating a space for the application of skills in practice (e.g., opening a business, finding a new job) which enhanced the potential and design of the TVET programmes.

Liberia's EPAG programme deliberately designed a pairing/small group approach between girls aiming to improve girls' skill development as well as to promote the creation of social capital, such as friendships, mutual trust and support networks among the girls (Subah-Belleh Associates 2010). These groups were paired deliberately consisting of 3-4 girls to encourage after school interaction and relationship building considering location and a mix of skill levels. The programme evaluation found this design element to be appreciated by the girl participants who reported helping each other inside and outside of the classroom activities. As one participant observed *'There were girls that could not write and felt ashamed to take part ... these girls were assisted by the others who could write ... and urged them to come to school regularly. With this persuasion and encouragement, we were able to overcome this challenge.'* The evaluation report further links these generated networks of girls to exit polls results describing the girl's newfound optimism and belief in themselves, with one trainer noting, *'now they look like sisters'* (Subah-Belleh Associates 2010).

Similar reports on the value and benefit of mentoring for business skills were found in the Kenyan GET Ahead for Women in Enterprise programme. In this context, mentoring was seen as a relevant design element to enhance business skill transfer and development by contextualising these skills in the experience of an established entrepreneur. Finally, in the ELA programme in Sierra Leone embedding TVET components within existing women groups and collectives proved a feasible design strategy. This design strategy provided a direct link to the theme 1 discussed above.

Theme 3: TVET and its design benefits from the inclusion of and coupling with LSCB programmes

The qualitative evidence on programme design suggests an integration of TVET and LSCB capacity building programmes to be of promise. The assessments of stand-alone TVET programmes and TVET components within LSCB programmes emphasise the interconnectedness of both types of skill sets; and programme participants, too, express a demand and need for both types of skills in order to enhance their life capabilities and opportunities. Within TVET programmes participants cited obstacles to the application of these skills to further their life being constrained by personal and structural issues that LSCB programmes focus on; vice-versa, within LSCB programmes, participants expressed a need for practical applications as well as a pathway for how new skills could lead to economic opportunities.

In Liberia's EGAP programmes, qualitative data and feedback from trainers and implementers highlighted the need to enhance the training's 'social aspect' targeting outcomes such as the girls' self-esteem and social competencies, which fall in the design realm of LSCB programmes rather (Subah-Belleh Associates 2010). This sentiment was echoed by the programme's exit poll data which confirmed girls' desire and appreciation for more of the life skills-related topics. Similar feedback emerged from the TVET programmes in Kenya and Sierra Leone with participants emphasising the need for practical opportunities to apply skills in real world situations, including life skills competencies rather than only technical skills (Stangl et al 2015; Bandiera et al 2018).

This more general theme of wanting to apply gained skills in daily contexts seemed to reflect a desire for practical tools and applications to further one's capabilities, personal development and growth. In combination with the qualitative data on LSCB programmes, which mirror this theme, this points to the direction that capacity building design can benefit from intervention packages with multiple components aimed at structural change. Within the qualitative data on TVET programmes, there is indication that this is indeed feasible from a design perspective and promising in practice. All three included TVET programmes referenced a deliberate multicomponent or multi-faceted design approach to support women's empowerment beyond the mere skill building component:

- The EGAP in Liberia reported four implemented performance enhancing strategies— the pairing/small group approach, mentoring, childcare services, and transport/meal allowance—which are recommended as a design package for future interventions.
- The ELA programme in Sierra Leone referred to its multifaceted approach to simultaneously tackle multiple disadvantages young women face, related to having agency over their bodies and barriers to accumulating human capital. This approach combines the provision of life skills, vocational skills and microfinance, delivered within a social club and safe space for women.
- The GET Ahead for Women in Enterprise programme in Kenya described itself as a management training programme with a strong gender component; that is, the programme is designed specifically with a gender lens throughout and aims to address both the practical and strategic needs of low-income women in enterprise by strengthening their basic business and people management skill as well as supporting networking and relationship building.

Theme 4: A bundle of key implementation considerations apply across contexts

Across the TVET programmes a number of implementation considerations were reported repeatedly as contributing factors to programme success. These implementation considerations are fairly programmatic and granular. They are reported below as a bundle of considerations suggested across individual studies along the implementation cycle:

- Design and implement a tailored and sensitive awareness-raising and outreach campaign to attract participants into the programme. For example, female entrepreneurs in the GET Ahead for Women in Enterprise mentioned the sensitive and respectful approach of recruiters that gave them a positive first impression of the programme including that they were approached directly at their own place of business. As part of this programme, implementers further ran an invitation choice experiment to test the design of three different intervention types. The results of this trial were, however, inconclusive.

- Identify and select a central and safe location that is easily accessible (e.g., via public transport, within walking distance of participants). In Kenya, some women part of the GET Ahead for Women in Enterprise had to leave their homes as early as 3:00am in order to reach the training venue via public transport; a fact which undermined participation and satisfaction. Programmes in Sierra Leone and Liberia mitigated this issue by identifying and renting facilities within villages and central transport hubs.
- Provision of an allowance to cover transport and possibly offset opportunity costs of attending. This implementation criterion was regarded as the most important aspect by women participants in the GET Ahead for Women in Enterprise in Kenya. In Liberia, too, female trainees rated the provided allowance as a key incentive and enabler for attendance in the EPAG programme.
- Provision of childcare services at the training venue. This implementation criterion featured prominently within the Kenyan and Liberian TVET programmes and was equally regarded as a strong enabler of attendance by female participants
- Training curricula and materials require careful adaptation to local context and translation into local languages. Participants of both the Liberian EGAP and Kenyan GET Ahead for Women in Enterprise complained about the use of foreign languages and irrelevant examples and contexts in the provided materials.
- The skilfulness and approach of programme trainers was reported as a key implementation element and driver of programme satisfaction across the included studies.

The above present a list of programmatic implementation factors reported in the qualitative evidence base. Within these the Liberian EGAP and Kenyan GET Ahead for Women in Enterprise made deliberate attempts to mitigate the gendered nature of some of these factors. This included the use of gender-sensitive outreach and communication materials in both programmes; the provision of childcare in both programmes; the selection of the programme location for women's access in Liberia; and the adaptation and design of a gender-specific curriculum in Kenya. Arguable, it should be feasible to apply a gendered lens to all identified implementation considerations in future programme design.

Discussion

The quantitative meta-analysis of TVET programmes revealed positive significant effects on the capacity of women entrepreneurs; women's increased access to and ownership of assets, credit, and income; and women having improved attitudes, self-image, and confidence. The remainder of outcomes were non-significant. This included empowerment related outcomes such as women having more control over their bodies; women's freedom of movement and association; women's participation in decision-making at the household and community level and direct measures of women's empowerment, which were all found to be non-significant.

A complementary qualitative synthesis was carried out to explore explanations and recommendations that may help design more effective TVET interventions that can go beyond narrow outcomes on skills development and positive attitudes targeting broader empowerment objectives:

- The qualitative evidence base indicated a need to design additional programmes components within TVET in order to target outcomes beyond individual skill

development and improved attitudes. In particular, TVET programmes were designed as women-only spaces—a design feature strongly preferred by women but flagged by implementers and decision-makers as a potential barrier for programmes to translate into broader empowerment outcomes. Not involving men by design in programmes might exclude a pathway for change and negate the translation and application of skills development into progress on household power relations and women’s empowerment within communities.

- A design feature suggested by the qualitative synthesis which might support a translation of skills development outcomes refers to the use of mentoring and ‘collective-building’ formats such as women’s self-help groups and clubs. These programme components can allow women to engage in peer-to-peer exchanges and to form relationships and networks that allow support to each other beyond the technical skills development. These components thereby add a programme dimension more aligned with targeting broader empowerment outcomes and creating multiple pathways for change.
- Overall, the qualitative evidence contextualised with the quantitative results does indicate that, unless designed for specifically within the programme composition and implementation, outcomes beyond the direct objective of TVET interventions (i.e., skills) are difficult to achieve; in particular if they depend on changing entrenched social norms and practices.

Summary of findings and discussion

We included 10 studies in eight countries in South Asia and Sub-Saharan Africa. We were able to examine effects on the following outcomes: Women have increased access and ownership to assets, credit and income, women can access decent work (formal and informal employment), improved capacity of women entrepreneurs, increased participation in decision making by women at the household or community level, including during crisis response, women have improved attitudes, self-image and confidence. Our included studies report against all secondary outcomes (Resources, Agency and Achievement) and seven immediate outcomes for our review. Overall, the GRADE assessments generally indicate a range of very low to high certainty in this body of evidence.

The 10 quantitative studies were further detailed by five linked qualitative studies on the impact of TVET programmes. The meta-analysis results align with the qualitative results; both suggest that women-only TVET programmes have a higher positive and significant impact. Additionally, qualitative findings suggest that bundling components together, such as the training and sensitisation, can result in more positive and long-lasting outcomes. Designing programmes with a peer exchange or social skills component also holds potential to bolster positive outcomes; girls or women are able to apply skills in practice and with the support of their peers. Table 39 below presents the GRADE review of our findings:

Table 39: GRADE summary of findings and certainty of evidence on TVET

Certainty assessment							Sample size	Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(AB1) Increased capacity of women to understand and use financial, banking, and business services effectively										
1	RCT	not serious	serious ^a	not serious	not serious	none	2400	SMD 0.09 SD higher (0.03 lower to 0.15 higher)	⊕⊕⊕○ MODERATE	Important, but not critical
(AB2) Women have increased access and ownership to assets, credit and income.										
10	RCT - 10	very serious ^b	very serious ^c	not serious	not serious	publication bias strongly suspected	7181	SMD 0.06 SD higher (0.01 higher to 0.11 higher)	⊕○○○ VERY LOW	Important, but not critical
(AC1) More women engaged in other micro, small, and medium-sized enterprises.										
1	RCT	not serious	serious ^a	not serious ^d	serious ^e	none	5398	Two small, positive effect estimates with a 95% CI range of -0.05 to 0.19	⊕⊕○○ LOW	Important, but not critical
(AC2) Initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities										
7	RCT - 6 QED - 1	serious ^f	very serious ^g	not serious	serious ^h	none	91564	SMD 0.07 SD higher (0.07 lower to 0.21 higher)	⊕○○○ VERY LOW	Important, but not critical
(AC3) Improved capacity of women entrepreneurs										
5	RCT - 5	very serious ⁱ	not serious	not serious	not serious	none	34761	SMD 0.14 SD higher (0.09 higher to 0.19 higher)	⊕⊕○○ LOW	Important, but not critical
(BA1) Women have improved success in the workplace.										
1	RCT	very serious ^j	serious ^a	not serious	not serious	none	2474	SMD 0.23 SD higher (0.15 higher to 0.31 higher)	⊕○○○ VERY LOW	Important, but not critical
(BA2) Women have more and better control over their bodies and sexual health.										
2	RCT - 1 QED - 1	serious ^p	very serious ^q	not serious	serious ^r	none	21941	SMD 0.26 SD higher (0.36 lower to 0.88 higher)	⊕○○○ VERY LOW	Critical

Certainty assessment							Sample size	Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(BA3) Women have increased freedom of movement and association.										
5	RCT - 4 QED - 1	very serious ^s	not serious	not serious	not serious	publication bias strongly suspected	12015	SMD 0.04 SD higher (0.02 lower to 0.11 higher)	⊕○○○ VERY LOW	Limited Importance
(BA5) Women have more positive attitude towards taking action to claim their rights.										
1	RCT	very serious ^w	serious ^a	not serious	not serious	none	1210	SMD 0.03 SD higher (0.08 lower to 0.14 higher)	⊕○○○ VERY LOW	Limited Importance
(BA6) Reduced percentage of women agreeing with certain reasons that justify violence against women and girls										
2	RCT - 1 QED - 1	very serious ^x	not serious	not serious	not serious	none	13870	SMD 0.03 SD lower (0.13 lower to 0.06 higher)	⊕⊕○○ LOW	Critical
(BA7) Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them.										
3	RCT - 2 QED - 1	very serious ^y	not serious	not serious	not serious	none	1558	SMD 0.04 SD higher (0.01 lower to 0.08 higher)	⊕⊕○○ LOW	Important, but not critical
(BA8) Reduced instances of child or forced marriage										
2	RCT - 2	very serious ^y	not serious	not serious	not serious	none	15424	SMD 0.17 SD higher (0 to 0.34 higher)	⊕⊕○○ LOW	Critical
(BB1) Increased participation in decision making by Women at the household or community level, including during crisis response										
4	RCT - 2 QED - 2	very serious ^z	very serious ^{aa}	not serious	serious ^{ab}	none	9763	SMD 0.01 SD lower (0.12 lower to 0.1 higher)	⊕○○○ VERY LOW	Limited Importance
(BB2) Women participate more in their community.										
1	QED	serious ^{ac}	serious ^a	not serious	not serious	very strong association	804	SMD 1.11 SD higher (0.97 higher to 1.25 higher)	⊕⊕⊕⊕ HIGH	Limited Importance
(BB4) Women increasingly form their own organisations.										

Certainty assessment							Sample size	Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
1	RCT	not serious	serious ^a	not serious	serious ^{ad}	none	5400	0 (0 to 0)	⊕⊕○○ LOW	Limited Importance
(CA2) Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration										
2	QED - 2	very serious ^{ae}	serious ^{af}	not serious	not serious	none	24735	SMD 0.29 SD higher (0.03 higher to 0.55 higher)	⊕○○○ VERY LOW	Limited Importance
(CA8) Increased community support for women's and children's human, economic and legal rights										
2	RCT 1 QED 1	very serious ^x	serious ^{ag}	not serious	not serious	none	992	SMD 0.07 SD higher (0.04 higher to 0.18 higher)	⊕○○○ VERY LOW	Important, but not critical
(CB3) Women have improved attitudes, self-image and confidence.										
6	RCT 5 QED 1	very serious ^{ah}	very serious ^{ai}	not serious	not serious	none	3993	SMD 0.17 SD higher (0.01 higher to 0.33 higher)	⊕○○○ VERY LOW	Limited Importance
(CB5) Decreased violence/discrimination at the household level										
1	QED	serious ^{aj}	serious ^a	not serious	serious ^{ak}	none	804	SMD 0.09 SD higher (0.05 lower to 0.22 higher)	⊕○○○ VERY LOW	Critical
(CB6) Safer and more secure household, communities and areas/territories for women, girls, men, and boys										
1	RCT	serious ^{al}	serious ^a	not serious	not serious	none	5046	0 (0 to 0)	⊕⊕○○ LOW	Critical
(CB7) Reduced frequency and distribution of types of violence by an intimate partner										
2	RCT - 2	very serious ⁱ	not serious	not serious	serious ^k	none	25864	SMD 0.09 SD higher (0.02 lower to 0.2 higher)	⊕○○○ VERY LOW	Critical
(CB8) Improved quality of relationships between women and their household and community members										
3	RCT - 3	very serious ^m	serious ⁿ	not serious	serious ^o	none	2602	SMD 0.01 SD higher (0.08 lower to 0.1 higher)	⊕○○○ VERY LOW	Limited Importance

Certainty assessment							Sample size	Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(CC1) Empowerment/Equality index										
4	RCT – 3 QED - 1	very serious ^t	very serious ^u	not serious	serious ^v	none	17030	SMD 0.13 SD higher (0.01 lower to 0.27 higher)	⊕○○○ VERY LOW	Important, but not critical

Note. CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

- a. All single studies downgraded once for inconsistency.
- b. Downgraded twice since although only one of 12 studies is not a RCT 5 of them are high risk of bias and 3 raise some concerns related to risk of bias
- c. Downgraded twice because studies are on both side of the threshold and CI are not overlapping. The $I^2 = 80.04\%$ and the Q-test concludes that the true outcomes appear to be heterogeneous
- d. Downgraded because there are point estimates on either side of the threshold with a wide berth of CIs.
- e. Downgraded because there are point estimates on either side of the threshold with a wide berth of CIs.
- f. Downgraded because three of seven included studies present high risk of bias cross various domain criteria, and an additional two are of some concerns related to risk of bias. Deviation from intended intervention is of particular concern.
- g. Downgraded twice because this group of studies estimates a wide range of effect sizes, with limited CI overlap and both positive and negative point estimates.
- h. While the CI are reasonably small, the CI cross the threshold, and the small point estimate is not representative of the range of estimates included in the meta-analysis.
- i. Downgraded twice because half the body of evidence comes from high risk of bias studies, and an additional two of six are from medium or uncertain risk of bias.
- j. Downgraded three times 1. Only two studies and both of them are high risk of bias 2. High risk on three RoB domains 3. Removing either study would impact meta-analysis result.
- k. Downgraded once because although sample size > 400 $p=0.10$.
- l. A high risk of bias was found in the unit of analysis criterion, and some concerns related to risk of bias in the selection and performance criteria.
- m. Downgraded two times 1. The three studies are RCTs but two of them are high risk and one raises some concerns related to risk of bias 2. Removing either of the high-risk studies would impact meta-analysis.
- n. Downgraded once 1. Effects sizes are on both side of the threshold and CI range is wide and not systematically overlapping 2. $I^2 = 45.70\%$ but according to q-test there was no significant amount of heterogeneity.
- o. Downgraded once because although sample size > 400 $p=0.89$.
- p. Downgraded to reflect the high and some concerns related to risk of bias studies in this group. There is particular concern with regards to deviation from intended intervention. The high risk of bias study strongly influences the final effect estimate.
- q. There is considerable variation in the effect estimates on both sides of the threshold. The difference is so large that the wide CIs do not cross.
- r. Downgraded due to wide CI that cross the threshold of interest widely on both sides.

- s. Downgraded twice because three of five studies in this group are of high risk for bias, and another is of some concerns related to risk of bias There is particular concern for selection bias and performance bias.
- t. Downgraded three times 1. Mix of RCTs and QED 2. All studies are high risk of bias covering five RoB domains 3. Removing either study would impact the meta-analysis
- u. Downgraded twice 1. we have effect sizes on both side of the threshold and a wide range of CI not systematically overlapping. 2. $I^2 = 91.17\%$ and according to Q-test the true outcomes appear to be heterogenous
- v. Downgraded once because although sample size > 400 $p=0.07$.
- w. High risk biases were found in the unit of analysis and selection bias criteria. The deviations from intended interventions and outcome measurement bias criteria were unclear.
- x. Downgraded twice because all evidence in this group comes from high risk of bias studies.
- y. Downgraded because half of the evidence comes from high risk of bias studies while the other half is uncertain.
- z. Downgraded 3 times 1. Mix of QED and RCTs 2. Three of the four studies are high risk of bias covering three different RoB domains 3. Removing either study would impact the results of the meta-analysis.
- aa. Downgraded twice 1. We have effect sizes on both side of the threshold and a wide range of CI not systematically overlapping. 2. $I^2 = 86.72\%$ and according to Q-test the true outcomes appear to be heterogenous.
- ab. Downgraded once because $p = 0.84$ although sample size > 400.
- ac. BB2 Women participate more in their community.
- ad. Downgraded because point estimates vary considerably, and CIs cross widely into both sides of the threshold.
- ae. Downgraded three times 1. The two studies are QED 2. One of the studies is high risk of bias and the other one is low risk 3. Removing either study would affect the meta-analysis.
- af. Downgraded once as the effect sizes are on the same side of the threshold and the CI overlap but $I^2 = 64.93\%$ indicates risk of heterogeneity.
- ag. Downgraded because of varied point estimates and CI that cross both sides of the threshold.
- ah. Downgraded three times 1. Although most of the studies are RCTs only one of them is low risk of bias and one raises some concerns related to risk of bias 2. Four studies out of six are high risk of bias covering three RoB domains 3. Removing any of the high risk would impact the results of the meta-analysis.
- ai. Downgraded twice 1. Effects sizes are on both side of the threshold and CI range is wide and not systematically overlapping 2. $I^2 = 96.13\%$ and Q-test concludes that the true outcomes appear to be heterogeneous.
- aj. some concerns related to risk of bias of reporting bias, and unclear deviations from intended interventions and outcome measurement bias.
- ak. Downgraded due to wide range of CIs.
- al. some concerns related to risk of bias was found in the unit of analysis criterion. The deviations from intended interventions and outcome measurement bias criteria were unclear.

4.4 Effects of interventions under the UNSCR protection pillar

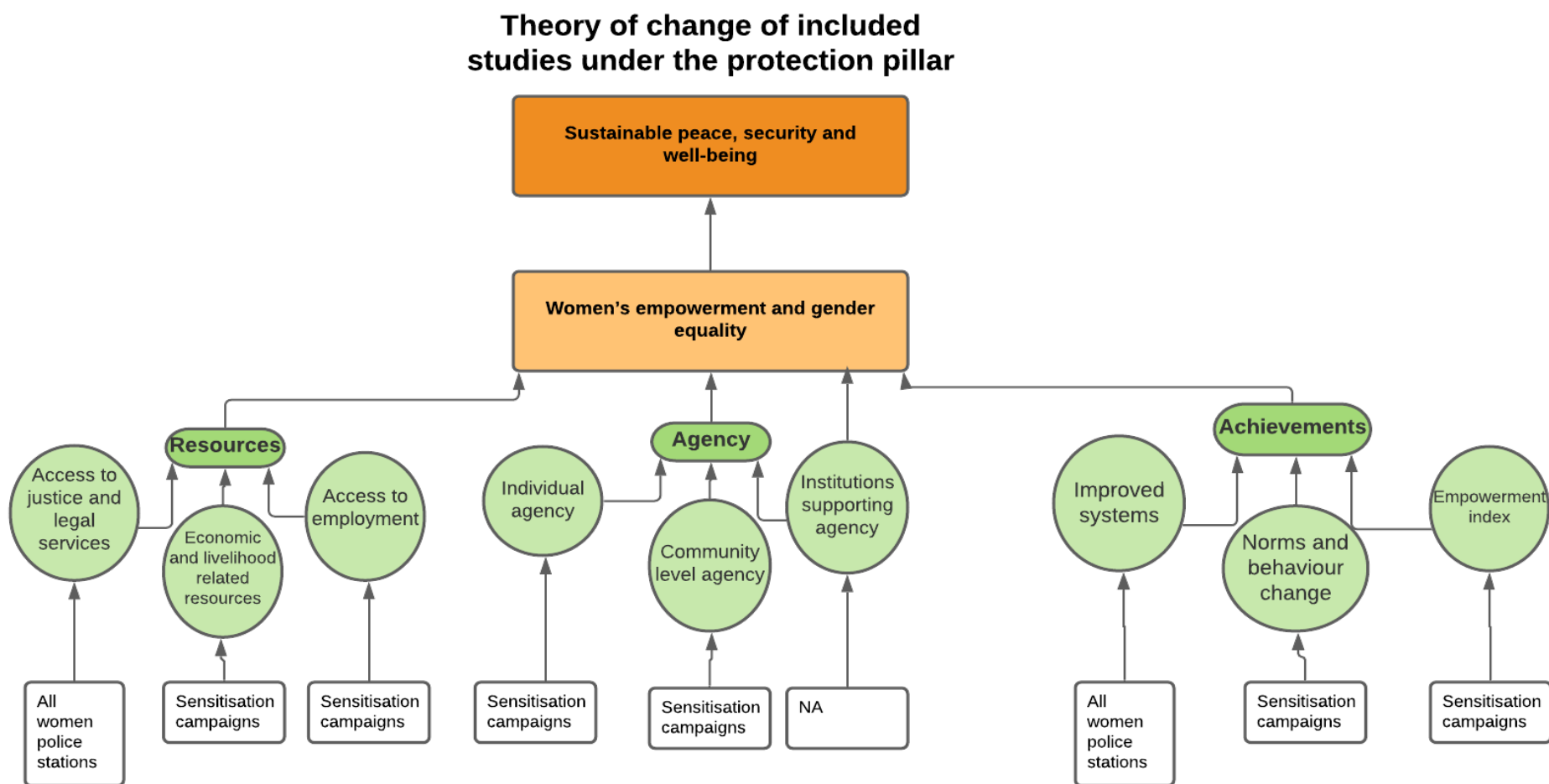
The protection pillar includes a set of interventions that create, facilitate access to or build awareness of and support for legal or social protections for women's and girls' rights. This also includes behavioural, legal and environmental interventions that aim to reduce women and girls' risk of experiencing SGBV.

In our systematic review, this pillar gathered the following types of interventions covering 8,306 beneficiaries:

- All women police stations
- Sensitisation campaigns

This section provides the findings of our synthesis of the 10 included studies evaluating the effect of these interventions on gender equality, women's empowerment and peaceful and inclusive societies. The section is organised by each intervention group. Each sub-section begins with a description of the intervention groups, their activities and theory of change, followed by descriptive results and the findings addressing our research questions. Figure 111 below provides a summary of the outcomes and targeted effects of the interventions included under the protection pillar:

Figure 111: Theory of change for included studies under the protection pillar



4.4.1 All women police stations

All women police stations are when policewomen can be posted at a standard station or all women station, the main distinction being that the latter is run entirely by women staff. All women stations are additive; a victim may register a crime in the standard police station near where the crime occurred or to the all-women station in the district headquarters. This form of policing provides safe spaces for women willing to use formal support services. They typically employ women officers purposely trained to handle gender-based violence crimes (Amaral et al. 2019). Other services such as psychological, social and legal advice, education and prevention or mediation can be offered in those stations (Perova and Reynolds 2017). Table 40 below presents the characteristics of the included studies.

How do all women police stations affect gender equality, women's empowerment and peace outcomes?

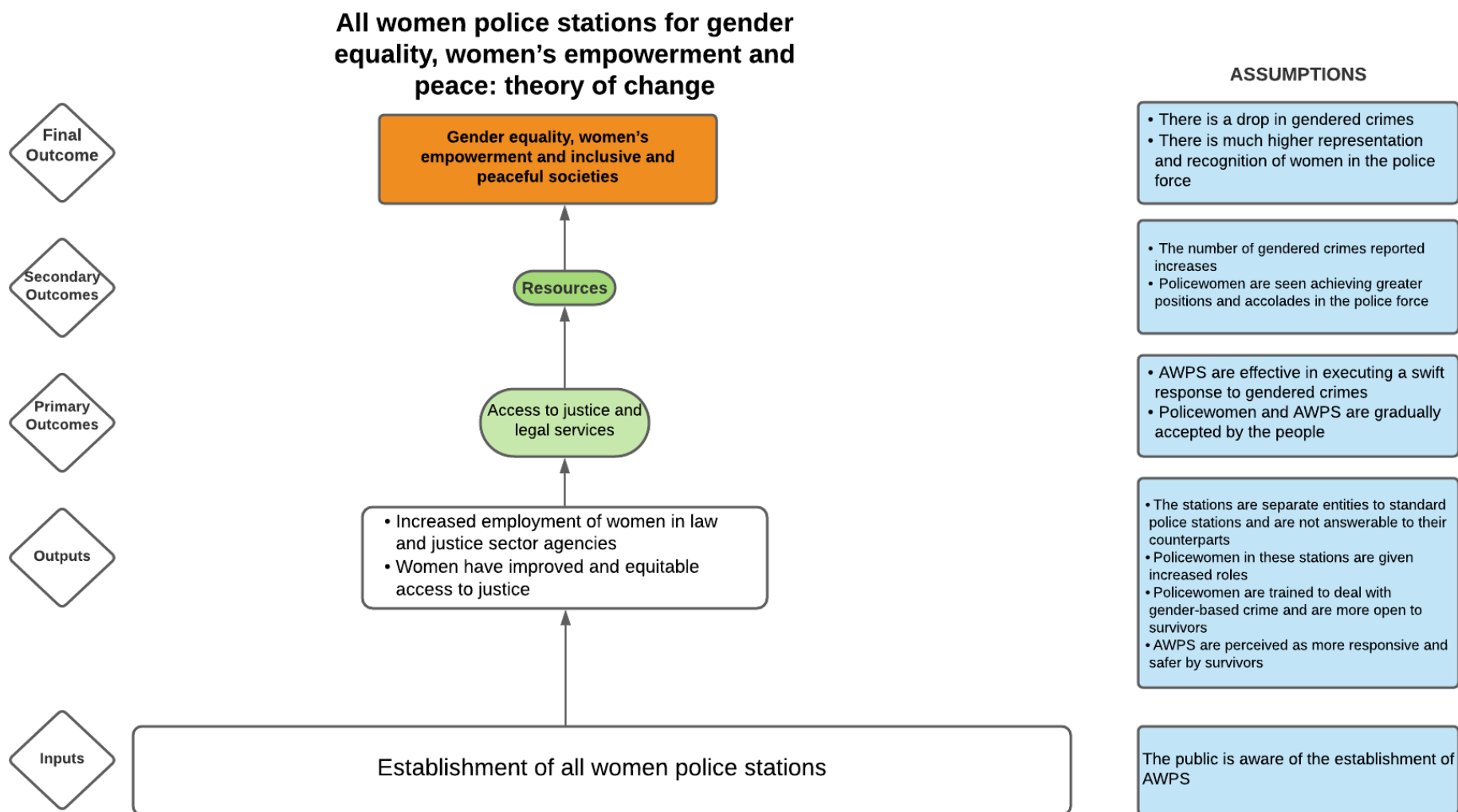
Figure 112 below maps out the causal chain of how all women police stations may improve gender equality, women's empowerment and peace outcomes.

All women police stations create gender equitable societies through two major pathways.

1) Safe spaces for women: All women police stations lower the cost of reporting crimes for women as they offer a more hospitable environment with women officers that are typically trained to handle gendered crimes and are separate from traditional police stations that are generally corrupt, dominated by men and are perceived to have a greater stigma against gendered crimes. 2) Representation: All women police stations promise to increase responsibilities of women in the police force and gradually remove previously held prejudices that may be hindering their advancement.

Our review only included one study (Jassal 2020) that explored the causal effects of AWPS. The study refuted the idea of AWPS leading to gender inclusive outcomes and argued for the opposite. They distinguish *inclusive* mechanisms of representation (e.g. quotas) from representation mechanisms that employ *separation*, such as the AWPS. They posit that AWPS offer segregation rather than integration and hence, hinder women's advancement in the police force and allow standard police stations to pass on gender-based crimes to women police stations that are relatively less widespread, increasing the cost of reporting for the victim.

Figure 112: All Women Police Stations for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included one study reported in one paper that evaluated one programme.

Population

All women police stations, by virtue of the concept, targeted women and girls. They comprise an all women staff and specialise in gender-based crimes.

The included studies evaluated one programme and trial in India. The low number of studies including a focus on all women police stations is mainly because this is a type of intervention primarily implemented in Latin America, where less countries met the inclusion criteria for our review.

Intervention, inputs and activities

The included study (Table 40) evaluated the following activity:

- All Women Police Stations (n = 1): Establishment of all women police stations in Haryana, India.

Table 40: AWPS features of included studies

Study	Activity/input	Length of treatment	Intervention frequency
Jassal - 2020	Establishment of all women police stations in the state of Haryana	-	-

Comparison

The included study compared treated groups to comparison groups receiving no intervention. The study did not include multiple treatment arms.

Outcomes

The included study reported on a number of relevant outcomes, including the following:

- Increased employment of women in law and justice sector agencies (n = 1): Increased recruitment practices for women, non-discrimination, more women police officers, judges, lawyers, community legal volunteers, etc. Through this employment, Women have an increased influence in this sector.
- Women have improved and equitable access to justice (n = 1): Women are able to file cases of injustice and raise cases, receive information, have judicial support and address offenses. Women's cases are not dropped, meaning that cases (especially relating to SGBV or women's rights, economic or otherwise) are followed up at all levels. This could mean the creation or editing of a referral pathway or improved communication mechanisms between actors and service providers

The division of the immediate and secondary outcomes in asset transfer is reported in Table 41 below:

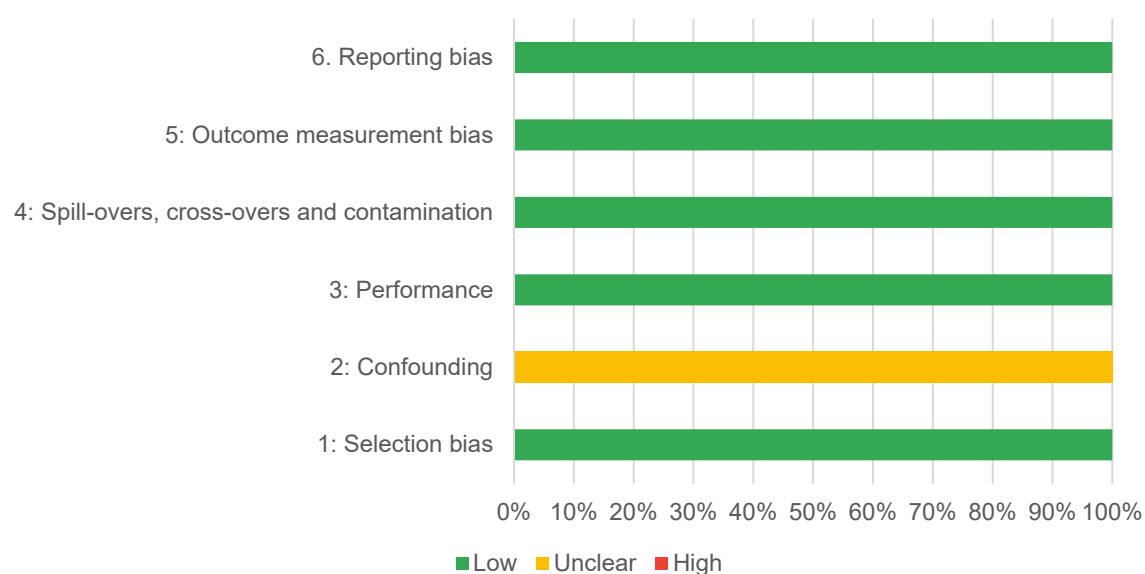
Table 41: Summary of Secondary and Immediate outcome

Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	1
	Economic and livelihood related resources	0
	Access to employment	0
Agency ability to define one's goals and act upon them and operationalised decision- making;	Individual agency	0
	Community level agency	0
	Institutions supporting agency	0
Achievement ways of being and doing which can be realised by different individuals	Improved systems	1
	Norms and behaviour change	0
	Empowerment index	0

Study design

The only study in this intervention group used a QED (Jassal 2020) which had no major concerns of risk of bias. We only identified one domain for which there may be some issues, this was the confounding domain given that authors did not report a balance test for observations at both sides of the cut-off point used for the regression discontinuity estimated to evaluate the impact of the AWPS intervention. .

Figure 113: QED AWPS RoB assessment



Qualitative studies, process evaluations and project documents

We have identified one additional document related to one programme covered by the All-Women Police Station group of studies. This single qualitative study was marked as moderate empirical quality. Given this insufficient number of linked qualitative studies (n=1) related to the All Women Police Station interventions, we were unable to conduct a qualitative evidence synthesis for this intervention group.

Synthesis of findings

The following subsection presents the results of effectiveness of all women police stations on gender equality, women's empowerment and peace outcomes.

Quantitative findings

Effects of all women police stations on increased employment of women in law and justice sector agencies

Jassal's (2020) quasi-experimental study in India was the only study evaluating the impact of all women police stations on increased employment of women in law and justice sector agencies. They examined the daily proportion of cases assigned to women. There was a large, and statistically significant negative effect ($g = -0.45$, [95% CI: -0.58 to -0.32]), and we assessed the study as having some risk of bias concerns.

Effects of all women police stations on women having improved and equitable access to justice

Jassal's (2020) quasi-experimental study in India was also the only study evaluating the impact of all women police stations on women having improved and equitable access to justice. There was a very small, and not statistically significant impact ($g = 0.07$, [95% CI: -0.06 to 0.19]), and we assessed the study as having some risk of bias concerns.

Summary of findings and discussion

We included one study in South Asia that evaluated the effect of all women police stations in India. We were able to examine effects on the following outcomes: Increased employment of women in law and justice sector agencies and women have improved and equitable access to justice. Our included studies report against two of the three secondary outcomes (Resources, Agency and Achievement) and two of the nine immediate outcomes in our review. Overall, the GRADE assessments generally indicate a very low to low certainty in this body of evidence. We identified only one linked qualitative study and were unable to conduct a qualitative evidence synthesis related to AWPS interventions. Table 42 below presents the GRADE review of our findings:

Table 42: GRADE summary of findings and certainty of evidence on all women police stations

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(AA2) Increased employment of women in law and justice sector agencies										
1	QED	serious ^a	serious ^b	not serious	serious ^c	strong association	949	SMD 0.44 SD lower (0.58 lower to 0.32 lower)	⊕⊕○○ LOW	Limited Importance
(CA7) Women have improved and equitable access to justice.										
1	QED	serious ^a	serious ^b	not serious	serious ^c	none	1898	SMD 0.06 SD higher (0.05 lower to 0.19 higher)	⊕○○○ VERY LOW	Important, but not critical

Note. CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

a. Downgraded once due to uncertainty surrounding confounding.

b. All single studies downgraded once for inconsistency.

c. Downgraded once because of very wide confidence intervals.

4.4.2 Sensitisation campaigns

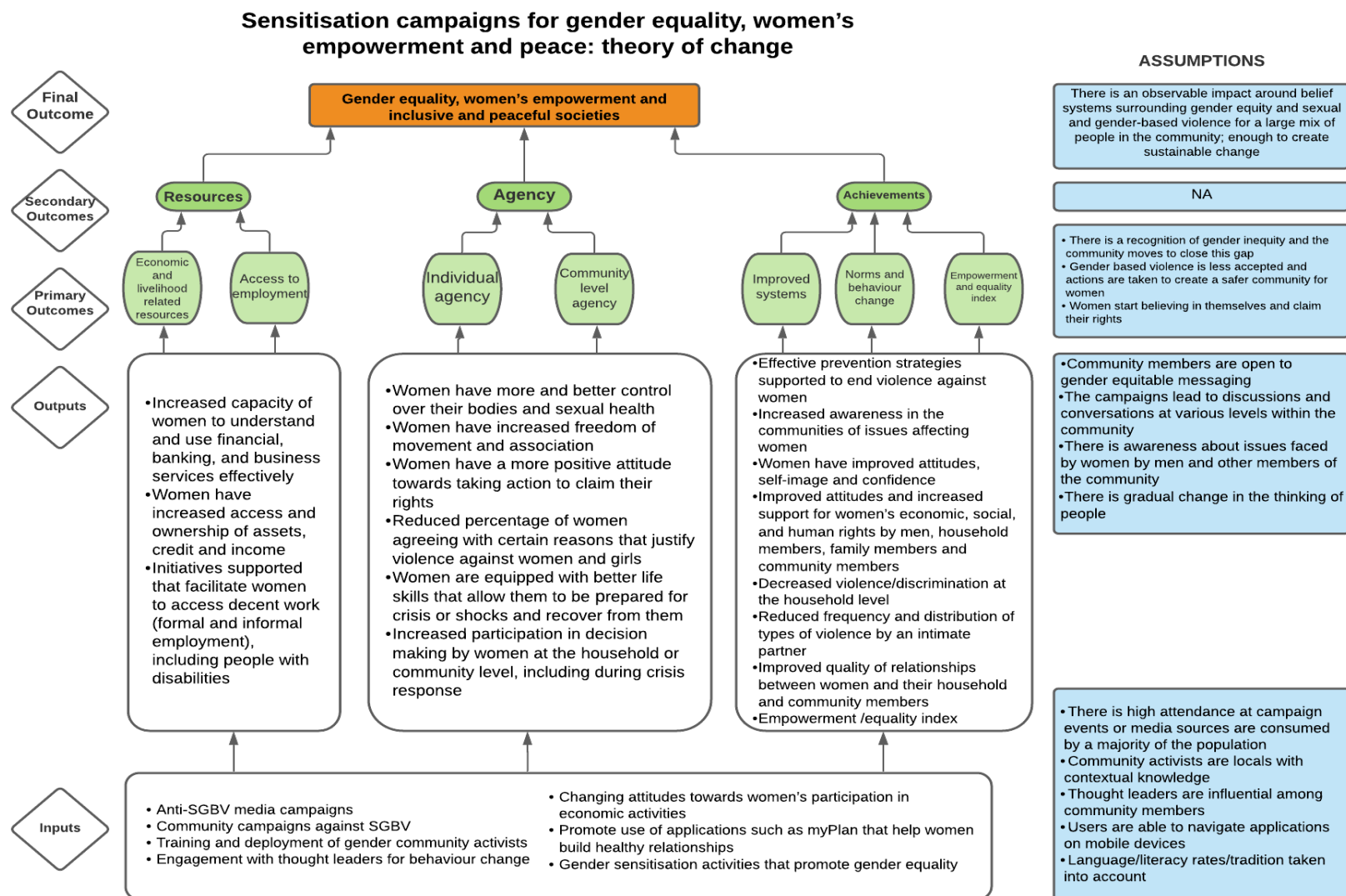
Sensitisation campaigns include the provision of access to information and messaging and/or raised awareness on specific topics through multimedia outlets like radio, posters and speakers (Gholami et al. 2017). Separate from discussion groups and community dialogues, sensitisation campaigns focus purely on the dissemination of messages. The interventions themselves offer no space for dialogue.

How do sensitisation campaigns affect gender equality, women's empowerment and peace outcomes?

Figure 114 below maps out the causal chain of how sensitisation campaigns may improve gender equality, women's empowerment and peace outcomes.

Sensitisation campaigns gradually uproot traditional norms and misconceptions surrounding sexual and gender-based violence and concepts of gender equity. Media campaigns, community campaigns, deployment of community activists and engagement with community leaders all sought to gradually change predispositions surrounding concepts of gender by initiating conversations and spreading awareness about such topics at different ecological levels. Changing beliefs lead to more economic initiatives for women, women having more control over their bodies, women having increased freedom of movement, increased decision making by women, support for effective prevention strategies to stop SGBV, decreased violence and discrimination at the household level and improved quality of relationships for women. Consequently, women have increased agency, and, at the community level, improved systems are employed and sustainable change in norms and behaviours is achieved. As a result, there is a more equitable and inclusive society.

Figure 114: Sensitisation campaigns for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included seven studies in nine different papers that evaluated the effect of six identified programmes. We included more than one paper that evaluated the same programme if the author(s) undertook different analyses or reported different outcomes over several papers.

Population

Sensitisation campaigns in our review, in general, targeted the collective consciousness of the members of a community to create a more inclusive community. Two interventions aimed to create change at the household level by targeting couples. Men's Action to Stop Violence Against Women focussed only on men. The included studies evaluated six different programmes and trials in South Asia and Sub-Saharan Africa: Bangladesh, India, Kenya, Rwanda and Uganda.

Intervention, inputs and activities

The included studies evaluated a range of sensitisation activities and inputs including:

- Anti-SGBV media campaign (n = 2): Utilised radio, television and other instruments of mass media to spread messages against SGBV and for gender equity.
- Community activists (n = 2): Trained and deployed community activists that are responsible for sensitising community members on violence against women and girls and promoting gender equitable norms.
- Community campaigns (n = 1): Inclusive community campaigns spreading gender equitable messaging.
- Engagement with thought leaders (n = 1): Created change by spreading awareness through community leaders that are well respected by people.
- Applications for SGBV preventions (n = 1): Used the myPlan app to sensitise women about healthy relationships and strategies to mitigate IPV
- Gender sensitisation activities (n = 1): Series of activities that led to the improvement in the status/empowerment of women and gender parity between women and men.
- Savings groups and safe spaces (n = 1): Specific use of such tools with the overall aim to promote women's partners acceptance of their participating in economic development activity, thus focusing on norm change at the household level

Table 43: Activities and design features of included sensitisation studies

Study	Activity/Input	Length of treatment	Intervention frequency
Abramsky – 2014	Training and deployment of gender community activists	33.5 months	Weekly
Das - 2012	Community campaigns against SGBV	-	-
Decker - 2020	Promote use of applications such as myPlan that help women build healthy relationships and flag unhealthy practices in their relationships	7 months	-
Dunkle - 2020	Changing attitudes towards women's participation in economic activities	36 months	-

Study	Activity/Input	Length of treatment	Intervention frequency
Green - 2020	Anti-SGBV media campaigns	2 months	Weekly
Green - 2018	Anti-SGBV media campaigns	2 months	Weekly
Quisumbing - 2020	Gender sensitisation activities to promote gender equality	29 months	Monthly
Schensul - 2015	Engagement with thought leaders for behaviour change	84 months	Weekly
Watts - 2015	Training and deployment of gender community activists	33.5 months	Weekly

Comparison

All our included studies compared treated groups to comparison groups receiving no intervention. Four studies included multiple treatment arms.

Outcomes

The included studies reported on a number of relevant outcomes, including the following:

- Women have a more positive attitude towards taking action to claim their rights (n = 3): Women feel entitled to be engaged and given the leadership capacity and knowledge to claim their rights and take action on relevant issues. Increased self-efficacy and autonomy. Increased opportunities for women to claim their rights including as a result of education and sensitisation. Women are engaged and given the leadership capacity and knowledge to claim their rights and take action on relevant issues.
- Reduced percentage of women agreeing with certain reasons that justify violence against women and girls (n = 2): Women are sensitised and empowered to recognise harmful social norms, particularly relating to intimate partner violence (IPV). As a consequence of this sensitisation, women are less keen to agree with reasons justifying wife-beating and are equipped with resources to claim their right and prosecute perpetrators.
- Increased awareness in communities of the issues affecting women (n = 2): Community members are easily able to recognise and support women, by having knowledge and understanding of these particular issues.
- Improved attitudes and increased support for women's economic, social and human rights by men, household and family (n = 4): These positive attitudes can be shaped by sensitisation and education to shift social norms, particularly relating to marginalised groups. As a result, they are aware of the specific needs of these groups and take an active part in reducing inequalities.
- Reduced frequency and distribution of types of violence by an intimate partner (n = 4): Types of violence in this case include four types of IPV: physical violence, sexual violence, stalking, and psychological aggression. These incidences are reduced over time due to a bundle of interventions that focus on both support for survivors and sanctioning perpetrators.

The division of the immediate and secondary outcomes is reported in Table 44 below:

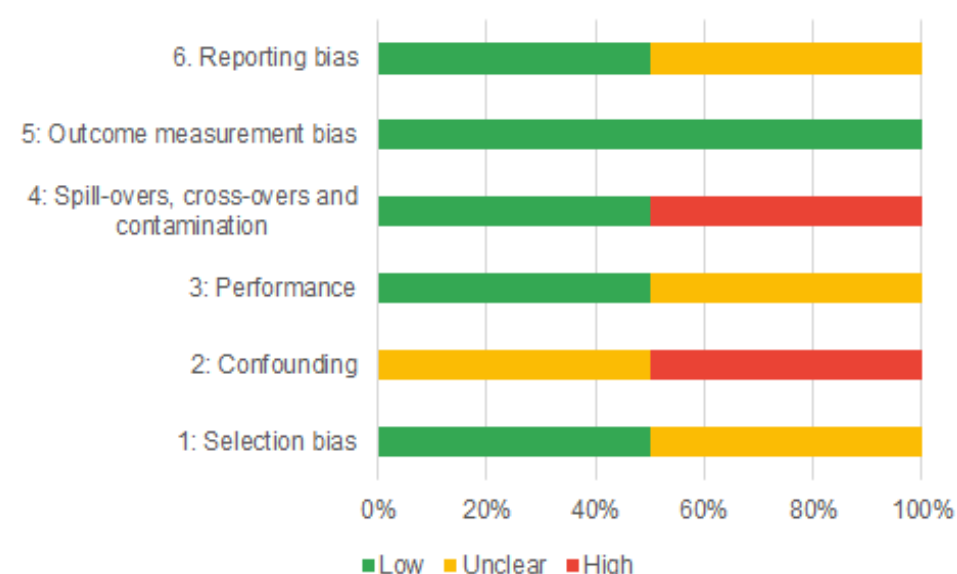
Table 44: Summary of secondary and immediate outcomes

Secondary Outcome Category	Immediate Outcome	Number of Studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	0
	Economic and livelihood related resources	1
	Access to employment	1
Agency ability to define one's goals and act upon them and operationalised decision- making	Individual agency	4
	Community level agency	1
	Institutions supporting agency	0
Achievement ways of being and doing which can be realised by different individuals	Improved systems	2
	Norms and behaviour change	6
	Empowerment index	2

Study design

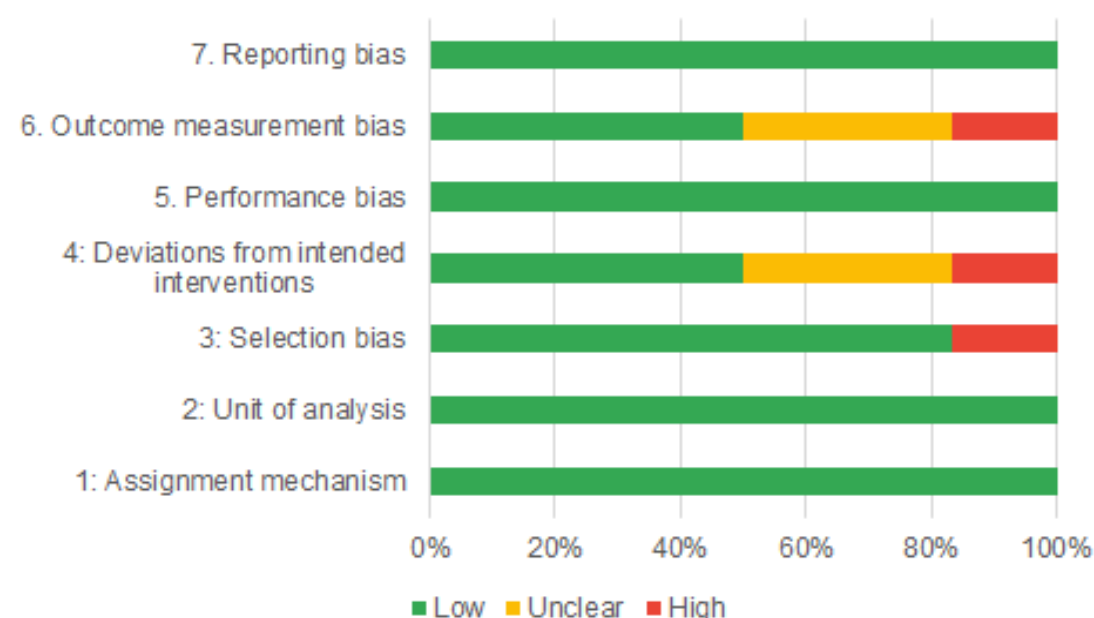
Two sensitisation campaigns studies used a QED (Das 2012; Schensul 2015). Both studies were assessed as having a high risk of bias. Reasons for this assessment included not reporting a Hausman test to suggest there is no evidence of endogeneity of the instrument (Das 2012), and potential contamination of other activities run by NGOs present in the area of intervention (Schensul 2015). We did not identify any other major issues, however in each of the following categories at least one study did not provide enough information to justify the absence of selection, performance or reporting bias.

Figure 115: Sensitisation campaign studies QED RoB assessment



Six studies used an experimental design. This included Abramsky (2014), Decker (2020), Dunkle (2020), Green (2020), Quisumbing (2020) and Watts (2015). We assessed two of these studies as having a low risk of bias, another two as having a high risk of bias, and the rest had some concerns. As detailed in Figure 116 below, high risk of bias was observed in selection, deviations from intended intervention, and outcome measurement domains. Selection bias could have occurred because of differential attrition between treatment and control groups (Dunkle 2020), deviations from interventions were observed because part of the control group in a waitlist received the intervention at the same time as the treatment group (Abramsky 2014), and measurement error was an issue because of potential reporting bias in self-reported outcomes (Abramsky 2014). However, these limitations were not common to more than one study and we did not identify any issues with the assignment mechanism of the included studies, the unit of analysis or the reporting of analysis and outcomes.

Figure 116: Sensitisation campaigns RCT RoB assessment



Qualitative studies, process evaluations and project documents

We identified 10 additional documents related to four programmes covered by the sensitisation campaign studies:

- myPlan Kenya (Kenya): one descriptive quantitative document
- Indashyikirwa - 'Agents of Change' (Rwanda): three qualitative studies and two process evaluations
- Men's Action to Stop Violence Against Women (MASVAW) (India): one qualitative study
- SASA! (Uganda): two qualitative studies and one descriptive quantitative study

Four of the 10 studies were marked as high quality, with the remainder of studies each falling into the moderate (n=3) and low (n=3) categories.

Synthesis of findings

The following subsection presents the results of effectiveness of sensitisation groups.

Quantitative findings

Effects of sensitisation campaigns on women having increased access and ownership to assets, credit and income

Dunkle and colleagues' (2020) experimental study in Rwanda was the only study evaluating the impact of sensitisation campaigns on women having increased access and ownership to assets, credit and income. Their report included two effects measured at 12 and 24 months that fell into this outcome category (any earned income and any household debt payments). The effects ranged from very small, positive point estimates at 12 months ($g = 0.03$, [95% CI: -0.08 to 0.13]) to small, positive effects at 24 months ($g = 0.17$, [95% CI: 0.05 to 0.30]). We assessed the study as having high risk of bias.

Effects of sensitisation campaigns on initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities

Quisumbing's (2020) experimental study in Bangladesh was the only study evaluating the impact of sensitisation campaigns on initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities. Their report included 6 effects that fell into this outcome category for each of two treatment arms (nutrition and agricultural). The effects measured were, the number of hours spent on work, work balance and time spent working. The effects ranged from very small and not significant point estimates ($g = 0.00$, [95% CI: -0.11 to 0.11]) when looking at first differences, to large, positive and statistically significant effects ($g = 0.58$, [95% CI: 0.47 to 0.69]) when comparing means. We assessed the study as having some risk of bias concerns.

Effects of sensitisation campaigns on women having more and better control over their bodies and sexual health

Decker and colleagues' (2020) experimental study in Malawi was the only study evaluating the impact of sensitisation campaigns on women having more and better control over their bodies and sexual health. There was a very small, not statistically significant, point estimate ($g = 0.01$, [95% CI: -0.21 to 0.23]), and we assessed the study as having some risk of bias concerns.

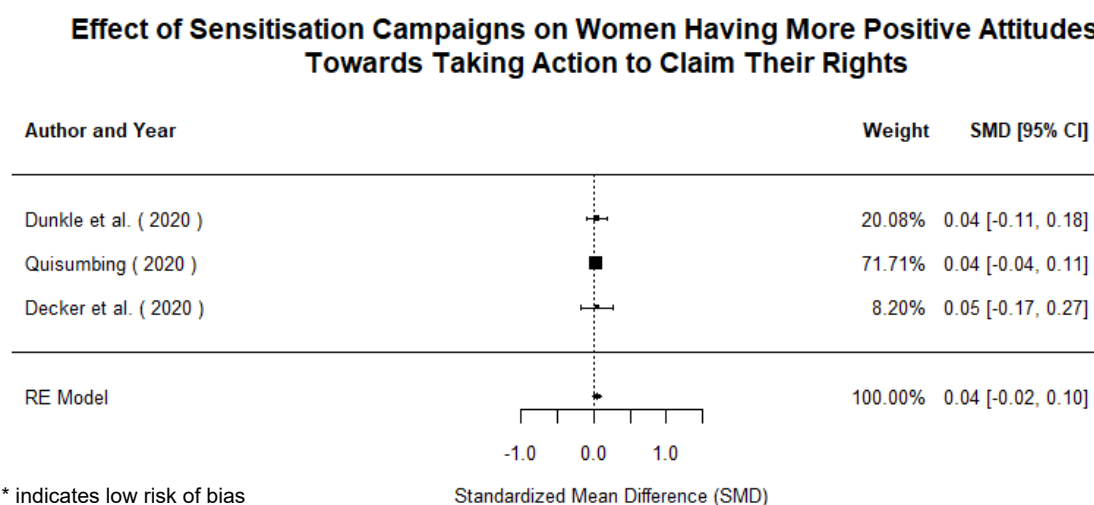
Effects of sensitisation campaigns on women having increased freedom of movement and association

Quisumbing's (2020) experimental study in Bangladesh was the only study evaluating the impact of sensitisation campaigns on women having increased freedom of movement and association. Their report included six effects that fell into this outcome category for each of the two treatment arms (e.g., Visiting important locations). The effects ranged from large, negative effects ($g = -0.67$, [95% CI: -0.78 to -0.56]) to large, positive effects ($g = 2.32$, [95% CI: 2.19 to 2.45]). We assessed the study as having some risk of bias concerns.

Effects of sensitisation campaigns on women having more positive attitudes towards taking action to claim their rights

We included a total of $k = 3$ studies in the analysis. We assessed none of the studies as low risk of bias, two as some concerns, and one as high risk of bias. The observed outcomes ranged from 0.04 to 0.05. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.04$ (95% CI: -0.02 to 0.10). Therefore, the average outcome did not differ significantly from zero ($z = 1.19$, $p = 0.23$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 117 SCBA5.

Figure 117: SCBA5: Forest plot showing the observed outcomes and the estimate of the random-effects model

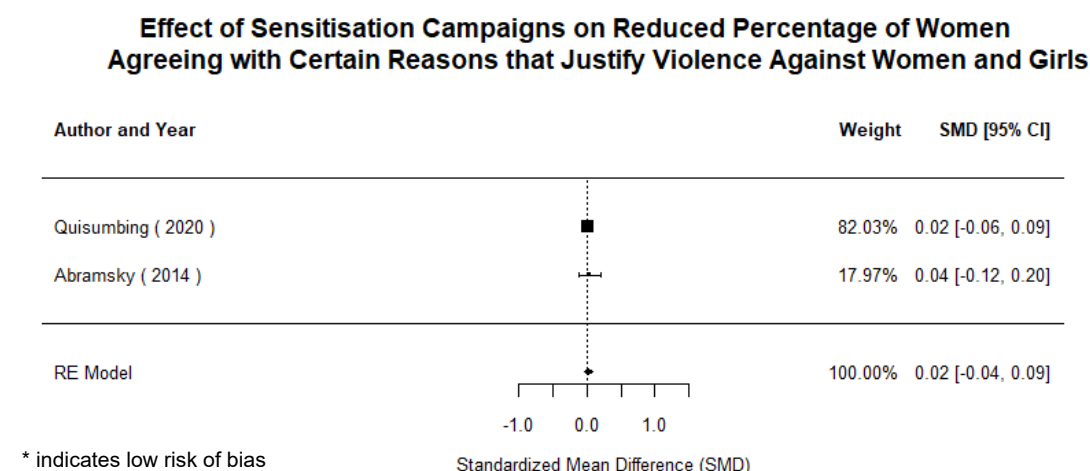


According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(2) = 0.01$, $p = 1.00$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.39 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies and no heterogeneity, moderator analyses were not appropriate.

Effects of sensitisation campaigns on reducing the percentage of women agreeing with certain reasons that justify violence against women and girls

Two studies examined the impacts of sensitisation campaigns on reducing the percentage of women agreeing with certain reasons that justify violence against women and girls, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as having some concerns of risk of bias and the other as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.02$ (95% CI: -0.04 to 0.09). Therefore, the average outcome did not differ significantly from zero ($z = 0.66$, $p = 0.51$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 118 SCBA6. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.06$, $p = 0.81$, $\tau^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid. None of the studies was appraised as having a low risk of bias.

Figure 118: SCBA6: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of sensitisation campaigns on women being equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them.

Decker and colleagues' (2020) experimental study in Malawi was the only study evaluating the impact of sensitisation campaigns on women being equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them. There was a medium, but not statistically significant, point estimate ($g = 0.21$, [95% CI: -0.01 to 0.43]), and we assessed the study as having some risk of bias concerns.

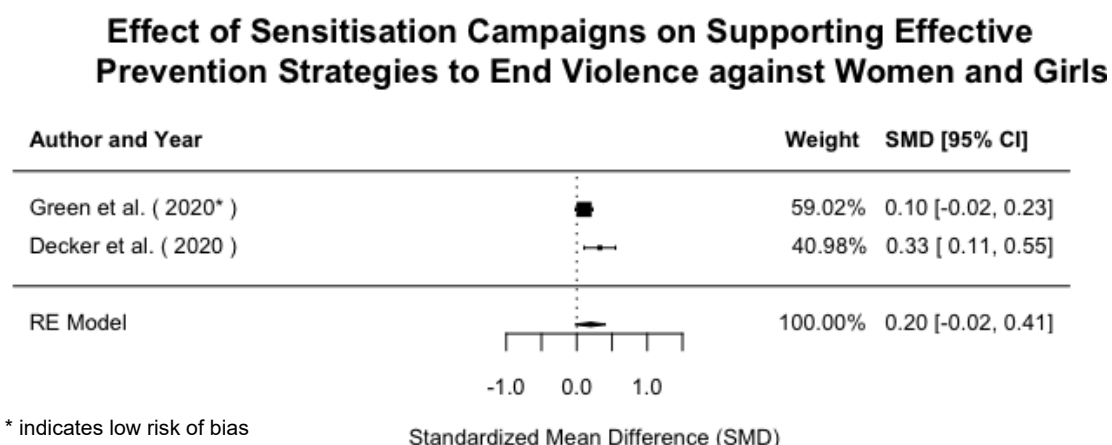
Effects of sensitisation campaigns on increased participation in decision making by women at the household or community level

Quisumbing's (2020) experimental study in Bangladesh was the only study evaluating the impact of sensitisation campaigns on increased participation in decision making by women at the household or community level, including during crisis response. For each of two treatment arms (nutrition and agricultural training), their report included 12 effects that fell into this outcome category (e.g., Access to and decisions on financial service). The effects ranged from very small, positive, but not statistically significant point estimates ($g = 0.01$, [95% CI: -0.07 to 0.08]) to large, positive effects ($g = 17.70$, [95% CI: 17.19 to 18.21]). We assessed the study as having some risk of bias concerns.

Effects of sensitisation campaigns on supporting effective prevention strategies to end violence against women and girls

Two studies examined the impacts of sensitisation campaigns on supporting effective prevention strategies to end violence against women and girls, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as low risk of bias and the other as some concerns. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.1958$ (95% CI: -0.0193 to 0.4109). Therefore, the average outcome did not differ significantly from zero ($z = 1.7845$, $p = 0.0743$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 119 SCCA6. Given the small number of studies, this result should be interpreted with caution. According to the Q-test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 2.19814$, $p = 0.0842$, $\tau^2 = 0.0165$, $I^2 = 66.4600\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid. One of the studies was appraised as having a low risk of bias.

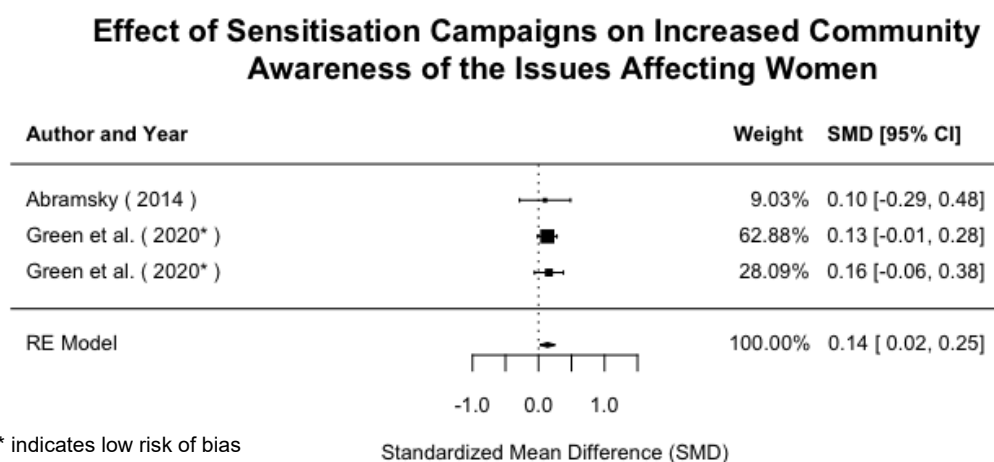
Figure 119: SCCA6: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of effect of sensitisation campaigns on increased community awareness of the issues affecting women

We included a total of $k = 3$ studies in the analysis. We assessed one of the studies (two of the effects) as low risk of bias and the other as high risk of bias. The observed outcomes ranged from 0.10 to 0.16. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.14$ (95% CI: 0.02 to 0.25). Therefore, the average outcome differed significantly from zero ($z = 2.32$, $p = 0.02$). Here we include two independent effects from Green and colleagues (2020), one for men ($g = .13$) and one for women ($g = .16$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 120 SCCB1. Two of the three studies were appraised as low risk of bias.

Figure 120: SCCB1: Forest plot showing the observed outcomes and the estimate of the random-effects model

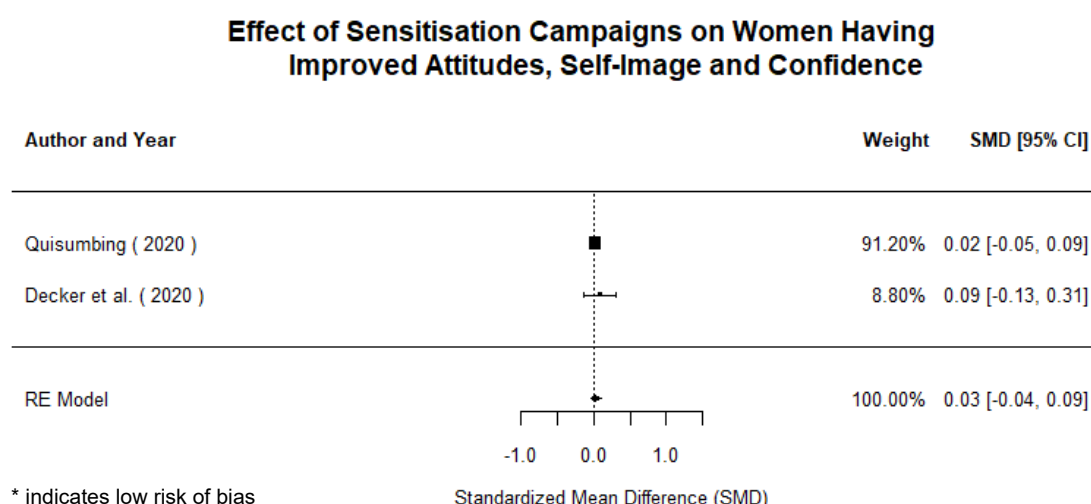


According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(2) = 0.08$, $p = 0.96$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.39 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be overly influential. With only three effects and no heterogeneity, moderator analyses were not appropriate.

Effects of sensitisation campaigns on women having improved self-image and confidence

Two studies examined the impacts of sensitisation campaigns on reducing the percentage of women agreeing with certain reasons that justify violence against women and girls, thus we included $k = 2$ studies in the analysis. We assessed both of the studies as having some concerns of risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.03$ (95% CI: -0.04 to 0.09). Therefore, the average outcome did not differ significantly from zero ($z = 0.86$, $p = 0.39$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 121 SCCB3. Given the small number of studies, this result should be interpreted with caution. According to the Q-test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.29$, $p = 0.59$, $\tau^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid. Neither of the studies were appraised as having a low risk of bias.

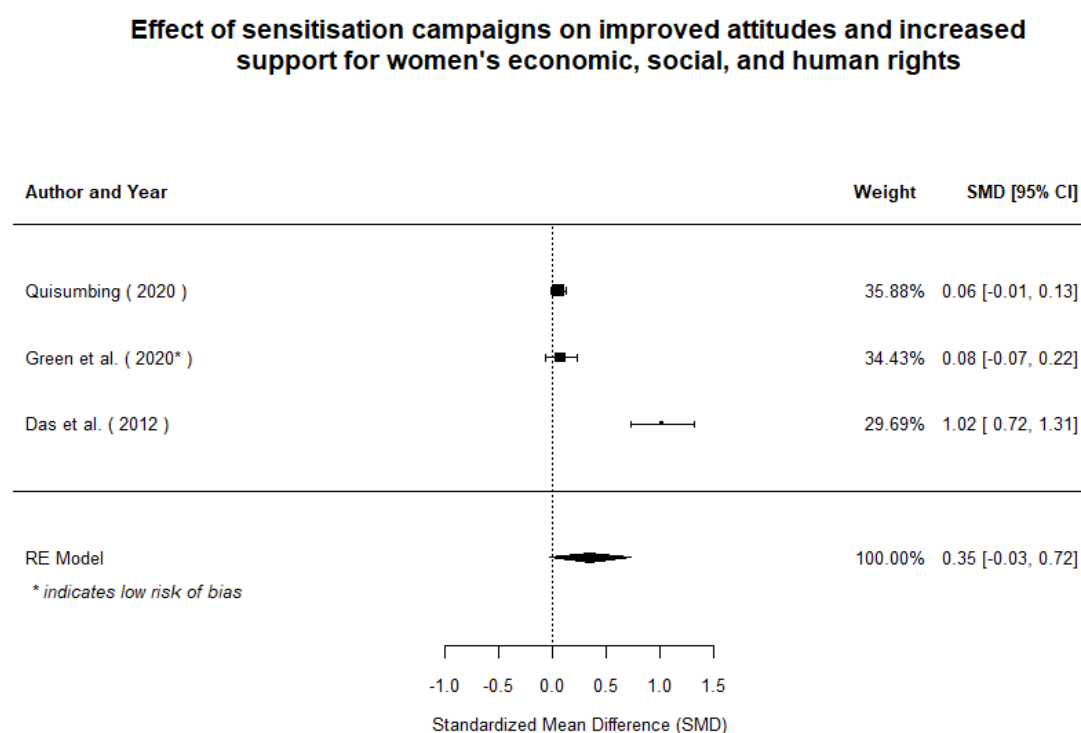
Figure 121: SCCB3: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of sensitisation campaigns on improved attitudes and increased support for women's economic, social, and human rights

We included a total of $k = 3$ studies in the analysis. We assessed one of the studies as low risk of bias, one as some concerns, and one as high risk of bias. The observed outcomes ranged from 0.06 to 1.02 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.35$ (95% CI: -0.03 to 0.72). Therefore, the average outcome did not differ significantly from zero ($z = 1.83$, $p = 0.07$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 122 SCCB4.

Figure 122: SCCB4: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(2) = 39.03$, $p < 0.01$, $\hat{\tau}^2 = 0.10$, $I^2 = 94.88\%$). An examination of the studentised residuals revealed that one study (Das et al. 2012) had a value larger than ± 2.39 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

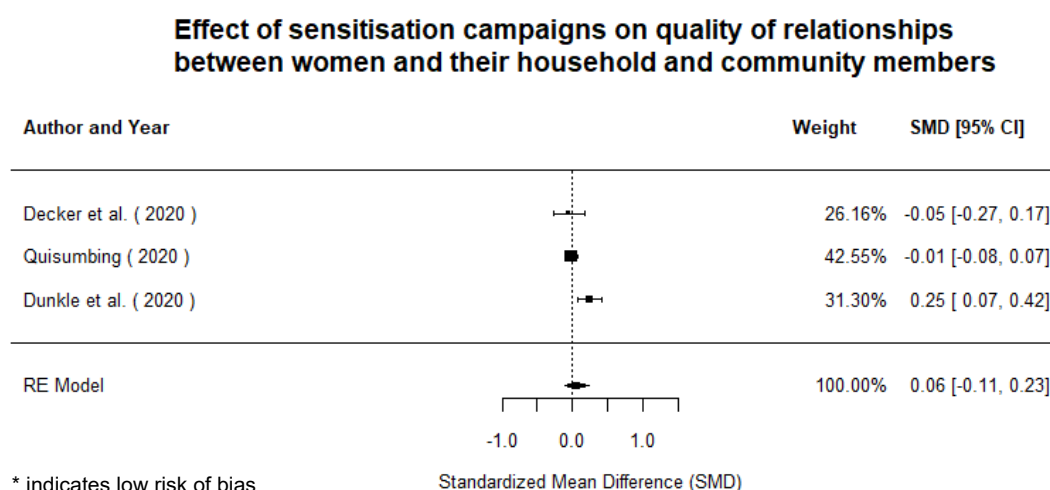
Effects of effect of sensitisation campaigns on decreased violence/discrimination at the household level

Dunkle and colleagues' (2020) experimental study in Rwanda was the only study evaluating the impact of sensitisation campaigns on decreased violence/discrimination at the household level. Their report included two effects that fell into this outcome category: children in household witnessing IPV at 12 and 24 months of exposure. For both waves the effect was medium, negative and significant ($g = -0.22$, [95% CI: -0.37 to -0.07]). However, we assessed the study as having high risk of bias.

Effect of sensitisation campaigns on quality of relationships between women and their household and community members

We included a total of $k = 3$ studies in the analysis. We assessed none of the studies as low risk of bias, two as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.05 to 0.25 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.06$ (95% CI: -0.11 to 0.23). Therefore, the average outcome did not differ significantly from zero ($z = 0.70$, $p = 0.49$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 123 SCCB8. None of the studies were appraised as low risk of bias. Accordingly, these results should be applied cautiously.

Figure 123: SCCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model

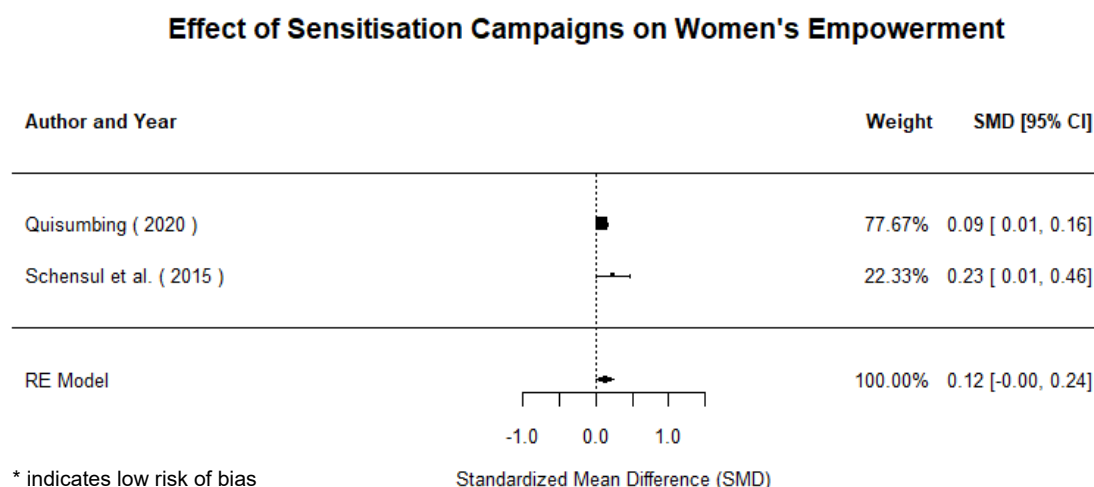


According to the Q -test, the true outcomes appear to be heterogeneous ($Q(3) = 8.51$, $p = 0.04$, $\hat{\tau}^2 = 0.01$, $I^2 = 64.74\%$). An examination of the studentised residuals revealed that one study (Dunkle et al. 2020) had a value larger than ± 2.50 and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Dunkle and colleagues (2020) would reduce the overall average effect ($\hat{\mu} = -0.01$ (95% CI: -0.15 to 0.36), making the effect negative but still non-significant ($z = 0.83$, $p = .40$). According to the Cook's distances, none of the studies could be overly influential. With only three studies, moderator analyses were not appropriate.

Effects of sensitisation campaigns on women's empowerment index

Two studies examined the impacts of sensitisation campaigns on women's empowerment index, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as having some concerns of risk of bias and the other as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.12$ (95% CI: -0.001 to 0.24). Therefore, the average outcome did not differ significantly from zero ($z = 1.94$, $p = 0.05$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 124 SCC1. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 1.45$, $p = 0.23$, $\tau^2 = 0.003$, $I^2 = 31.23\%$). With only two studies, moderator analyses were not possible, and tests of publication bias are not valid.

Figure 124: SCCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model



Qualitative findings

We conducted a thematic synthesis on the 10 linked qualitative studies to the included sensitisation campaign interventions. As indicated above, this thematic synthesis aims to identify themes related to the interplay of intervention design, intervention implementation, target population, and contextual variables with intervention outcomes and effects. In total, we identified 17 descriptive themes, which we configured into five analytical themes (Online Appendix N). These five analytical themes present the synthesis results and are discussed in more detail below.

Theme 1: Patriarchal attitudes and gendered norms hinder sensitisation campaigns whilst religious messages can facilitate the process of attaining gender equality sensitisation.

Constructive religious messages and principles can encourage sensitisation and changing attitudes towards gender equality. More specifically, while rigid gender norms and attitudes act as a potential barrier to gender sensitisation, religious messages can act as a useful channel. Due to gendered norms, women tend to refrain from allowing their husbands to participate in domestic work. Further, women may not be comfortable discussing intimate issues in public and confiding in others. Based on findings from the qualitative evidence base, men who were more actively and visibly involved in domestic work were considered outliers. Men may be reluctant to loosen control over their wives and allow them to work outside the home (Mogford, Irby and Das 2015).

Traditional family structures and rigid patriarchal attitudes were reported to curtail sensitisation across communities. For instance, Men's Action for Stopping Violence Against Women (MASVAW) respondents in India reported the difficulty of bypassing existing norms and roles in the context of extended families (Mogford et al. 2015). In this way, "...when a MASVAW man living in an extended family" confronted the status quo, he was perceived as being an opponent of family members' lifestyles. Gender norms made it difficult for women to adjust to the sensitisation of men into roles that were traditionally reserved for women. In fact, some spouses were not comfortable with the increased involvement of their husbands in household chores due to the family pressure they faced as a result of allowing this (Mogford et al. 2015).

Curriculums designed to help adjust power dynamics in relationships, identify the causes of IPV and improve knowledge on the rights of women seemed to be successful in discouraging IPV and enhancing sensitisation to the IPV problem in a Rwandan context (Stern and Nyibizi 2018). The Indashyikirwa Intimate Partner Violence Prevention Programme (Indashyikirwa) was observed to aid men's understanding of the perpetration of sexual IPV and sensitised women to their own rights to refusal. One Indashyikirwa man stated: *"I thought, my wife is mine, I married her; so, there is nothing which can prevent me from having sexual intercourse with her. But then, after receiving the lesson on sexual intercourse, we found out that she can sometimes be feeling unwell when you force her to have sexual intercourse, in that case you are committing violence."* Women enrolled in the programme also reported improved spousal relations and knowledge owing to better understanding their own rights to sexual consent (Stern et al., 2020a). Indashyikirwa couples described better-quality relationships due to a better understanding of IPV as they adjusted their behaviour and focus on their caregiving responsibilities within the relationship (Stern and Nyiritunga 2017).

The role of reflection through religious messages and principles supported participants to understand and internalise that the genders were all equal. Essentially, religion acted as a conduit for participants to better understand how power imbalances between genders could be improved. One religious leader expressed: *"The Indashyikirwa training has been very useful to us even though it is hard in the beginning because there are some who didn't understand. There has been a challenge of mentality but because we are used to learning the word of God, we know everyone should have the equal right regardless of their gender. That has been very helpful to us,"* (Stern, Heise and Cislaghi 2020). The training component of Indashyikirwa was a valuable tool because it provided couples with a participatory environment and problem-solving skills platform.

Theme 2: Culture of openness and community engagement supports sensitisation to gender issues.

Open dialogue forums and communication between and within communities can aid the resolution of gender-related conflict as people collaborate to understand gender issues and reduce IPV. Third parties such as other community members realise that there are greater incentives to intervene during conflict as it may help build respect and trust in relationships. Women are additionally encouraged to discuss their experiences and sensitise their communities to violence against women and other gender-related issues.

By using participatory training and community-based activism, the Indashyikirwa programme encouraged a culture of community engagement to activate learning and third-party intervention. In this way, participants supported those that suffered from IPV. One community leader expressed that: *"We studied laws protecting women. That has made us more confident because if I see a victim of violence, it would not cause me any problems to look for a leader and tell him there is a person who is a victim of violence there. After receiving that training, I felt I also have to play a role"*. An opinion leader offered that: *"We used to face cases of violence in our work but ignored how to classify the types of violence. I and my colleagues who had the training, we came to understand this. ...we immediately understand what it is about and how to solve it quickly"* (Stern, et al. 2020b). While Indashyikirwa respondents from Chatterji and colleagues (2020) reported that more intimate settings were conducive to thorough engagement, others from Stern and colleagues (2020b) voiced greater confidence in raising awareness and

openly discussing IPV: *“Now I openly speak out and I use the power that I have in me and I feel there is something that I can do to make my family developed. That is a very big thing.”* The Indashyikirwa also found value in using community forums such as parent’s evenings to sensitise communities on the negative consequences of perpetrating IPV (Stern and Niyibizi 2018). Ultimately, community-wide coordination and planning may have ensured that implementation was relatively easy as one activist noted: *“When I prepare a discussion, I ask for help from the village leader. The village leader tells the community members to come for a meeting; comes to start a meeting for you, then you continue with the discussion and he closes for you. These are thanks to the quarterly meeting you had together, about taking measures and solutions”*.

SASA! in Uganda used community mobilisation to mitigate the scourge of IPV and to develop quality of relationships between couples. The engagement of couples by SASA! staff equipped them with problem-solving skills that encouraged greater mutual respect and more open communication channels (Starmann et al. 2017). By helping participants understand the main tenets of a healthy relationship through community engagement, some couples reported positive responses to the intervention as two men observed in their relationships respectively: *“[W]e started smiling, we started talking and discussing issues well together,”* and *“What I am most happy about is the agreeing and understanding each other...it shows love in the relationship.”* Engagement with SASA! staff seemed to work well to support greater consideration of equality and choice between partners, openness around sensitive issues and shifts in gender norms and roles. Regarding household finances, a SASA! man in Uganda observed: *“I don’t hide anything from her . . . it was SASA! that taught us all that... They tell us that it is good for each one of you to inform the other about your income. You have to let them know how much you earn and together you decide how to use that money. You discuss what to buy, how much to save and you don’t use another’s money without their consent because you have power. You have to agree and if one of the parties does not want [to] then you don’t force her”* (Kyegombe et al. 2014).

Theme 3: Awareness, conflict resolution and communication skills training contribute to greater group sensitisation and awareness to IPV.

Trainings to build awareness of IPV and conflict resolution communication skills are observed to be important in the sensitisation of men and boys given the lack of understanding of various forms of IPV among men. Training presents an opportunity for men to challenge traditional gender boundaries. Partner involvement in this training enhances openness about personal experience and combined decision-making, particularly regarding finances.

Post-intervention, many participants of the Indashyikirwa programme expressed greater awareness in understanding IPV in its various forms after being sensitised to the violation behind it (Stern and Niyibizi 2018). Men were sensitised to the perverse nature of the act of violence on both women and children; as one woman remarked on the behaviour of her partner: *“He has many regrets and said I wish I could get back the time that I wasted and said this will not happen again. Sometimes the child even tells him, ‘do you remember the way you used to beat me?’ And my husband tells him, ‘I promise I will not beat you again.’”* (Stern et al. 2020). Analogous to this were the qualitative findings that behaviour change communication influenced their conflict resolution skills. The curriculum provided in the Indashyikirwa allowed participants to self-reflect and discover

the underlying causes for IPV as one man in the programme recounted: *“After drinking I felt, I am a man and I would commit violence against my wife but after having studied about alcohol, I realised it prevents me from reaching any household development. Now I am no longer a drunkard. I am now a person who builds his family and my family has noticed that I have changed. We became Indashyikirwa.”*

Training men to understand the value of women’s roles better was an important exercise for MASVAW participants (Mogford et al. 2015). In particular, allowing men to imagine “the average day in a woman’s life” was instrumental in helping to provoke reflection around gender roles and norms. In the SASA! programme, a former perpetrator of physical IPV reflected that: *“They told me, ‘what you are doing is not good. You will kill your wife, and after killing her, you won’t find another wife like her.’ So, I understood that it is a shame and until now it has never happened again.”* Essentially, the qualitative evidence proposed that equipping men with conflict resolution skills was helpful in addressing attitudes to IPV (Stern and Niyibizi 2018). Additionally, financial skills were reported to be a useful addition to the training as it allowed better communication between couples regarding finances. As one man pointed out: *“I used to feel I had to be the one to say how to spend money and that my wife didn’t have a say because she had not worked for that money. But after having seen that when I go to work my wife does other housework, we put together what we have both done, we put the money at the bank or at the VSLAs since we are members of the group. We learned to save and how to spend the rest of the money. That has helped us.”* (Stern and Nyiratunga 2017). In this way, sensitising couples to find the source of spousal frictions was used to smoothen the quality of their relationships.

Theme 4: Greater contact frequency and programme length aids a positive shift in perceptions about gender norms and equality while regular and sustained interaction are key requirements for enhanced programme implementation

Greater time exposure to programme activity can improve programme effectiveness by allowing participants to adjust their beliefs and adopt better social norms over time. But sensitisation programme donors need to consider that such intervention types can be costly and protracted when undertaken. The qualitative evidence suggests that programme outcomes are more favourable when dissemination delays are minimised, targeting rigour is ensured, and activity sequencing is carefully planned. To maximise sensitisation campaigns’ influence, consistent interaction and discussions with implementation staff and peers is a required design feature. Joint engagement by relevant stakeholders and concerted activism enhanced implementation. In particular, sensitisation campaigns designed to aid and improve interpersonal relationships and change couple processes were observed to be of promise. The role of colleagues or wives in providing inspiration and motivation is crucial for sensitisation campaigns aimed at transforming male attitudes and behaviours. Moreover, beneficiaries report notable reduction in anxiety levels and greater motivation to participate under such conditions.

The longer and more frequent the programme activities, the better received the intervention was by participants as this can allow enough time to learn about and apply new skills. Qualitative responses from SASA! participants in Rwanda cited the length and frequency of exposure to intervention activities on couples as key factors for participant follow through (Starmann et al. 2017). Responses from SASA! implementation staff and community activists suggested that programme impact was likely curtailed by insufficient

time for in-depth activism across communities (Chatterji et al. 2020). Similarly, Indashyikirwa programme partners indicated that the activism component of this intervention was fairly truncated and thus insufficient to convincingly alter gender norms that are embedded within society. Furthermore, from a funding perspective, the duration of an intervention designed to shift societal norms has to be carefully considered in order to effect change (Chatterji et al. 2020).

Effective programme implementation relies heavily on sustained interaction with programme staff and designed material. The formation of sustainable relationships with community activists in the SASA! programme reportedly contributed to positive changes in personal and couple relationships (Starmann et al. 2016). SASA! participants observed that the continued guidance and support of community activists empowered couples to broaden communication avenues for quarrelsome topics to be dealt with jointly. The proximity, availability and consistency of community activists allowed SASA! participants to trust and use their advice on the resolution of relationship conflict. In the case of Indashyikirwa, the cursory use of community activists hindered their contribution as their ability to mobilise the community was curtailed by hasty implementation and poor readiness. An additional constraint in this case was the tailoring of interventions to successfully embed them in an Indian setting (Chatterji et al. 2020).

The regularity of interaction with programme staff is reported to gradually sensitise programme participants. The frequency of activities by SASA! participants was highlighted by various respondent couples as a motivating factor to continue with adherence and participation (Starmann et al. 2017). One woman responded: “[T]hey warned him, I think they all scared him...he realised that he had to change” when asked about how regular interaction with community activists had enhanced her relationship with her partner. It was noteworthy that the benefits of steady interaction were realised by couples even when only one partner adhered to this aspect of the SASA! intervention. Moreover, allowing both partners to interact with community activists was reported to yield greater success as couples were able to build healthier relationships with constant support from programme staff.

Theme 5: Community leaders including men and opinion leaders are key actors in developing and using sensitisation strategies.

Opinion leaders and men are key actors in the success of sensitisation programmes in communities. When opinion leaders realise their influence over societal outcomes and citizen behaviours, sensitisation programmes can help in focusing their efforts in a more responsible and positively influential manner geared towards gender equality.

Sensitisation programmes offer men the opportunity to learn and self-correct negative behaviours that are grounded in beliefs around masculinity. Personal histories and relationships can be a barrier in improving gender equality, but concerted sensitisation efforts can improve this.

In order to achieve greater sensitisation, community leaders and target groups are frequently appointed as agents of change to lead programme activities and positively influence other participants. As a field member of the implementation staff in the Indashyikirwa programme in Rwanda observed: *“Several opinion leaders shared testimonies of change in their relationships by the end line interviews, including men more actively supporting domestic and caregiving roles, more equitable decision-making, and improved*

conflict resolution,” (Stern et al. 2020b). The qualitative findings indicated that opinion leaders were excited to engage actively with other participants and support programme staff in identifying and embedding knowledge on gender inequalities in their communities. Their involvement encouraged them to prioritise IPV prevention and wider engagement on gender issues with the community. Similarly, the involvement of men helped sensitisation campaigns gain traction in their localities as they were equipped with the knowledge to be champions for change on gender and equity issues (Mogford et al. 2015).

Participants of the MASVAW intervention in India communicated that they were empowered to assess their own complicity in the perpetuation of gender stereotypes and inequalities in their own societies. Some of the reported prominent behavioural changes amongst these men were greater acceptance of women’s physical autonomy, more informed views on sexual intercourse with their partners and their domestic involvement in activities that were traditionally set aside as being strictly for women. Furthermore, all participants in the MASVAW programme the reported greater satisfaction in their marriages through greater knowledge of gender issues.

A useful dynamic in the MASVAW model of interaction was selecting men to engage with other men (Mogford et al. 2015). Participants viewed this as successful because in addition to broadening their understanding of gender equality and empathy with women, they could reflect on the role of their own masculinity in peddling misinformation and unyielding attitudes to women. Moreover, opinion leaders and men were seen as successful agents of change in the sensitisation of their communities. As another opinion leader selected in the Indashyikirwa programme expressed: *“A leader must be a role model in his own family, a role model in his extended family, a role model where you live, where you work because you cannot be a bad example while you are a leader. After being trained, I gained a lot; starting from my family, even if we did not have conflicts, but it helps me be humble, and to spread a good message.”*

Discussion

Quantitative meta-analysis showed positive results for sensitisation campaigns across a number of outcomes. While there are cases of statistical insignificance at the 95 per cent confidence level (all outcomes except community awareness), various insights from the qualitative findings can supplement the quantitative results. Positive quantitative effects from sensitisation campaigns cover outcomes such as: 1) women having more positive attitudes towards taking action to claim their rights (n = 3), 2) reducing the percentage of women agreeing with certain reasons that justify violence against women and girls (n = 2), 3) supporting effective prevention strategies to end violence against women and girls (n = 2), 4) women having improved self-image and confidence (n = 2), 5) improved attitudes and increased support for women’s economic, social, and human rights (n = 3), 6) increased community awareness of the issues affecting women (n = 3), 7) reduced frequency and distribution of types of violence by an intimate partner (n = 5), 8) quality of relationships between women and their household and community members (n = 4), 9) women’s empowerment index (n = 2). Some qualitative findings that can explain these findings are as follows:

- Based on the qualitative findings, sensitisation campaigns can be limited and restricted by social factors such as gender norms and roles. Changes that are related to altering human behaviour through learning and communication commonly can take time to realise.

- Culturally developed patriarchal attitudes can stem the effectiveness of sensitisation campaigns because they are more difficult to change.
- To enhance effectiveness, sensitisation campaigns can adopt strategies that expose participants to longer and more frequent programme activities such as training and counselling. More concerted efforts to sensitise communities on a larger scale can lead to more empathetic interventions by community members and support staff in spousal conflicts.
- Hurried and disorganised implementation of sensitisation processes can hinder programme participation by communities.
- Issues of personal image, knowledge on rights and empowerment can be addressed gradually as they depend on the nature of close relationships and personal confidence from successful programme impact.
- Opinion leaders can be used as agents of change to teach about IPV and instil useful skills for couples to rely on.
- Men who are respected community leaders can act as mentors for IPV perpetrators and help to facilitate the gradual breakdown of societal norms.

Is there a more effective form of media for sensitisation campaigns?

While the qualitative evidence base assessed does not provide a clear recommendation on the most reliable medium for the dissemination sensitisation campaigns, there are some notable observations:

- Media used for sensitisation campaigns are largely in-person platforms and intimate sessions with a large focus on couples and men as perpetrators.
- Respected leaders in the community are used to counsel and sensitise couples to IPV, teach on the tenets of healthy relationships, impart conflict resolution skills and knowledge and mediate on conflicts within relationships.
- Using communal and public platforms is feasible to reach more participants at the same time while intimate settings with trusted leaders can help to more actively address obstacles to healthy personal relationships.
- A medium that is easily accessible and conscious to socioeconomic conditions related to access (to phones, places etc.) can lead to greater participation.

Summary of findings and discussion

We included seven studies in five countries in South Asia and Sub-Saharan Africa that evaluated the effect of sensitisation campaigns. The largest positive and significant impact in the category of sensitisation campaigns was on women having more positive attitudes towards taking action to claim their rights, though the magnitude of the effect was still small. There was a similar small positive and significant impact on increased awareness in communities of the issues affecting women. There were also many outcomes for which there were positive but non-significant summary effects. These included reduced percentage of women agreeing with reasons that justify violence against women and girls, increased support for prevention strategies to end violence, women having improved attitudes, self-image and confidence, improved attitudes and increased support for women's economic, social, and human rights by men/household/community members, improved quality of relationships between women and their household/community members, and women's empowerment. In addition, there were two negative but non-significant effects for reduced frequency of both physical and sexual IPV. Overall, the GRADE assessments generally indicate a range of very low to moderate certainty in this body of evidence.

The ten linked qualitative studies and their resulting five analytical themes also offered a key look into what features of sensitisation campaigns can aid or hinder outcomes. The engagement of community leaders is a promising facilitator of sensitisation and engagement by the community. However, entrenched patriarchal norms will have to result in increased buy-in by these stakeholders for appropriate uptake at the community level. Open dialogue forums may support this. Similarly, this aspect combined with regular and sustained interaction can also enhance the potential of programmes to achieve their intended outcomes. Overall, rich and in-depth qualitative data can sustain the claims made by the meta-analysis by pointing to recurring as well as contextual themes that have occurred across interventions. Table 45 below presents the GRADE review of our findings.

Table 45: GRADE summary of findings and certainty of evidence on sensitisation campaigns

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(AB2) Women have increased access and ownership to assets, credit and income.										
1	RCT	very serious ^a	serious ^b	not serious	not serious	none	1417	Four positive effect sizes with a 95% CI range of -0.08 to 0.30	⊕○○○ VERY LOW	Important, but not critical
(AC2) Initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities										
1	RCT	serious ^c	serious ^b	not serious	serious ^d	none	3245	Two zero and six positive effect sizes, all near zero with one significant outlier	⊕○○○ VERY LOW	Important, but not critical
(BA2) Women have more and better control over their bodies and sexual health.										
1	RCT	not serious	serious ^b	not serious	not serious	none	312	SMD 0.01 SD higher (0.21 lower to 0.23 higher)	⊕⊕⊕○ MODERATE	Critical
(BA3) Women have increased freedom of movement and association.										
1	RCT	serious ^e	serious ^b	not serious	very serious ^f	none	2739	Two negative, two zero, and seven positive effect sizes with a 95% CI range of -0.78 to 2.45	⊕○○○ VERY LOW	Limited importance
(BA5) Women have more positive attitude towards taking action to claim their rights.										
3	RCT – 3	serious ^g	not serious	not serious	not serious	none	3817	SMD 0.18 SD higher (0.06 lower to 0.30 higher)	⊕⊕⊕○ MODERATE	Limited importance
(BA6) Reduced percentage of women agreeing with certain reasons that justify violence against women and girls										
2	RCT – 2	serious ^h	serious	not serious	not serious	none	3338	SMD 0.02 SD higher (0.04 lower to 0.09 higher)	⊕⊕○○ LOW	Critical
(BA7) Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them.										
1	RCT	not serious	serious ^b	not serious ⁱ	serious ⁱ	none	312	SMD 0.21 SD higher (0.01 lower to 0.44 higher)	⊕⊕○○ LOW	Important, but not critical
(BB1) Increased participation in decision making by Women at the household or community level, including during crisis response										
1	RCT	serious ^j	serious ^b	not serious	very serious ^k	none	2739	29 positive effect estimates with widely varying CIs	⊕○○○ VERY LOW	Limited importance
(CA6) Effective prevention strategies supported to end violence against women and girls										
2	RCT - 2	not serious	serious ^l	not serious	serious ^m	none	1348	SMD 0.2 SD higher (0.02 lower to 0.41 higher)	⊕⊕○○ LOW	Critical
(CB1) Increased awareness in communities of the issues affecting women										
2	RCT - 2	serious ⁿ	not serious	not serious	not serious	none	1143	SMD 0.14 SD higher (0.02 higher to 0.25 higher)	⊕⊕⊕○ MODERATE	Limited importance
(CB3) Women have improved attitudes, self-image and confidence.										
2	RCT – 2	not serious	serious ^o	not serious	not serious	none	3557	SMD 0.03 SD higher (0.04 lower to 0.09 higher)	⊕⊕⊕○ MODERATE	Limited importance
(CB4) Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members										
3	RCT – 2	serious ^p	very serious ^q	not serious	serious ^r	none	4154	SMD 0.35 SD higher	⊕○○○	Limited

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
	QED – 1							(0.03 lower to 0.72 higher)	VERY LOW	importance
(CB5) Decreased violence/discrimination at the household level										
1	RCT	serious ^s	serious ^b	not serious	not serious	none	674	Two negative effect estimates with a 95% CI range of -0.37 to -0.07	⊕⊕○○ LOW	Critical
(CB7) Reduced frequency and distribution of types of violence by an intimate partner - Physical violence (PV)										
4	RCT - 4	very serious ^t	serious ^u	not serious	serious ^v	none	4018	SMD 0.06 SD lower (0.18 lower to 0.06 higher)	⊕○○○ VERY LOW	Critical
(CB7) Reduced frequency and distribution of types of violence by an intimate partner – Sexual violence (SV)										
3	RCT - 3	very serious ^a	serious ^b	not serious	serious ^v	none	1380	SMD 0.01 SD lower (0.07 lower to 0.13 higher)	⊕○○○ VERY LOW	Critical
(CB8) Improved of relationships between women and their household and community members										
3	RCT - 3	serious ^w	serious ^x	not serious	serious ^y	none	3555	SMD 0.06 SD higher (0.011 lower to 0.23 higher)	⊕○○○ VERY LOW	Limited importance
(CC1) Empowerment/Equality index										
2	RCT – 1 QED – 1	very serious ^z	not serious	not serious	not serious	none	3035	SMD 0.12 SD higher (0 to 0.24 higher)	⊕⊕○○ LOW	Important, but not critical

Note. CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

- Downgraded because of a high overall risk of bias. Selection bias is likely. Outcome measurement bias unclear.
- All single studies downgraded once for inconsistency.
- Downgraded once because of a medium overall risk of bias. Deviation from intended intervention and outcome measurement bias are assessed as unclear.
- For seven out of eight effect sizes there is large overlap of confidence intervals (CI) and low variance of point estimates, centered around the median of -0.07. There is however one outlier with an estimate of (-)0.58 and a confidence interval not overlap with any of the seven studies.
- Downgraded once because of a medium overall risk of bias. Deviation from intended intervention and outcome measurement bias are assessed as unclear.
- Very wide variance of point estimates across the twelve effect sizes at both sides of the 0.00 threshold, ranging from -0.67 to 2.3 b) Very limited overlap of confidence intervals (CI).
- Downgraded because one study presents high risk of bias and the others are unclear. There is risk across all three studies of deviation from intended intervention.
- Downgraded because half of the evidence comes from a high risk of bias study with issues relating to reporting bias and deviations from intended outcomes.
- The 95% confidence interval for the only effect size spans from no effect (-0.01) to large effect (0.43).
- Based on only one study assessed as having a medium overall risk of bias. Deviation from intended intervention and outcome measurement bias are assessed as unclear.
- Very wide variance of point estimates across the 27 effect sizes, ranging from 0 to 17 b) Very limited overlap of confidence intervals (CI).
- Downgraded to reflect large CIs and one study's point estimate not falling within the CI of the other.
- Downgraded because of the RE model's large CI's which cross both sides of the threshold.

- n. Downgraded once to reflect that one-third of the included studies presents high risk of bias due to outcome measurement and deviation from intended intervention.
- o. Both studies' estimates are near to the threshold, but a wide CI in the Decker study, taken in cognizance of there only being two total studies, warrants a downgrade.
- p. Downgraded because this groups has a high risk of bias QED whose point estimate varies hugely from the other two (lower risk of bias) RCTs.
- q. a) Wide variance of point estimates and no overlap of CI between the one QED study, and the other two RCTs. b) The Q test for heterogeneity which tests the null hypothesis that all studies have the same underlying magnitude of effect, has a very low p-value ($p < 0.01$), significant at a 99% confidence level, which may indicate a problem with heterogeneity. I^2 indicates that 95% of the variability in effect estimates may be due to heterogeneity rather than sampling error (chance).
- r. a) Rate down if confidence intervals of RE Model cross the threshold: Yes B) Rate down if one study: No, three studies, C) Sample size: Large, not rated down.
- s. Based on only one study assessed as having a high overall risk of bias. Selection bias is likely. Outcome measurement bias unclear.
- t. Out of four studies: two have high RoB, two have some concerns related to risk of bias. RoB assessments indicate that three out of four studies have been rated as (probably) not free from outcome measurement bias, and two have been marked as likely to have deviations from intended interventions.
- u. a) Relatively wide variance of point estimates across studies b) Minimal overlap of confidence intervals (CI), which suggests variation may be more than what one would expect by chance alone c) The Q test for heterogeneity which tests the null hypothesis that all studies have the same underlying magnitude of effect, has a relatively low p-value ($p < 0.06$), significant at a 90% confidence level, which may indicate a problem with heterogeneity. I^2 indicates that 59% of the variability in effect estimates may be due to heterogeneity rather than sampling error (chance).
- v. a) Rate down if confidence intervals of RE Model cross the threshold: Y B) Rate down if one study: No four studies, C) Sample size: large, not rated down
- w. Out of three studies: one has a high RoB, two have some concerns related to risk of bias. For one study selection bias has been identified as a likely risk. Outcome measurement bias has been marked as unclear for two studies, while deviations from intended interventions are marked as unclear for one. I consider this as crucial limitation for one criterion and some limitations more criteria sufficient to lower confidence in the estimate of effect - by one level.
- x. a) Relatively wide variance of point estimates across studies b) Minimal overlap of confidence intervals (CI), which suggests variation may be more than what one would expect by chance alone c) The Q test for heterogeneity which tests the null hypothesis that all studies have the same underlying magnitude of effect, has a relatively low p-value ($p < 0.04$), significant at a 95% confidence level, which may indicate a problem with heterogeneity. I^2 indicates that 65% of the variability in effect estimates may be due to heterogeneity rather than sampling error (chance).
- y. a) Rate down if confidence intervals of RE Model cross the threshold: Yes B) Rate down if one study: No, three studies, C) Sample size: >400, so not rated down.
- z. Downgraded twice because both studies present some risk of bias, and one is a high risk of bias QED. When that study is removed, the point estimate of the RE model changes significantly.

4.5 Effects of interventions under the UNSCR prevention pillar

The prevention pillar is a set of interventions that build capacities and systems to support the gender responsiveness and inclusivity of violence prevention and conflict transformation processes. This also includes efforts to hold perpetrators of violence accountable through formal or informal means.

In our systematic review, this pillar bundled the following type of interventions covering 3,872 beneficiaries:

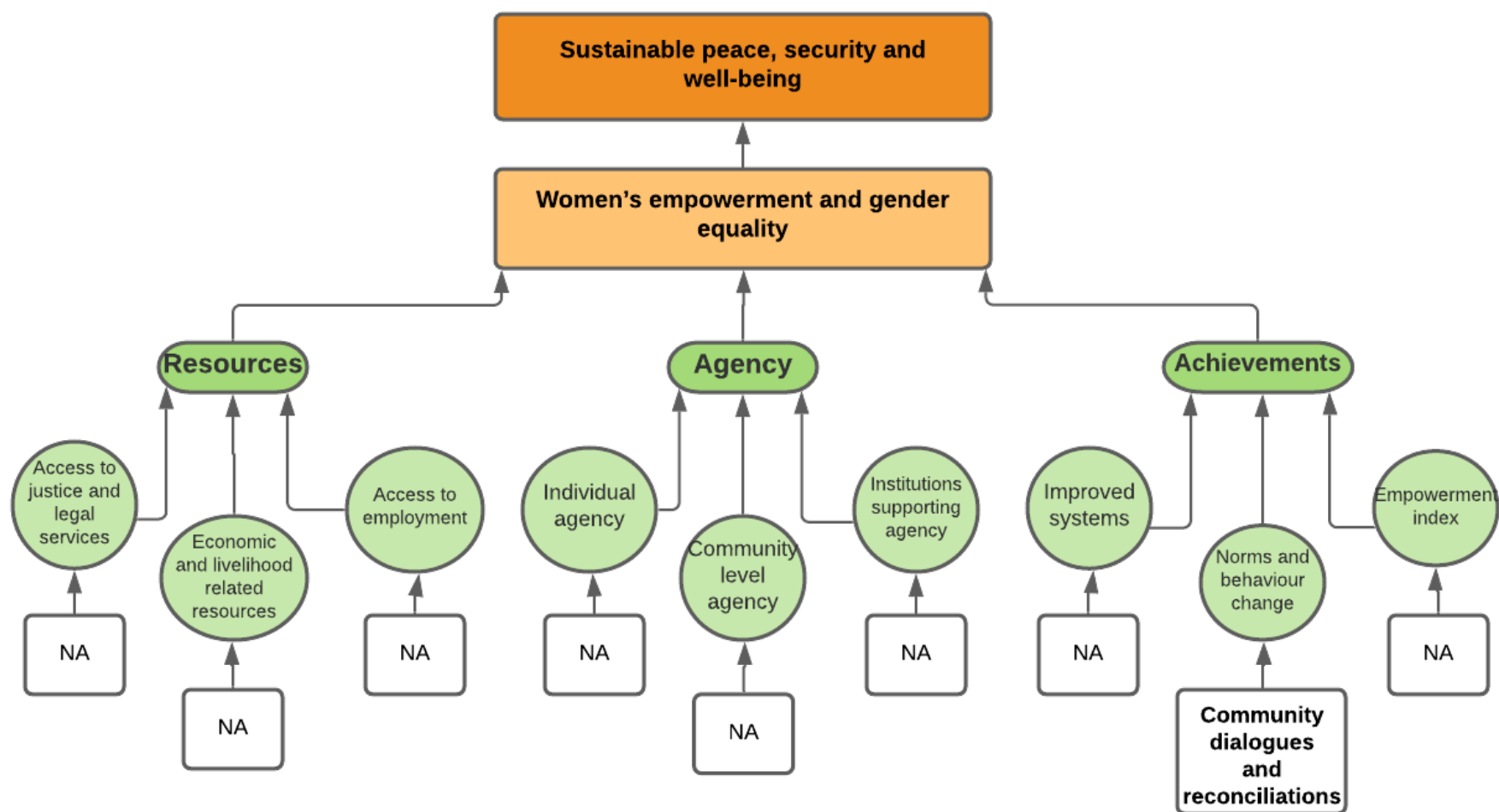
- Community dialogues and reconciliations

This section provides the findings of our synthesis of the two included studies evaluating the effect of these interventions on gender equality, women's empowerment and peaceful and inclusive societies. The section is organised by each intervention group. Each sub-section starts with a description of the intervention groups, their activities and theory of change, followed by descriptive results and the findings addressing our research questions.

Figure 125 below provides a summary of the outcomes and targeted effects of the interventions included under the prevention pillar:

Figure 125: Theory of change of the included studies under the prevention pillar

Theory of change of included studies under the prevention pillar



4.5.1 Community dialogues and reconciliations

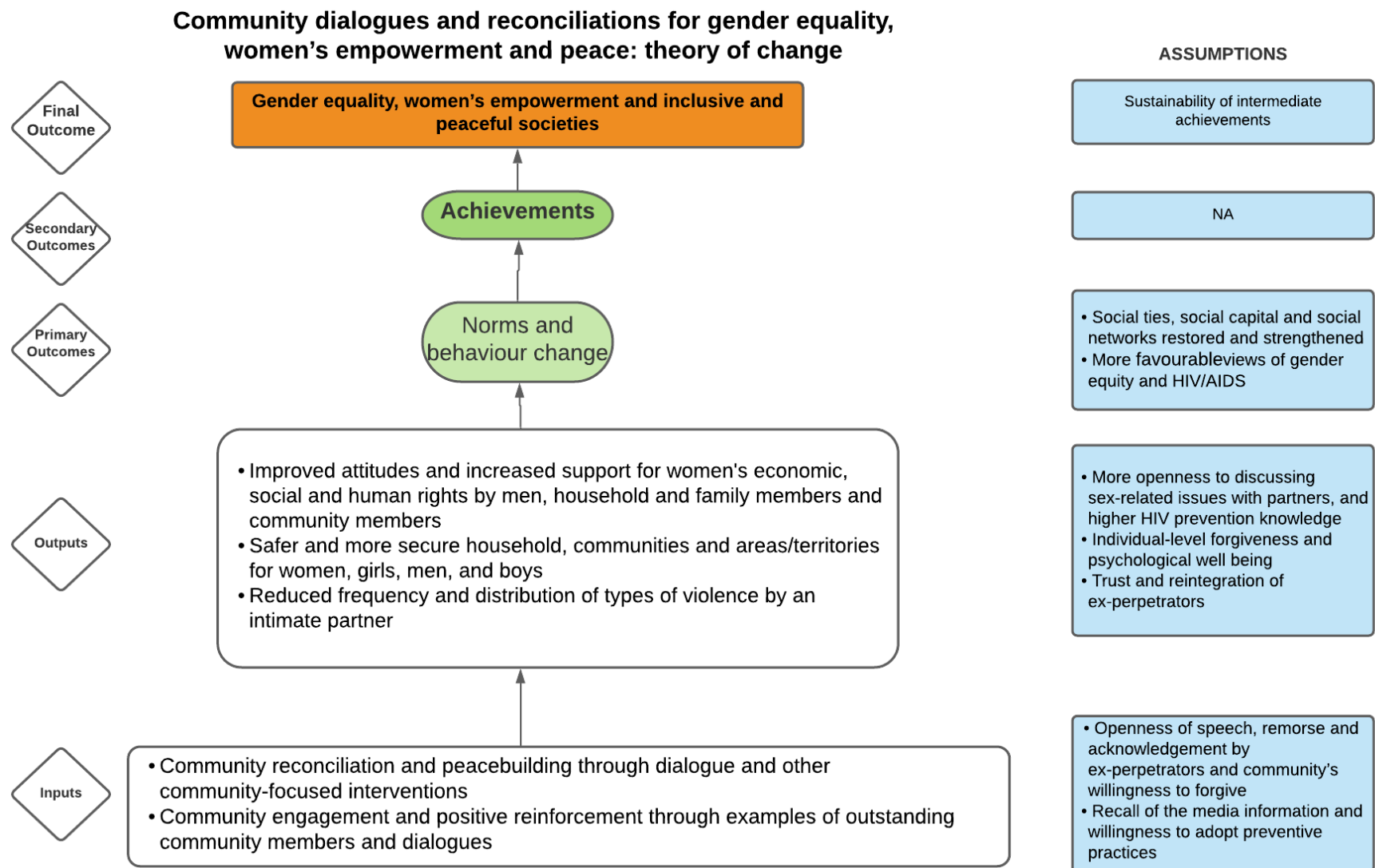
Community dialogues and reconciliations draw participants from different sections of a community and create opportunities for exchanging information and perspectives, clarifying viewpoints, and developing solutions to issues of interest to the community (Search for Common Ground 2016). They differentiate from discussion groups because they can involve groups with conflicting interests while discussion groups involve members with shared interests. They often, but not exclusively, occur in conflict/post-conflict settings where survivors, perpetrators and community members are brought together in order to increase inter-group trust. Communal dialogue projects are influenced by the contact theory and encompass expected changes at both personal level (empathy, trust) and intergroup levels (positive behaviours, intergroup respect and trust) (CDA Collaborative Learning 2019).

How do community dialogues and reconciliations affect gender equality, women's empowerment and peace outcomes?

Figure 126 below maps out a causal chain of how community dialogues and reconciliations may improve gender equality, women's empowerment and peace outcomes. Cultural norms are hard to break because pressures to conform with the acceptable, and the avoidance of social isolation prevent actors from expressing divergent/deviant views (Scheufele and Moy 2000). New social norms can emerge by a process of normative influence through communication networks (Kincaid 2004). The opportunity for exchange of information on a certain topic accelerates convergence, a new view of the world, a critical consciousness among participants. This consciousness is a powerful agent of change for a new way of thinking that can evolve into new social norms (Kincaid 2000). Dialogue groups and the media are means by which exposure to such communication can take place.

As shown in the figure below, achieving required changes in gender-related behaviour and norms is a high-level achievement for affected communities, and it can be made possible through community dialogues that foster an improved attitude and increased support for women's rights, reduce violence by intimate partners and create safer households and communities, when the necessary community- and individual-level acceptance and networks are in place.

Figure 126: Community dialogues and reconciliations for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included two studies reported in two different papers that evaluated the effects of two programmes. Each of the papers evaluated a different programme carried out in a different country.

Population

Both programmes evaluated involved men and women, boys and girls at the community level. The included studies evaluated two different programmes and trials in Sub-Saharan Africa, including Mozambique and Sierra Leone.

Intervention, inputs and activities

The included studies evaluated a range of different community dialogues and reconciliation activities including:

- Community reconciliation and peacebuilding (n = 1): This intervention occurred at the level of sections, which are clusters of up to 10 contiguous villages. Two groups are established: a reconciliation committee consisting of village chiefs, religious and youth leaders, as well as some war survivors and former combatants; and an outreach committee, consisting mostly of youths. The reconciliation forum consists of a two-day bonfire ceremony, at which victims share their experiences and perpetrators ask for forgiveness for their war crimes. It is capped by a 'cleansing' ceremony conducted by Sierra Leone's 'secret societies', where perpetrators who have expressed remorse are cleansed of their sins. After the ceremony, Fambul Tok (FT) also establishes a series of local institutional structures to further heal the community. It sets up a Peace Tree, as a symbolic gesture, which provides a focal point for resolving disputes. In some treated areas, it also creates communal farms on land set aside as a pledge towards reconciliation. Finally, it helps establish a Peace Mothers group to promote women's economic activities and discuss gender-targeted atrocities perpetrated during the war.
- Community engagement and positive reinforcement through examples of outstanding community members and dialogues (n = 1): The programme had two main components: facilitated community dialogues (Tchova Tchova Histórias de Vida: Diálogos Comunitários (TTHV) sessions) using the African Transformation gender tool adapted to HIV/AIDS prevention and the TT radio magazine. The tool included nine video and written profiles of real Mozambican trendsetters. In the profiles, the men, women, and couples tell their stories of how they overcame gender, cultural, and social barriers, such as domestic violence, alcohol abuse, and the subordination of women, to make positive changes in their lives that impacted HIV treatment and prevention. To expand the reach of TTHV activities, the TT radio magazine provided reinforcing messages and featured testimonies of TTHV participants who had made changes in their lives and modelled positive behaviours. The radio magazine included 34 programmes of 12 minutes each, in local languages, covering topics such as condom use, sexual networks, domestic violence, dialogue on HIV between partners, HIV stigma, and gender and HIV, among others.

Table 46: Community dialogues and reconciliation activities design features

Study	Activity/Input	Length of treatment	Intervention frequency
Cilliers (2015)	Community reconciliation and peacebuilding through dialogue and other community-focused interventions	5 months	Weekly
Figueroa (2016)	Community engagement and positive reinforcement through examples of outstanding community members and dialogues	18 months	Weekly

Comparison

All included studies compared treated groups to comparison groups receiving no intervention. One study included multiple treatment arms.

Outcomes

The included studies reported on a number of relevant outcomes, including the following:

- Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members (n = 2): These positive attitudes can be shaped by sensitisation and education to shift social norms, particularly relating to marginalised groups. As a result, they are aware of the specific needs of these groups and take an active part in reducing inequalities.
- Communities have a more positive attitude towards women/marginalised groups (n = 1): Communities are informed and more aware of issues, characteristics and actions for women/marginalised groups and are able to act on these issues. This includes media actors (traditional and internet) publish more pieces (multimedia inclusive) focusing on these issues through gender-transformative and gender-specific lens.

Both of the studies (n = 2) measured norms and behaviour through improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members.

One of the studies (n = 1) also measured norms and behaviour through safer and more secure households, communities and areas/territories for women, girls, men, and boys; as well as through reduced frequency and distribution of types of violence by an intimate partner.

The division of the immediate and secondary outcomes in asset transfer is reported in Table 47 below:

Table 47: Summary of secondary and immediate outcomes for community dialogues and reconciliations

Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	0
	Economic and livelihood related resources	0
	Access to employment	0
Agency ability to define one's goals and act upon them and operationalised decision-making;	Individual agency	0
	Community level agency	0
	Institutions supporting agency	0
Achievement ways of being and doing which can be realised by different individuals	Improved systems	0
	Norms and behaviour change	2
	Empowerment index	0

Study design

One of the community dialogues and reconciliation studies used a QED (Figueroa 2016) which had some concerns related to risk of bias. The authors do not test if the results are sensitive to the existence of hidden bias, nor do they use different matching methods or report sensitivity or robustness checks, so that we could not justify the absence of confounding bias or reporting bias. However, we did not identify any major limitation and we assessed the study as having a low risk of selection bias, performance bias and outcome measurement issues.

The other community dialogues and reconciliation study used an experimental design (Cilliers 2015) which had a high risk of selection bias because of issues with the implementation that led to spill-overs. We did not identify any other issue or limitation related to any of the other domains of the RoB (assignment mechanism, unit of analysis, confounding, performance, outcome measurement and analysis reporting). Random assignment was based on a public lottery and baseline characteristics were balanced between treatment and control and the authors report low rates of attrition.

Qualitative studies, process evaluations and project documents

We identified only one additional document related to one programme.

- Fambul Tok's reconciliation programme (Sierra Leone): one descriptive quantitative study

This single document was ranked as high empirical quality during the appraisal process. We identified only one linked qualitative study and were unable to conduct a qualitative evidence synthesis related to community dialogues.

Synthesis of findings

The following subsection presents the results of effectiveness of dialogue groups and reconciliations on gender equality, women's empowerment and peace outcomes.

Quantitative findings

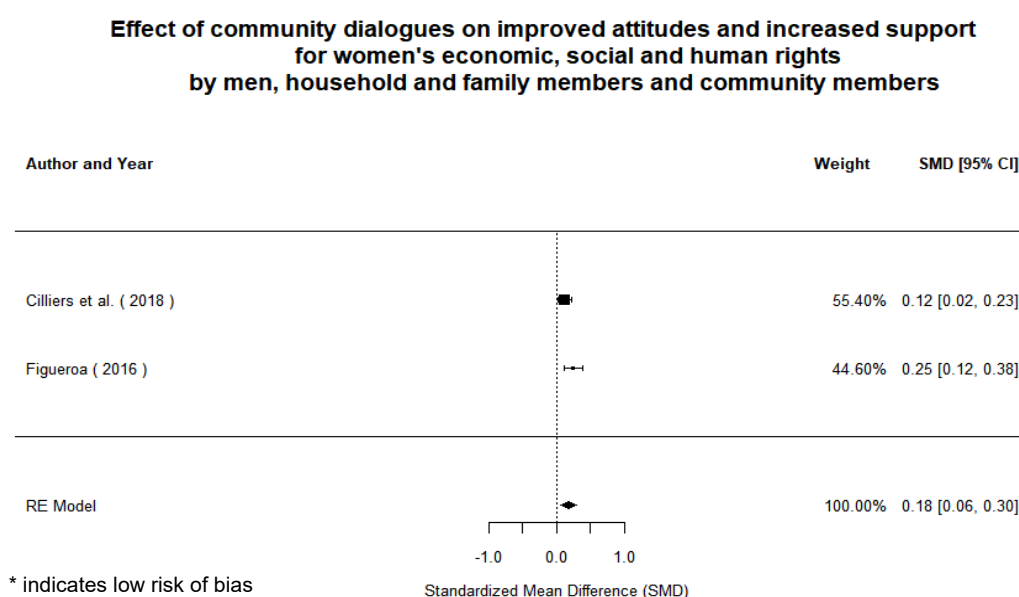
Effects of community dialogues and reconciliation on communities having a more positive attitude towards women/marginalised groups

Cilliers and colleagues' (2018) experimental study in Sierra Leone was the only study evaluating the impact of community dialogues and reconciliations on communities having a more positive attitude towards women/marginalised groups. Their report included two effects that fell into this outcome category (i.e., index of attitudes towards women and attitudes towards wife beating). The effects were both very small, positive and not statistically significant point estimates ($g = 0.09$, [95% CI: -0.01 to 0.19]) and ($g = 0.04$, [95% CI: -0.07 to 0.14]). We assessed the study as having high risk of bias.

Effects of community dialogues and reconciliation on improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members

Two studies reported disaggregated data for improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as having some concerns of risk of bias and the other as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.18$ (95% CI: 0.06 to 0.30). Therefore, the average outcome was significantly different from zero ($z = 2.89$, $p = 0.004$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 127 CDCB4. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 2.19$, $p = 0.139$, $\hat{\tau}^2 = 0.0042$, $I^2 = 54.33\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid.

Figure 127: CDCB4. Forest plot showing the observed outcomes and the estimate of the random-effects model



Summary of findings and discussion

We included two studies in two Sub-Saharan African countries that evaluated the effect of community dialogues and reconciliations. There was a paucity of evidence for this intervention category, thus we were only able to complete a quantitative synthesis for one outcome. We found a small positive but significant effect on improved attitudes and increased support for women's economic, social and human rights by men/households/community members. Overall, the GRADE assessments generally indicate a very low to low certainty in this body of evidence.

We identified only one linked qualitative study for community dialogues and were unable to conduct a qualitative evidence synthesis related to this intervention. Table 48 below presents the summary of our quantitative findings along with the GRADE certainty of evidence ratings:

Table 48: GRADE summary of findings and certainty of evidence on community dialogues and reconciliations

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(CB2) Communities have a more positive attitude towards women/marginalised groups.										
1	QED	very serious ^c	serious ^d	not serious	serious	none	2982	Two positive effect estimates with a 95% CI range of -0.01 to 0.19	⊕○○○ VERY LOW	Limited Importance
(CB4) Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members										
2	RCT - 1 QED - 1	very serious ^a	not serious	not serious	not serious ^b	none	3872	SMD 0.18 SD higher (0.06 higher to 0.3 higher)	⊕⊕○○ LOW	Limited Importance

Note. CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

- a. This analysis has been downgraded twice due to high risk of bias in both studies. This includes one study rated as very high risk that selection bias and deviations from intended interventions.
- b. Downgraded once to account for a p-value of 0.139.
- c. Downgraded to very serious because of significant issues related to deviation from intended intervention and other criteria.
- d. All single studies downgraded once for inconsistency.

4.6 Effects of interventions under multiple UNSCR pillars: participation and protection

Multi-pillar interventions intersect across multiple pillars. In our study, multi-pillar interventions both targeted the participation and protection pillars:

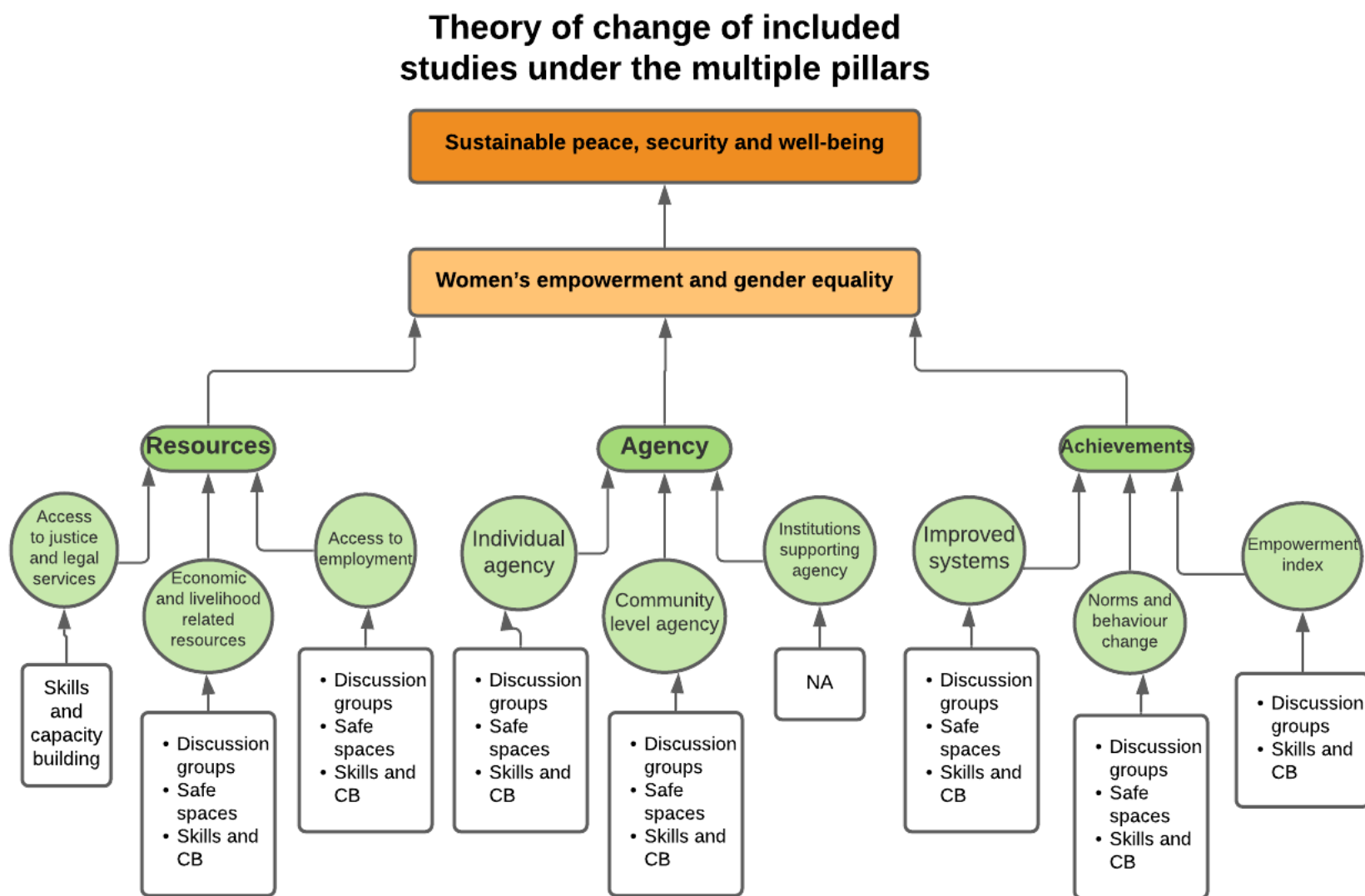
- The participation pillar includes interventions that create opportunities for, build acceptance of, or strengthen capacities for the equal participation and full involvement of women and girls in political, economic and social institutions and decision-making processes.
- The protection pillar includes interventions that create, facilitate access to, or build awareness of and support for legal or social protections for women's and girls' rights. This also includes behavioural, legal and environmental interventions that aim to reduce women and girls' risk of experiencing SGBV.

In our systematic review, this pillar gathered the following interventions covering 47,232 beneficiaries:

- Discussion groups
- Safe spaces
- Life, social and livelihood skills and capacity building

This section provides the findings of our synthesis of the 30 included studies evaluating the effect of these interventions on gender equality, women's empowerment and peaceful and inclusive societies. The section is organised by each intervention group. Each sub-section starts with a description of the intervention groups, their activities and theory of change, followed by descriptive results and the findings addressing our research questions. Figure 128 below provides a summary of the outcomes and targeted effects of the interventions included under the multiple pillars:

Figure 128: Theory of change of the included studies under multiple pillars



4.6.1 Discussion groups

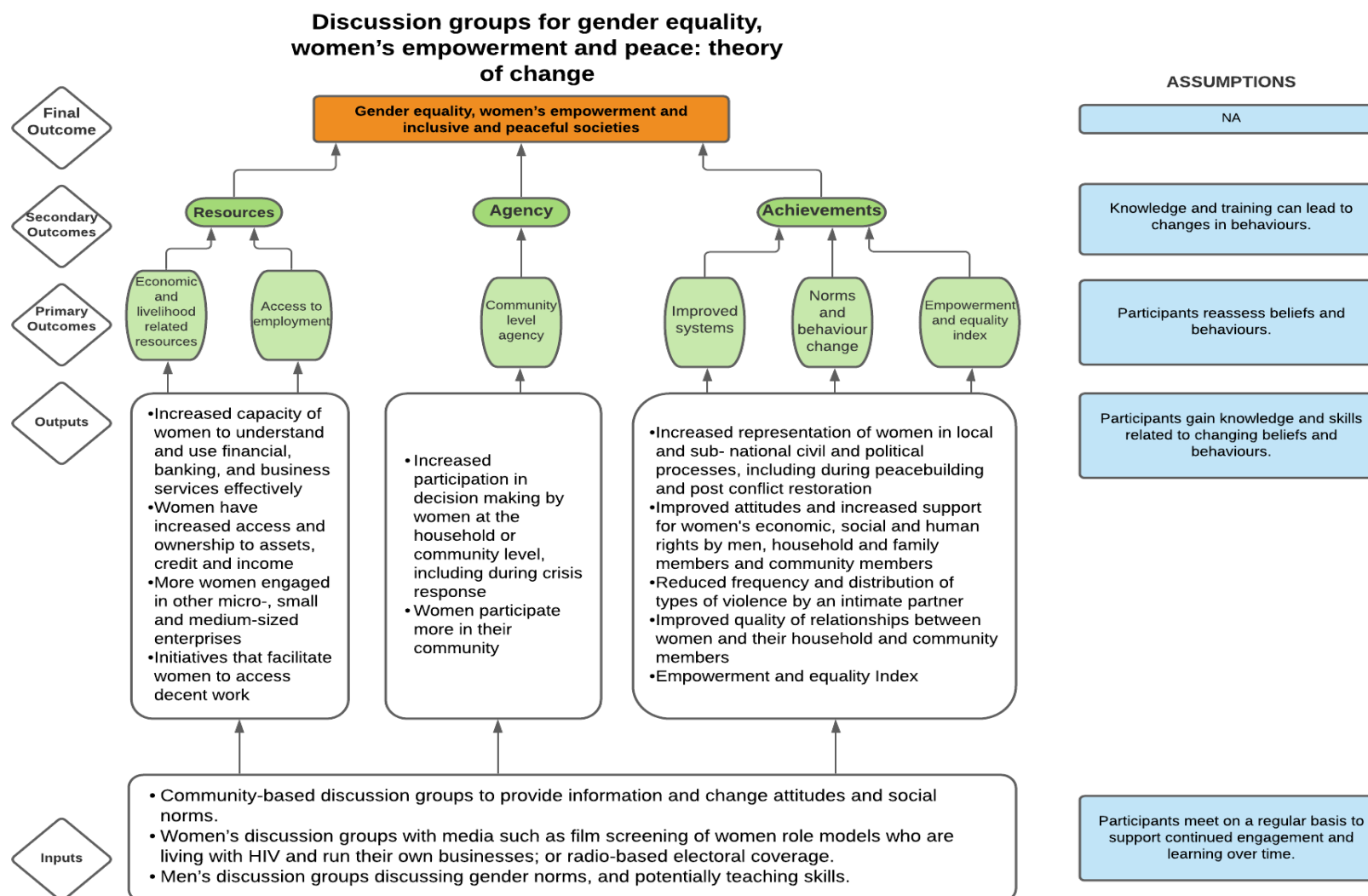
Discussion groups are meetings, facilitated by a leader, amongst community members with common interests or characteristics which aim to expand understanding, improve attitudes, or change behaviours through conversation. The emphasis is on people with social capital, authority and influence, whose views and behaviours, if directed at reconciliation, can result in positive change for the wider community (CDA Collaborative Learning 2019).

How do discussion groups affect gender equality, women's empowerment and peace outcomes?

Figure 129 below maps out the causal chain of how discussion groups may improve gender equality, women's empowerment and peace outcomes.

Engaging women and men in discussion groups that convey information through a curriculum, media, or other means would aim to improve participants' knowledge and skills. Gaining knowledge would help participants reassess their beliefs and behaviours in favour of more equitable social norms and relationships. Participants' reflection would lead to gradual and overall change in their beliefs and behaviours, and contribute to advancing gender equality, women's empowerment, and inclusive and peaceful societies. Assumptions associated with this theory of change include participants meeting on a regular basis, such as weekly or monthly, to build dialogue and engagement with the topics over time. While knowledge and training are assumed to lead to changes in behaviours at the individual level and potentially at the household and community levels, longer timeframes may be needed to capture community-level change, and interventions may need to focus beyond individual participants to change community norms. For example, a discussion group that regularly engages men about gender norms and teaches conflict management skills would aim to help participants reflect on their beliefs and identify ways that they could change their behaviours, such as improving the quality of their relationships with women in their households and/or at the community level. Changes in attitudes and beliefs about gender relations could lead to changes in behaviour and reduce incidents of SGBV, IPV and related consequences of harmful social norms, though the latter would emerge gradually. If achieved, reductions in violence could contribute to a more peaceful and equitable society.

Figure 129: Discussion groups for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included four studies from five different papers that evaluated the effect of two programmes in L&MICs. We included more than one paper that evaluated the same programme if the author(s) reported different outcomes over several papers.

Population

The targeted populations for discussion group interventions included women only, men only, women and men, or women and/or men who had undergone testing for HIV or who had tested positive for HIV. Discussion group interventions generally targeted changes in beliefs, attitudes and behaviours at the individual level, though some interventions, such as those focusing on political participation or gender relations, also targeted change at the household or community levels.

The included studies evaluated two different programmes in four countries in Sub-Saharan Africa: Côte d'Ivoire, DRC, Liberia and Uganda.

Intervention, inputs and activities

The included studies evaluated a range of different discussion group activities and inputs including the following:

- Discussion groups related to SGBV (n = 3): Groups generally separated by gender convened to discuss attitudes and behaviours about SGBV and IPV, gender relations and gender rights. Some groups also received skills training.
- Discussion groups with radios (n = 1): Women were provided radios and listened to electoral coverage.
- Short films at HIV/AIDS clinics (n = 1): Women at these clinics participated in discussion groups and watched short films featuring women entrepreneurs living with HIV.

Table 49: Discussion group activities design features of included studies

Study	Activity/Input	Length of treatment	Intervention frequency
Wagman et al. – 2015	Discussion groups related to SGBV	35 months	NA
Vaillant et al. – 2020	Discussion groups related to SGBV	NA	Weekly
Vaillant et al. – 2016	Discussion groups related to SGBV	NA	NA
Hossain et al. - 2014	Discussion groups related to SGBV	5 months	Weekly
Mvukiyehe - 2017	Discussion groups with radios	NA	Weekly
Lubega et al. – 2017	Short films at HIV clinics	10 months	Monthly

Comparison

All the included studies compared treated groups to comparison groups receiving no intervention.

Outcomes

The included studies reported on a number of relevant outcomes, including:

- Women have increased access and ownership to assets, credit and income (n = 2): Women are able to apply for, receive and manage assets/credit and income and have support to manage, claim and execute their assets without pressure or influence from external actors, including male family members, husbands and cultural leaders.
- Increased participation in decision making by women at the household or community level, including during crisis (n = 2): Women take part in all or any step of the decision-making process at the household community and district level, but also are able to meaningfully take part and have influence on the final decision, including in crisis response.
- Improved attitudes and increased support for women's economic, social and human rights by men, household and family (n = 2): These positive attitudes can be shaped by sensitisation and education to shift social norms, particularly relating to marginalised groups. As a result, they are aware of the specific needs of these groups and take an active part in reducing inequalities.
- Reduced frequency and distribution of types of violence by an intimate partner (n = 1): Types of violence in this case include four types of IPV: physical violence, sexual violence, stalking, and psychological aggression. These incidences are reduced over time due to a bundle of interventions that focus on both support for survivors and sanctioning perpetrators.

The separation of the immediate and secondary outcomes is reported in Table 50 below:

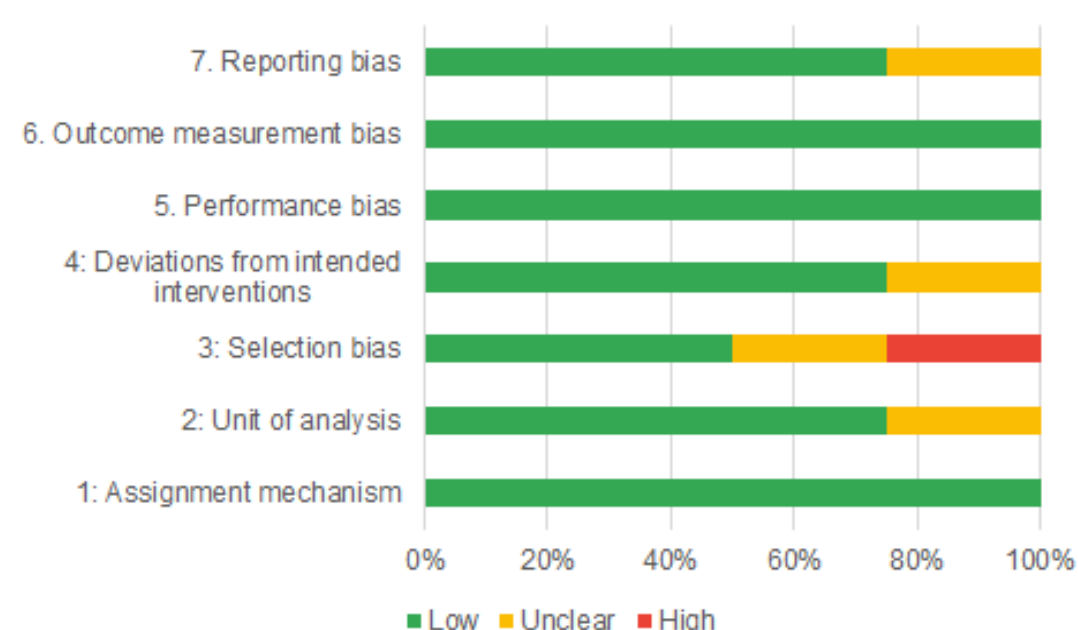
Table 50: Summary of secondary and immediate outcomes for discussion groups

Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	0
	Economic and livelihood related resources	2
	Access to employment	1
Agency ability to define one's goals and act upon them and operationalised decision-making;	Individual agency	1
	Community-level agency	2
	Institutions supporting agency	0
Achievement ways of being and doing which can be realised by different individuals	Improved systems	1
	Norms and behaviour change	2
	Empowerment index	1

Study design

The four included studies evaluating discussion groups used an experimental design (Hossain 2014; Lubega 2017; Mvukiyehe 2017; Vaillant 2020). One of the four studies was assessed as having a high risk of bias. As detailed in Figure 130 below, high risk of bias was identified in the selection domain because of a large difference in attrition rates between the treatment and control groups (Hossain 2014). Two studies were assessed as having some concerns of risk of bias. The first did not provide enough information for us to justify low risk of deviations from intended interventions, no mention of spill-overs, crossovers or contamination was observed, and no information of geographical dispersion that may have dissuaded implementation issues (Vaillant 2020). We assessed another study as unclear on selection and reporting bias because the authors do not discuss attrition and do not provide enough information to justify the absence of selective reporting of the analysis and the results. We did not identify any limitations related to the assignment mechanics, performance and behavioural changes in participants and not participants relate to the knowledge of being observed and compared, or outcome measurement.

Figure 130: Discussion Groups studies RCT RoB assessment



Qualitative studies, process evaluations and project documents

We identified two additional documents related to one programme covered by the discussion group section of included studies.

- Engaging Men through Accountable Practice (EMAP) (DRC): one descriptive quantitative and one qualitative study

One study was marked as high empirical quality and the other was marked as moderate. We identified only two qualitative studies and were unable to conduct a qualitative evidence synthesis related to the discussion group intervention.

Synthesis of findings

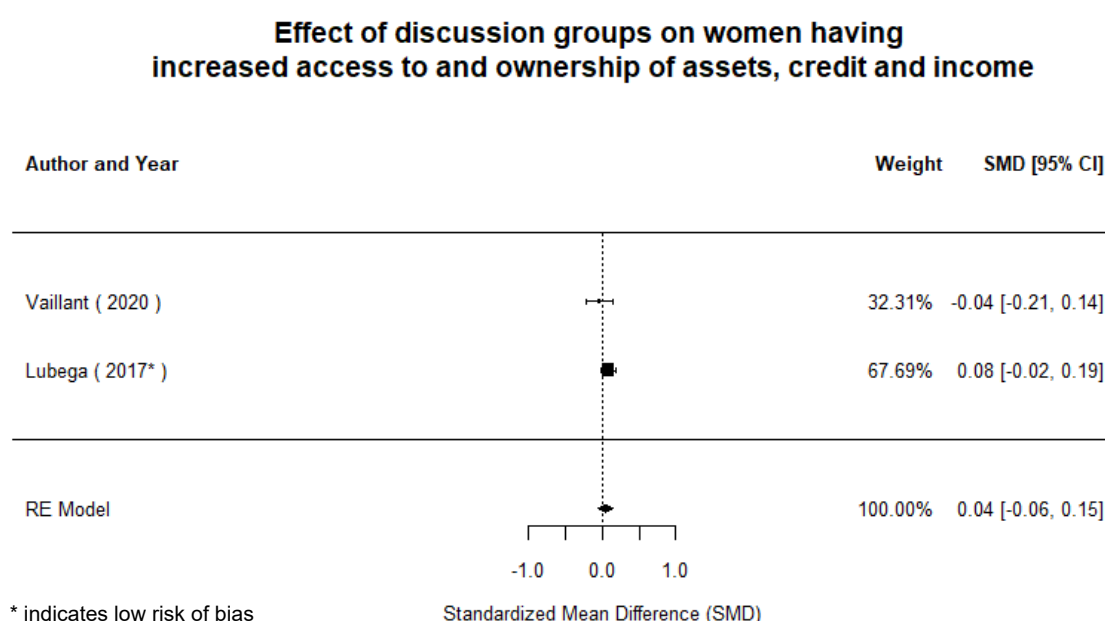
The following subsection presents the results of effectiveness of discussion groups on gender equality, women's empowerment and peace outcomes.

Quantitative findings

Effects of discussion groups on women having increased access to and ownership of assets, credit, and income

Two studies examined the effects of discussion groups on women having increased access to and ownership of assets, credit, and income, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as low risk of bias and the other as some concerns. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.04$ (95% CI: -0.06 to 0.15). Therefore, the average outcome did not differ significantly from zero ($z = 0.80$, $p = 0.42$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 131 DGAB2. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 1.37$, $p = 0.24$, $\hat{\tau}^2 = 0.002$, $I^2 = 26.92\%$). One study was appraised as having *some concerns* with risk of bias (Vaillant 2020) and the other (Lubega 2017) was assessed as having a low risk of bias. With only two studies, moderator analyses were not possible and tests of publication bias are not valid.

Figure 131: DGAB2. Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of discussion groups on engaging more women in other micro, small, and medium-sized enterprises

Lubega's (2017) experimental study in Uganda was the only study evaluating the impact of discussion groups on increasing women's engagement in other micro, small, and medium-sized enterprises. Their report included three effects that fell into this outcome category (e.g., whether a woman operates an enterprise or open a retail shop). The effects ranged from a null point estimate ($g = 0.00$, [95% CI: -0.10 to 0.10]) for whether a woman operates an enterprise to a small, positive point estimate ($g = 0.13$, [95% CI: -0.13 to 0.40]) for whether women open a retain shop among the subgroup of women who start a business. We assessed the study as having low risk of bias.

Effects of discussion groups on whether women have more positive attitude towards taking action to claim their rights.

Mvukiyehe's (2017) experimental study in Liberia was the only study evaluating the impact of discussion groups on whether women have more positive attitudes towards taking action to claim their rights. Their report included 19 effects that fell into this outcome category (e.g., such as attending strategic meetings or strategic voting). The effects ranged from a large negative effect ($g = -0.52$, [95% CI: -0.74 to -0.30]) for an overall index measuring political behaviour to a small, positive point estimate ($g = 0.15$, [95% CI: -0.07 to 0.36]) for whether women had their ballot marked by someone else. We assessed the study as having some risk of bias concerns.

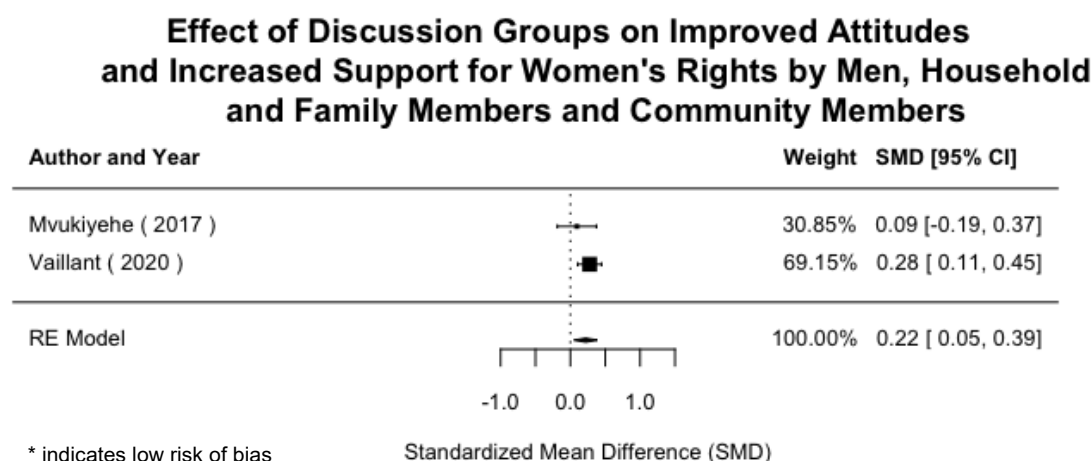
Effects of discussion groups on increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration.

Mvukiyehe's (2017) experimental study in Liberia was the only study evaluating the impact of discussion groups on increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration. There was a large, and statistically significant, negative effect ($g = -0.47$, [95% CI: -0.76 to -0.18]), and we assessed the study as having some risk of bias concerns.

Effects of discussion groups on improved attitudes and increased support for women's rights by men, household and family members and community members

Two studies examined the effects of discussion groups on improved attitudes and increased support for women's rights by men, household and family members and community members, thus we included $k = 2$ studies in the analysis (see Table 51). We assessed both of the studies as having some concerns of risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.2199$ (95% CI: 0.0542 to 0.3857). Therefore, the average outcome differed significantly from zero ($z = 2.6003$, $p = 0.0093$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 132 DGCB4. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 1.19$, $p = 0.27$, $\hat{\tau}^2 = 0.0027$, $I^2 = 16.20\%$). Both studies were appraised as having *some concerns* with risk of bias and these results should be applied cautiously. With only two studies, moderator analyses were not possible and tests of publication bias are not valid.

Figure 132: DGCB4. Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of discussion groups on an improved quality of relationships between women and their household and community members.

Vaillant 's (2020) experimental study in Democratic Republic of Congo was the only study evaluating the impact of discussion groups on improved quality of relationships between women and their household and community members. There was a small, but not statistically significant, point estimate ($g = 0.15$, [95% CI: -0.02 to 0.32]), and we assessed the study as having some risk of bias concerns.

Effects of discussion groups on women's empowerment index

Mvukiyehe's (2017) experimental study in Liberia was the only study evaluating the impact of discussion groups on women's empowerment index. There was a small, but not statistically significant, point estimate ($g = -0.11$, [95% CI: -0.39 to 0.17]), and we assessed the study as having some risk of bias concerns.

Summary of findings and discussion

We included four studies in four Sub-Saharan African countries that evaluated the effect of discussion groups. Only two outcomes had sufficient evidence to conduct a quantitative synthesis. There was a significant positive effect of discussion groups on improved attitudes and increased support for women's economic, social and human rights by men/household/community members. There was a positive but non-significant point estimate for of discussion groups on women having increased access to and ownership of assets, credit and income. Overall, the GRADE assessments generally indicate a very low to moderate certainty in this body of evidence.

We identified no linked qualitative studies and were unable to conduct a qualitative evidence synthesis related to the discussion groups intervention. The summary of our quantitative findings along with the GRADE certainty of evidence ratings are presented in Table 51.

Table 51: GRADE summary of findings and certainty of evidence on discussion groups

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations				
(AB2) Women have increased access and ownership to assets, credit and income.										
2	RCT – 2	not serious	serious ^a	not serious	not serious	none	2515	SMD 0.04 SD higher (0.06 lower to 0.15 higher)	⊕⊕⊕○ MODERATE	Important, but not critical
(AC1) More women engaged in other micro, small, and medium-sized enterprises										
1	RCT	not serious	serious ^b	not serious	not serious	none	1502	One null effect and two positive effect estimates with a 95% CI range of -0.12 to 0.39	⊕⊕⊕○ MODERATE	Important, but not critical
(BA5) Women have more positive attitude towards taking action to claim their rights.										
1	RCT	serious ^c	serious ^b	not serious	not serious	none	190	SMD 0.19 SD higher (0.09 lower to 0.47 higher)	⊕⊕○○ LOW	Limited importance
(CA2) Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration										
1	RCT	serious ^d	serious ^b	not serious	serious ^e	none	190	One null, 4 negative, and 15 positive effect sizes with a 95% range of -0.76 to 0.37	⊕○○○ VERY LOW	Limited importance
(CB4) Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members										
2	RCT – 2	not serious	serious ^f	not serious	serious ^g	none	1245	SMD 0.22 SD higher (0.05 higher to 0.39 higher)	⊕⊕○○ LOW	Limited importance
(CB8) Improved quality of relationships between women and their household and community members										
1	RCT	serious ^h	serious ^b	not serious	serious ⁱ	none	1051	SMD 0.15 SD higher (0.02 lower to 0.32 higher)	⊕○○○ VERY LOW	Limited importance
(CC1) Empowerment/Equality index										
1	RCT	serious ^j	serious ^b	not serious	serious ^k	none	194	SMD 0.11 SD lower (0.39 lower to 0.17 higher)	⊕○○○ VERY LOW	Important, but not critical

Note. CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

- a. Although the point estimates are quite close to null, this has been downgraded because of wide confidence intervals suggesting inconsistency.
- b. Downgraded once for one study in the group.
- c. Downgraded because the study presents uncertain risk of bias with regards to selection bias and reporting bias.
- d. Downgraded because the study presents uncertain risk of bias with regards to selection bias and reporting bias.
- e. Downgraded because this single study reports a wide range of effect sizes measuring this outcome, including on both sides of the threshold.
- f. Downgraded to reflect significantly different point estimates whose CIs, despite overlapping, are very wide.
- g. Downgraded because the point estimate's CIs are wide and vary considerably relative to one of the studies in the analysis.
- h. Downgraded because the study presents uncertain risk of bias due to deviation from the intended intervention.
- i. Downgraded to reflect the very wide range of CIs surrounding the point estimates.
- j. Downgraded because the study presents uncertain risk of bias with regards to selection bias and reporting bias.
- k. Downgraded to reflect the very wide range of CIs surrounding the point estimates.

4.6.2 Safe spaces

Safe spaces are formal or informal gathering places where women and girls feel physically and emotionally safe, access services, participate in training and skills development activities, join peer support networks, mentorship etc. Safe spaces are group-centred mechanisms to articulate and accommodate the preferences of women and the girl-child. Safe spaces provide participants with a welcoming atmosphere that can enhance their opening up on sensitive issues as particular to them, individually and/or collectively. Safe spaces shoulder the responsibility of knowledge dissemination through developmental activities and mentoring (UNFPA 2015).

How do safe spaces affect gender equality, women's empowerment and peace outcomes?

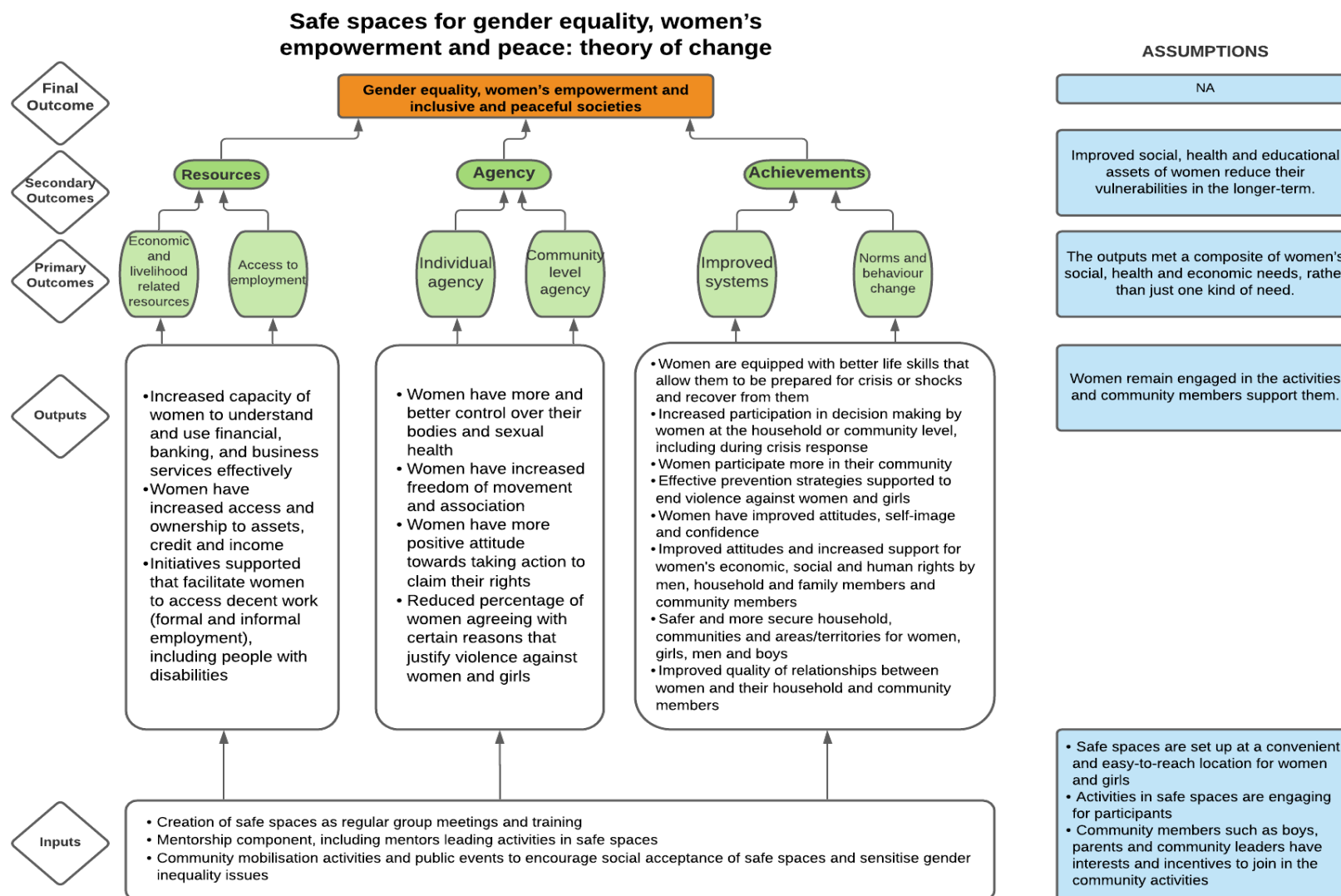
Figure 133 below maps out the causal chain of how safe spaces may improve gender equality, women's empowerment and peace outcomes.

The main inputs are safe space programmes that involve regular group sessions for education, discussions and peer networking activities. Curriculums mainly focus on literacy, sexual and reproductive health, financial education and other life skills. The activities are typically led by mentors from local communities, who support and build trust with programme participants within and outside of safe spaces. In some cases, programme components include activities for community members such as boys, parents and community leaders to sensitise gender inequality issues and encourage their support and acceptance of safe spaces for women. The assumptions are that locations of safe spaces are easily accessible by women, and both safe space participants and community members have sustained interest and incentives to participate in the activities.

The outputs of these activities include women's improved knowledge on a range of topics such as life skills, greater access to assets and more decent work, better control over their bodies and sexual health, self-efficacy, and freedom of movement and participation in their households and communities. The community members will improve their attitudes towards women's economic, social and human rights. Relationships between women and their household and community members will be enhanced. It is crucial that participants stay engaged in the safe space programmes, and community members keep supporting them.

These will result in immediate outcomes, such as improved access to economic and livelihood related resources, access to employment, individual and community level agency, improved systems, and norms and behaviour change. It is assumed that the outputs meet a combination of women's social, health, and economic assets rather than one kind of asset only. For example, it was claimed that only improving girls' knowledge of sexual and reproductive health itself is insufficient to yield any expected outcomes, as their vulnerable economic situation overshadows their knowledge (Austrian and Muthengi 2014). Finally, these immediate outcomes will catalyse women's improved resources, agency and achievements, which will then lead to better gender equality, empowerment and inclusive and peaceful societies.

Figure 133: Safe spaces for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included six studies reported in six different papers that evaluated the effect of five programmes. We included more than one paper that evaluated the same programme if the author(s) reported different outcomes over several papers.

Population

All programmes targeted adolescent girls aged between 10-19, except the *Powering up Biruh Tesfa* programme in Ethiopia where girls under 10 years old were included (Erulkar and Medhin 2017). Around half of the programmes explicitly targeted girls in low-income areas (n = 3). Two programmes in Ethiopia and Egypt targeted out-of-school girls only (Erulkar and Medhin 2017; Sieverding and Elbadawy 2016). One programme in Zambia targeted only girls who had never married at baseline (Austrian et al. 2020).

The included studies evaluated five different programmes and trials in MENA and Sub-Saharan Africa, covering six countries including: Egypt, Ethiopia, Niger, Tanzania, Uganda and Zambia.

Intervention, inputs and activities

The included studies evaluated a range of different safe space activities and inputs including:

- Safe spaces with access to economic and/or health assets (n = 4): Safe space programmes that involve regular group sessions and training led by mentors, with the provision of vouchers for free or subsidised health services, and/or assistance to open savings accounts for girls.
- Safe spaces with community activities (n = 2): Safe space programmes that involve regular group sessions and training led by mentors, with community activities (e.g., meetings, seminars or public events) to encourage acceptance of safe spaces and sensitisation of gender inequality issues within communities.

Table 52: Safe spaces activities and design features of included studies

Study	Activity/Input	Length of treatment	Intervention frequency
Austrian and Muthengi - 2014	Safe spaces with regular group meetings, training, mentoring, provision of savings accounts	12 months	Weekly
Austrian et al. - 2020	Safe spaces with regular group meetings, training, mentoring, provision of savings accounts and vouchers for free health services	24 months	Weekly
Buehren et al. -2017	Safe spaces with recreational activities and education, mentoring, community activities (e.g., periodic meetings with village members)	24 months	Daily

Study	Activity/Input	Length of treatment	Intervention frequency
Erulkar and Medhin - 2017	Safe spaces with education, mentoring, provision of vouchers for free or subsidised health services	6 months	Daily
Mercy Corps - 2015	Safe spaces with regular group meetings and education, mentoring, provision of savings account and conditional lentil ration	8-20 months	Weekly
Sieverding and Elbadawy - 2016	Safe spaces with education, mentoring, community activities (e.g., seminars, public events), provision of conditional food ration	36 months	Daily

Comparison

All included studies are comparing treated groups to comparison groups receiving no intervention. Four studies included multiple treatment arms.

Outcomes

The included studies reported on a number of relevant outcomes, including the following:

- Women have increased access and ownership to assets, credit and income (n = 4): Women are able to apply for, receive and manage assets/credit and income and have support to manage, claim and execute their assets without pressure or influence from external actors, including male family members, husbands and cultural leaders.
- Women have more and better control over their bodies and sexual health (n = 2): Women are able to make free and educated decisions about their health (not included in this study), sexuality and procreation.
- Women have a more positive attitude towards taking action to claim their rights (n = 1): Women feel entitled to be engaged and given the leadership capacity and knowledge to claim their rights and take action on relevant issues. Increased self-efficacy and autonomy. Increased opportunities for women to claim their rights including as a result of education and sensitisation. Women are engaged and given the leadership capacity and knowledge to claim their rights and take action on relevant issues.
- Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them (n = 4): Women develop skills, have access to resources, and make use of these skills and resources to increase their level of resilience and be equipped to face life challenges (health, education, finance, social relations, work etc.).
- Women have improved attitudes, self-image and confidence (n = 4): Women feel more entitled to claim their rights and needs in community and change social norms and behaviours. They are aware of the importance of their status in society and are empowered to take this role and make use of the available resources to guarantee their rights.

We report the division of the immediate and secondary outcomes in Table 53 below:

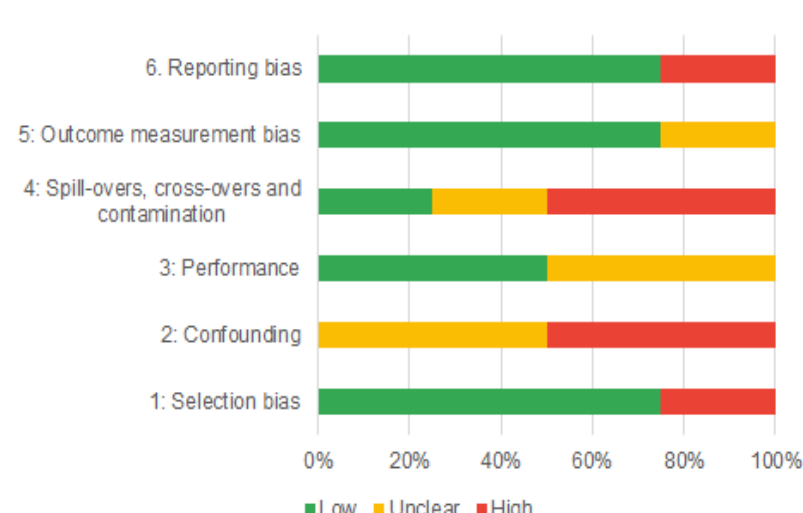
Table 53: Summary of secondary and immediate outcomes

Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	0
	Economic and livelihood related resources	4
	Access to employment	1
Agency ability to define one's goals and act upon them and operationalised decision-making;	Individual agency	5
	Community level agency	4
	Institutions supporting agency	0
Achievement ways of being and doing which can be realised by different individuals	Improved systems	2
	Norms and behaviour change	6
	Empowerment index	0

Study design

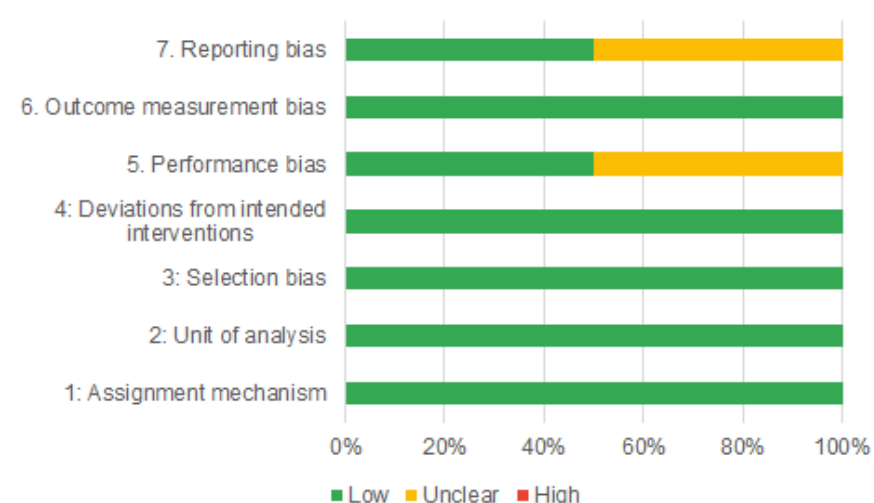
Four safe spaces studies used a QED. This included Austrian and Muthengi (2014), Erulkar and Medhin (2017), Mercy Corps (2015), Sieverding and Elbadawy (2016). We assessed three of these as having a high risk of bias. As detailed in Figure 134 below, high risk of bias was identified in four RoB domains (selection, confounding, spill-overs and reporting). Selection bias was an issue in a study using difference in differences without adjusting for any covariates (Sieverding and Elbadawy 2016), differential attrition posed risks of confounding (Sieverding and Elbadawy 2016; Austrian and Muthengi 2014), spill-overs were an issue in interventions where participants assigned to the control group received the intervention (Mercy Corps 2015), and no robustness tests were reported in one of the studies that used the difference in differences but did not match the treatment and comparison groups based on the observable characteristics, so that self-selection was not accounted for (Sieverding and Elbadawy 2016).

Figure 134: Safe spaces studies QED RoB assessment



Two safe spaces studies used an experimental design (Austrian 2020; Buehren 2017). Both studies scored some concerns related to risk of performance and reporting bias, yet we did not identify any major issue. As detailed in Figure 135 below, high risk of bias was not identified in any RoB domain of the sample of included studies, but the authors did not report enough information to justify the absence of performance bias in one study (Austrian 2020) and reporting bias in the other (Buehren 2017).

Figure 135: Safe spaces studies QED RoB assessment



Qualitative studies, process evaluations and project documents

We identified four additional documents related to three programmes covered by the safe spaces studies. Programmes are as follows:

- Ishraq (Egypt): one qualitative study and one process evaluation
- Adolescent Girls Empowerment Programme (AGEP) (Zambia): one qualitative study
- Sawki "Safe Space" and "Safe Space + Livelihoods" (Niger): one descriptive quantitative study

Half (n = 2) were marked as moderate quality and the remainder (n = 2) marked as low-quality documents.

Synthesis of findings

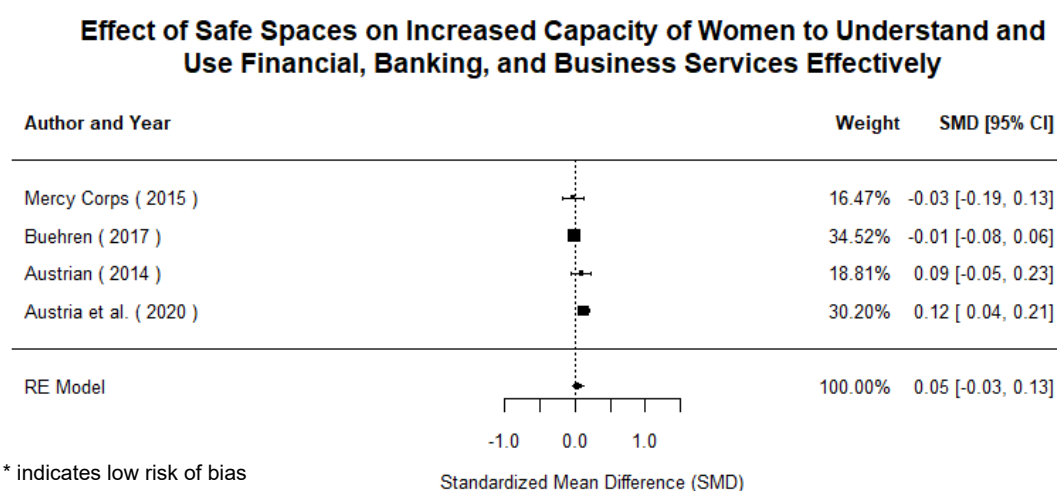
The following subsection presents the results of effectiveness of safe spaces on gender equality, women's empowerment and peace outcomes.

Quantitative findings

Effects of safe spaces on women's increased capacity to understand and use financial, banking, and business services effectively

We included a total of $k = 4$ studies in the analysis. The observed outcomes ranged from -0.03 to 0.12 . We assessed none of the studies as low risk of bias, two as some concerns, and two as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.05$ (95% CI: -0.03 to 0.13). Therefore, the average outcome did not differ significantly from zero ($z = 1.13$, $p = 0.26$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 136 SSAB1.

Figure 136: SSAB1: Forest plot showing the observed outcomes and the estimate of the random-effects model

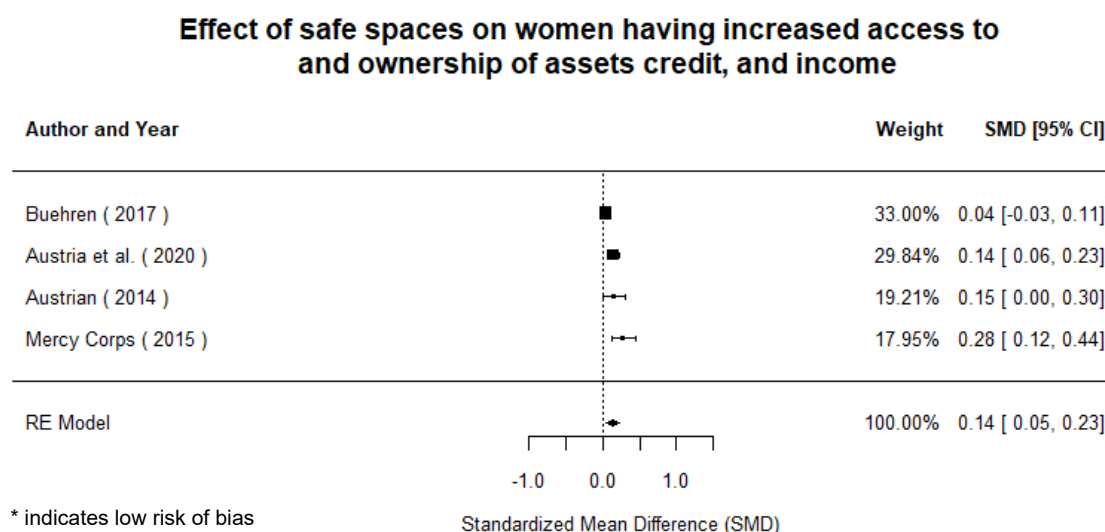


The Q -test for heterogeneity was not significant, but some heterogeneity may still be present in the true outcomes ($Q(3) = 6.78$, $p = 0.08$, $\hat{\tau}^2 = 0.00$, $I^2 = 55.77\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.50 and hence there was no sign of outliers in the context of this model. According to the Cook's distances, none of the studies could be overly influential. Given the non-significant test for heterogeneity, we did not proceed with moderator analyses.

Effects of safe spaces on women's increased access to and ownership of assets, credit, and income

We included a total of $k = 4$ studies in the analysis. We assessed none of the studies as low risk of bias, two as some concerns, and two as high risk of bias. The observed outcomes ranged from 0.04 to 0.28 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.14$ (95% CI: 0.05 to 0.23). Therefore, the average outcome differed significantly from zero ($z = 2.95$, $p < 0.01$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 137 SSAB2.

Figure 137: SSAB2: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(3) = 8.45$, $p = 0.04$, $\hat{\tau}^2 = 0.01$, $I^2 = 64.48\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.50 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be overly influential.

We found significant moderation by exposure to intervention ($\hat{\beta} = -0.007$, $p = 0.006$ [95% CI: -0.01 to -0.002]) such that each additional month of exposure decreased the impact by .007 standard deviation units. Effects did not differ by risk of bias, study design, or publication year. We were unable to test evaluation period (all were evaluated immediately after the end of the intervention), whether the model was adjusted for covariates, or change from baseline versus post-intervention changes.

Effects of safe spaces on improved capacity of women entrepreneurs

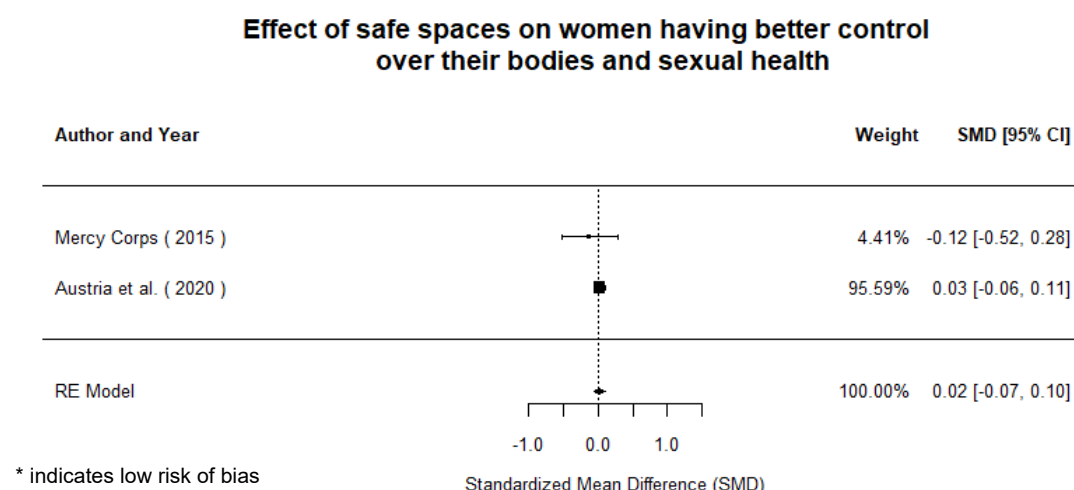
Buehren's (2017) experimental study in Tanzania was the only study evaluating the impact of safe spaces on improved capacity of women entrepreneurs. Their report included two effects that fell into this outcome category (e.g., whether the respondent indicated that she was planning to start a new income-generating activity). The effects are reported separately for each treatment village. For Treatment Village A, the effect was a very small, positive point estimate ($g = 0.02$, [95% CI: -0.05 to 0.09]), and for Treatment Village B the effect was a very small, positive point estimate ($g = 0.02$, [95% CI: -0.05 to 0.09]). We assessed the study as having some risk of bias concerns.

Effects of safe spaces on women's control over their bodies and sexual health

Only two studies reported the impact of safe spaces on women's control over their bodies and sexual health, thus we included $k = 2$ studies in the analysis. We assessed both studies as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.02$ (95% CI: -0.07 to 0.10). This positive effect did not differ significantly from zero ($z = 0.44$, $p = .66$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 138 SSBA2. Given the small number of studies, this result should be interpreted with caution.

According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.48$, $p = 0.49$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid. Neither of the studies was appraised as low risk of bias.

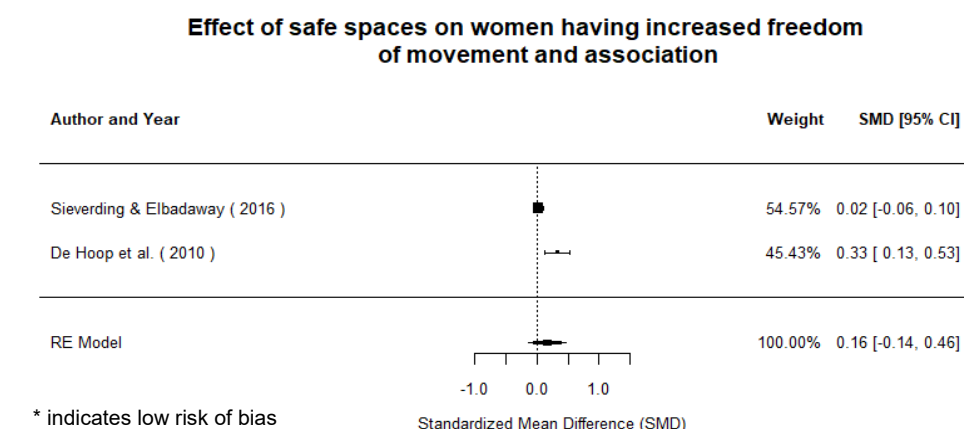
Figure 138: SSBA2: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of safe spaces on women's increased freedom of movement and association

Only two studies reported the impact of safe spaces on women's increased freedom of movement and association, thus we included $k = 2$ studies in the analysis. We assessed one of the studies as having some concerns of risk of bias and the other as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.16$ (95% CI: -0.14 to 0.29). This positive effect did not differ significantly from zero ($z = 1.05$, $p = .29$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 139 SSBA3. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was significant amount of heterogeneity in the true outcomes ($Q(1) = 7.73$, $p = 0.01$, $\hat{\tau}^2 = 0.04$, $I^2 = 87.07\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid. Neither of the studies was appraised as low risk of bias.

Figure 139: SSBA3: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of safe spaces on women's attitudes toward taking action to claim their rights

Mercy Corps' (2015) quasi-experimental study in Niger was the only study evaluating the impact of safe spaces on women having more positive attitude towards taking action to claim their rights. Their report included two effects that fell into this outcome category (e.g., whether women feel in control). The effects were for slightly different interventions. For the Sawki program (safe spaces), there was a very small, positive point estimate ($g = 0.06$, [95% CI: -0.10 to 0.22]), while for the Sawki program (safe spaces + livelihoods) the point estimate was a small and positive ($g = 0.12$, [95% CI: -0.04 to 0.28]). We assessed the study as having high risk of bias.

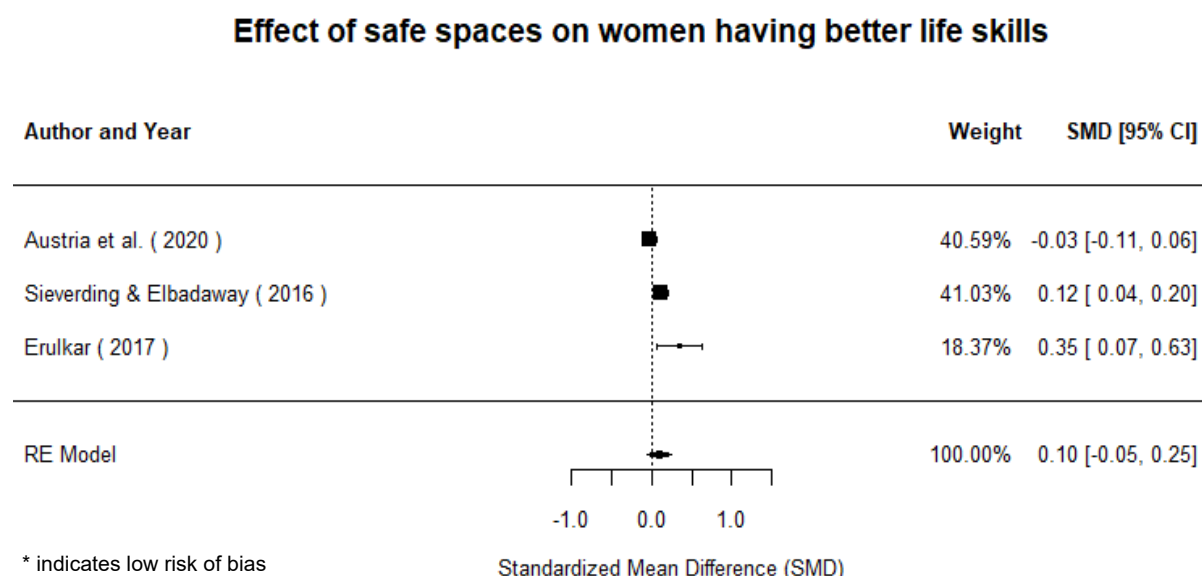
Effects of safe spaces on reduced percentage of women agreeing with certain reasons that justify violence against women and girls

Sieverding and Elbadaway's (2016) quasi-experimental study in Egypt was the only study evaluating the impact of safe spaces on reduced percentage of women agreeing with certain reasons that justify violence against women and girls. Their report included three effects that fell into this outcome category (e.g. an index asking when it is justified to beat a girl across different situations). The effects measured two outcomes in this category and at different periods after the intervention ended. The point estimates ranged from very small and negative ($g = -0.03$, [95% CI: -0.11 to 0.06]) to very small and positive ($g = 0.03$, [95% CI: -0.07 to 0.13]). We assessed the study as having high risk of bias.

Effects of safe spaces on women having better life skills

We included a total of $k = 3$ studies in the analysis. We assessed none of the studies as low risk of bias, two as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.03 to 0.35 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.10$ (95% CI: -0.05 to 0.25). Therefore, the average outcome did not differ significantly from zero ($z = 1.32$, $p = 0.19$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 140 SSBA7.

Figure 140: SSBA7: Forest plot showing the observed outcomes and the estimate of the random-effects model

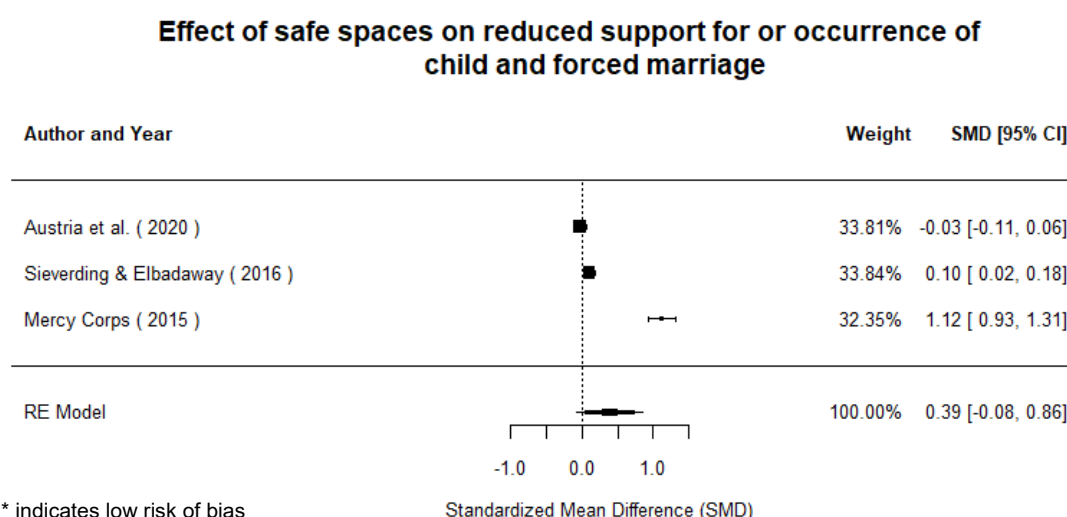


According to the Q -test, the true outcomes appear to be heterogeneous ($Q(2) = 9.87$, $p < 0.01$, $\hat{\tau}^2 = 0.01$, $I^2 = 79.74\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.39 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

Effects of safe spaces on reduced support for or occurrence of child and forced marriage

We included a total of $k = 3$ studies that examined the impact of safe spaces interventions on reduced support for or occurrence of child and forced marriage. We assessed none of the studies as low risk of bias, one as some concerns, and two as high risk of bias. The observed outcomes ranged from -0.03 to 1.12 , $\hat{\mu} = 0.39$. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.39$ (95% CI: -0.08 to 0.86). Therefore, the average outcome did not differ significantly from zero ($z = 1.62$, $p = 0.10$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 141 SSBA8.

Figure 141: SSBA8: Forest plot showing the observed outcomes and the estimate of the random-effects model



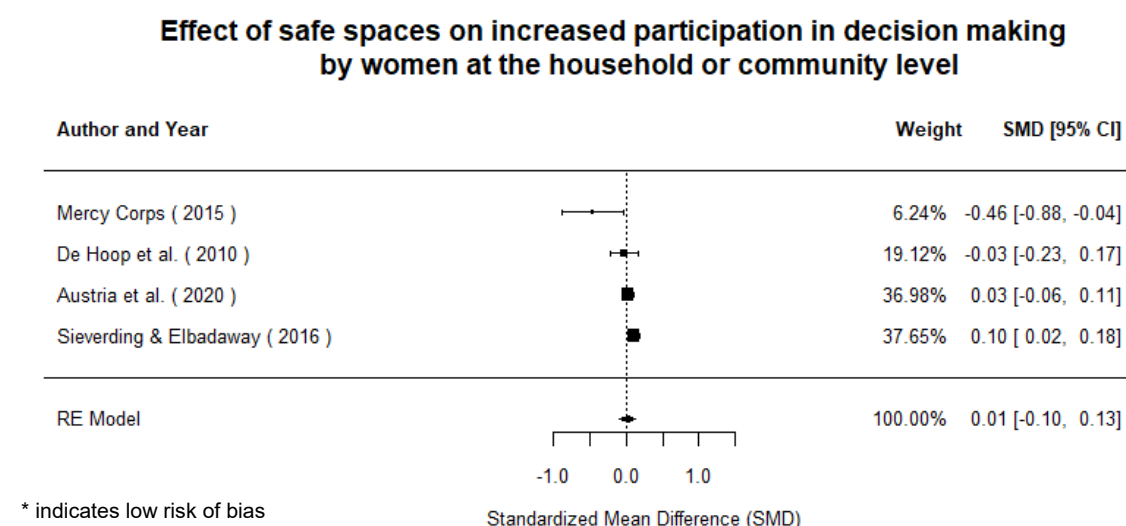
According to the Q -test, the true outcomes appear to be heterogeneous ($Q(2) = 115.54$, $p < 0.01$, $\hat{\tau}^2 = 0.17$, $I^2 = 98.27\%$). An examination of the studentised residuals revealed that one study (Mercy Corps 2015) had a value larger than ± 2.39 and may be a potential outlier in the context of this model. Indeed sensitivity analyses leaving each study out indicated that removing Mercy Corps (2015) would reduce the overall average effect ($\hat{\mu} = 0.04$ (95% CI: -0.09 to 0.16), but the effect would still be positive and non-significant ($z = 0.61$, $p = .54$). According to the Cook's distances, none of the studies could be overly influential. With only three studies, moderator analyses were not appropriate.

Effects of safe spaces on increased participation in decision making by women at the household or community level

We included a total of $k = 4$ studies in the analysis. We assessed none of the studies as low risk of bias, two as some concerns, and two as high risk of bias. The observed outcomes ranged from -0.46 to 0.10 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.01$ (95% CI: -0.10 to 0.13). Therefore, the average

outcome did not differ significantly from zero ($z = 0.24$, $p = 0.81$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 142 SSBB1.

Figure 142: SSBB1: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(3) = 7.96$, $p = 0.047$, $\hat{\tau}^2 = 0.01$, $I^2 = 62.29\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.50 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be overly influential.

Effects did not differ by risk of bias, publication year, or whether the study was examining post-intervention changes or changes from baseline. We were unable to examine study design, exposure to intervention, evaluation period or whether the model was adjusted for covariates as potential sources of heterogeneity.

Effects of safe spaces on women's participation in their community

Sieverding and Elbadaway's (2016) quasi-experimental study in Egypt was the only study evaluating the impact of safe spaces on women participating more in their community. There was small, and statistically significant, effect ($g = 0.12$, [95% CI: 0.04 to 0.20]), but we assessed the study as having high risk of bias.

Effects of safe spaces on prevention strategies to end violence against women and girls

Sieverding and Elbadaway's (2016) quasi-experimental study in Egypt was also the only study evaluating the impact of safe spaces on effective prevention strategies to end violence against women and girls. There was a small, and statistically significant, effect ($g = 0.12$, [95% CI: 0.04 to 0.20]), and we assessed the study as having high risk of bias.

Effects of safe spaces on whether relief and recovery initiatives in conflict and post conflict situations respond to the needs of women and girls

Mercy Corps' (2015) quasi-experimental study in Niger was the only study evaluating the impact of safe spaces on whether relief and recovery initiatives in conflict and post conflict situations respond to the needs of women and girls, especially vulnerable

groups. There was a small, but not statistically significant, point estimate ($g = -0.11$, [95% CI: -0.27 to 0.05]), and we assessed the study as having high risk of bias.

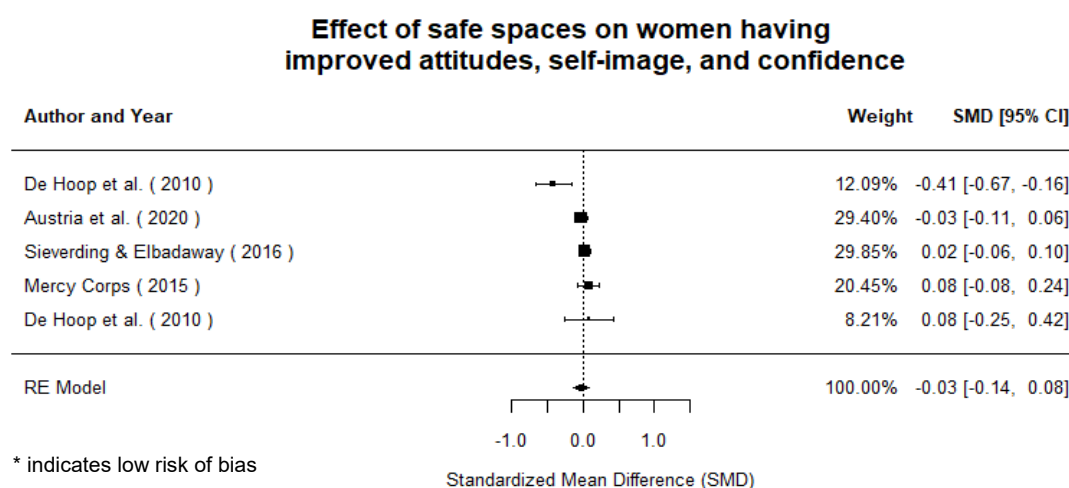
Effects of safe spaces on whether communities have a more positive attitude towards women/marginalised groups

Sieverding and Elbadaway's (2016) quasi-experimental study in Egypt was the only study evaluating the impact of safe spaces on whether communities have a more positive attitude towards women/marginalised groups. There was a very small, but not statistically significant, point estimate ($g = -0.02$, [95% CI: -0.10 to 0.06]), and we assessed the study as having high risk of bias.

Effects of safe spaces on women having improved attitudes, self-image, and confidence

We included a total of $k = 5$ studies in the analysis. We assessed none of the studies as low risk of bias, two as some concerns, and two as high risk of bias. De Hoop and colleagues (2010) contribute two effects from two independent samples, one sample of participants from villages with conservative social norms ($g = 0.41$), and the other from participants in villages with liberal social norms ($g = 0.08$). The observed outcomes ranged from -0.41 to 0.08 . The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.03$ (95% CI: -0.14 to 0.08). Therefore, the average outcome did not differ significantly from zero ($z = -0.52$, $p = 0.61$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 143 SSCB3.

Figure 143: SSCB3: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(4) = 11.52$, $p = 0.02$, $\hat{\tau}^2 = 0.01$, $I^2 = 65.27\%$). An examination of the studentised residuals revealed that one study (De Hoop et al. 2010) had a value larger than ± 2.58 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be overly influential.

Effects were not moderated by risk of bias, publication year, or whether the study was examining post-intervention changes versus changes from baseline. We were unable to test moderation by study design, exposure to the intervention, evaluation period, whether the model was adjusted for covariates,

Effects of safe spaces on attitudes and support for women's economic, social and human rights

Sieverding and Elbadaway's (2016) quasi-experimental study in Egypt was the only study evaluating the impact of safe spaces on improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members. Their report included two effects that fell into this outcome category (e.g. whether the mother was pro-sports for girls). The effects ranged from very small, positive point estimates ($g = 0.02$, 95% CI $[-0.06, 0.10]$) to very small, positive point estimates ($g = 0.02$, 95% CI $[-0.06, 0.10]$). We assessed the study as having high risk of bias.

Effects of safe spaces on safer and more secure household, communities and areas/territories for women, girls, men, and boys

Mercy Corps' (2015) quasi-experimental study in Niger was the only study evaluating the impact of safe spaces on safer and more secure household, communities and areas/territories for women, girls, men, and boys. Their report included six effects that fell into this outcome category (e.g., someone to rely on when you have sensitive problems). The effects examined three outcomes over different time periods and slightly different versions of the intervention (savings plus vs savings only). The effects ranged from very small, negative point estimates ($g = -0.02$, [95% CI: -0.18 to 0.15]) to medium, positive effects ($g = 0.27$, [95% CI: 0.12 to 0.43]). We assessed the study as having high risk of bias.

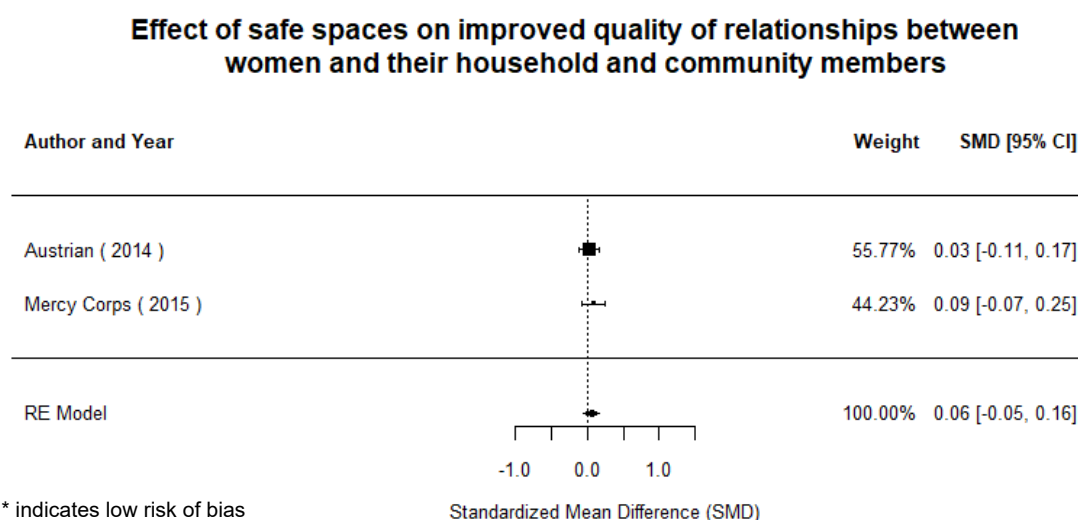
Effects of safe spaces on reduced frequency and distribution of types of violence by an intimate partner

Buehren's (2017) experimental study in Tanzania was the only study evaluating the impact of safe spaces on reduced frequency and distribution of types of violence by an intimate partner. Their report included 2 effects that fell into this outcome category (e.g. ever had sex unwillingly). The study reported effects for slightly different versions of the intervention. The point estimate for the Club Only intervention was very small and negative ($g = -0.06$, [95% CI: -0.13 to 0.01]) and the point estimate for the Club + Microfinance intervention was very small and positive ($g = 0.00$, [95% CI: -0.07 to 0.07]). We assessed the study as having some risk of bias concerns.

Effects of safe spaces on improved quality of relationships between women and their household and community members

Only two studies reported the impact of safe spaces on improved quality of relationships between women and their household and community members, thus we included $k = 2$ studies in the analysis. We assessed both studies as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.06$ (95% CI: -0.05 to 0.16). This positive effect did not differ significantly from zero ($z = 1.02$, $p = .31$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 144 SSCB8. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.28$, $p = 0.59$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid. Neither of the studies was appraised as low risk of bias.

Figure 144: SSCB8. Forest plot showing the observed outcomes and the estimate of the random-effects model



Qualitative findings

We conducted a thematic synthesis on the four linked qualitative studies to the included safe spaces interventions. As indicated above, this thematic synthesis aims to identify themes related to the interplay of intervention design, intervention implementation, target population, and contextual variables with intervention outcomes and effects. In total, we identified 11 descriptive themes, which we configured into three analytical themes (Online Appendix N). These three analytical themes present the synthesis results and are discussed in more detail below.

Theme 1: Safe space intervention design benefits from targeting girls before reaching the contextually eligible age for marriage and parental awareness raising initiatives to encourage uptake.

Cultural and gender practices can predispose girls to early forced marriages and teenage pregnancy. The qualitative evidence suggests to target girls before they reach their teens or ages that are culturally considered eligible for marriage. The Ishraq programme in Egypt targeted girls 13-15 years of age, which was reported as a factor that may have contributed to programme impacts (Ringler 2009) in a community where any girl above 15 years of age was considered for marriage (e.g., through traditional rites ceremonies to prepare teenagers). One community leader from Zambia's (AGEP) claimed: *"They do not value education and especially girls, so girls are married off at a very tender age... we have heard that within the circles, a girl can be married at actually 3 years. 4 years, and the man keeps waiting for that girl, and then when they reach puberty, they just get that girl. So, it is very, very common in this community... Community challenges... religious beliefs, where elderly men marry off young girls... we really need to find some interventions to get to them,"* (Duby, Zulu and Austrian 2016).

Negative perceptions about programme intentions and debates around childrens' rights can impede programme uptake, emphasising the need for programme design to incorporate appropriate awareness and communication strategies. For example, a mentor in the AGEP programme in Zambia discussed religious sentiments preventing participation: *"The first thing is the parents, since inception of the programme it was difficult even the time, we were distributing letters some parents used to refuse, for those*

that understood what we were explaining are the ones that came but some didn't and they never showed up, because they had already told their minds that we are Satanists. There was a lot of influence going on among the members of the community that led to most of the girls dropping out. Due to the rumour of us being Satanists." A parent in the same context in an interview complained about human rights for their children: *"Human rights has to be removed because a parent has to reinforce discipline in the children and sometimes parents are blamed of which it's not their fault but because of this human rights thing."* (Duby, Zulu and Austrian 2016).

Theme 2: Mentorship appears to be a key component of the intervention design and delivery.

Safe spaces interventions are a forum where interactions geared towards changing participants' beliefs and attitudes take place. Mentors are seen to play a crucial role in encouraging collaboration, interaction, learning from each other and bringing about positive attitudes and knowledge, which often cover issues beyond what is learnt through formal education. Safe spaces intervention and relevant group-based activities benefit from the role mentors play in this process. Mentors' personality, passion and purpose for the cause of the girls, as well as their capacity development in the areas of advocacy, outreach, communication and networking skills, support winning the confidence of parents to permit their children to participate.

the case of the Ishraq programme in Egypt, emphasised the role that mentors play in building trusted relations to support programme uptake: *"One of the most effective recruitment strategies was the promoters' home visits. Promoters met with families of eligible girls to describe the programme, what was expected from girls and their families and the valuable opportunity it represented for their daughters. Because of Ishraq's novelty and the community's entrenched social norms, more than one home visit often was needed to secure a girl's participation. By understanding more about who the promoters were and gaining confidence and trust in them, parents began to feel increasingly comfortable with the programme"* (Brady et al. 2007).

Through mentorship, girls may open up on matters they may not be able to share, even with their own parents. One of the participants of the Adolescent Girls Empowerment Programme (AGEP) in Zambia said: *"It has definitely changed my confidence... at the Safe Spaces we are instilled with so much self-esteem, even if someone annoyed you at home when you get to the Safe Spaces you feel better... when you get there your mentor will start to give you different lessons in the process, she will make you laugh about something or put you on the right path"* (Duby, Zulu and Austrian 2016). And about the safe spaces programme in Niger, this was reported, *"...mentors, usually an iya (traditional mentor), are central to the success of safe space objectives. While this finding needs to be further investigated, it also tells us that programs utilising safe spaces must invest in training and coaching their mentors to support girls and effectively deliver content"* (Mercy Corps 2015).

Theme 3: Provision of a suitable, accessible and safe location for meetings and as well as sufficient incentives encourage participation.

While safe spaces provide a forum where girls and women can openly discuss private issues that may feel unsafe to discuss publicly, it is important that meeting places also be physically safe. Meeting places require deliberate selection emphasising safe

accessibility. When meeting places are either too far from the reach of participants or located in dangerous locations, they are not suitable undermining the willingness of participants to attend and focus on programme activities, te. A mentor in the AGEP in Zambia reflected on this issue: *"We had a challenge with where we were meeting, we were not safe. At first, we were just meeting outside so there was no concentration, they (group members) were not participating, they were just looking around and doing other things. When you are in an enclosure, they concentrate..."* (Duby, Zulu and Austrian 2016).

Incentives and rewards are a promising strategy to encourage participation of the targeted population in programmes. However, the design of both needs to take into account a number of reported factors that can undermine their implementation: i) where incentives are not provided as and when due; ii) when the conditions tied to the award of rewards are too stringent for participants to fulfil; and iii) when incentives become the sole reason for participation. These factors can lead to complaints from participants, resulting in an additional burden for mentors or implementers of the programme. One of the mentors on the AGEP in Zambia commented about the challenge that incentives constituted to the programme this way: *"What used to discourage me is the issue to do with the prizes, they used to give me some trouble and they caused a lot of children to stop coming because they used to say a child can only get a prize if they have been attending the meetings continuously for ten weeks but you would find that some of them would miss some meetings and then when they come they find that their friends are being given prizes you find that they go back and never come back and secondly instead of giving the children the prizes at the time that they promised they would give them, they would delay, it used to take a long time and they would even start asking to say madam, when will we get the prizes, madam, we have been attending the meetings for the past so many weeks, when will we get our things? That used to discourage us, it used to be a challenge for us mentors... Most of them stopped because some of the children were coming for the meetings to get the prizes not necessarily to learn,"* (Duby, Zulu and Austrian 2016).

Discussion

Quantitative meta-analysis revealed that safe spaces interventions had positive significant effects on women's increased access to and ownership of assets, credit, and income. Although small positive effects were found in terms of increases capacity to understand financial services, increases mobility, control over body, marriage and sexual health, better life skills, participation in household/community level decision making and quality of relationship with household or community members, none of these were statistically significant. The intervention group was only covered by [4] studies and the associated GRADE analysis showed low confidence about the quality of the evidence. A complementary qualitative analysis was carried out to put forward a set of recommendations that may help design more effective safe spaces interventions

An explanation to the limited effects might be the fact that many of the outcomes involved changes in long-standing social and cultural norms and attitudes towards women or girls, and it may need a longer period to observe visible changes. In view of this, awareness building, at both household and community levels, were found to be a key design element. Working only with target participants may limit the potential for such changes if the household / community members are not ready to change their attitudes.

This supports the idea of pre-intervention research to design the best intervention mechanism to suit the settings.

Targeting the appropriate demographic groups was found to be a further crucial factor. With regard to bringing about positive changes in outcomes related to marriage and maternal health, control over body and sexual health, it is important to work with unmarried young girls and women. Awareness building among participants as well as their household members and the wider community at an early stage of the intervention plays a critical role as many of such outcomes were compounded by social and cultural norms and traditions.

As mentioned earlier, quantitative analysis showed small positive yet insignificant effects of safe spaces interventions on women's/girls' control over their bodies, knowledge about sexual health, better life skills and freedom of movement and association. These outcomes are often conveyed through collaborative learning and participation in group activities and mentorship provided via safe spaces interventions was found as an important driver in view of this. Mentors play a crucial role in encouraging collaboration, interaction, learning from each other and bringing about positive attitudes and knowledge. Mentors served as trustworthy source to whom participants could turn when in need. Parents would often feel comfortable to let their daughters participate in outdoor activities if accompanied by mentors.

Positive changes in women's and girls' lives expected from safe spaces interventions critically hinges on participation in activities such as meetings, trainings, and group learning/ activities. The qualitative synthesis revealed several concerns hindering women's/girls' participation which might have resulted in the statistically insignificant effects found in the quantitative analysis. Qualitative evidence found that distance to the safe spaces meeting location, timing of the meetings, availability of the meeting venues during the scheduled meeting times were among key challenges limiting participation. The physical facility where safe spaces meetings and activities would be carried out needs to be safe and suitable for participants. These operational issues need to be considered during design.

Summary of findings and discussion

We included six studies in six countries in MENA and Sub-Saharan Africa that evaluated the effect of safe spaces. While many outcome categories had sufficient evidence for quantitative synthesis, only one outcome showed a significant impact of safe spaces. Safe spaces interventions had positive significant effects on women's increased access to and ownership of assets, credit, and income. Though small positive effects were found for increased capacity to understand financial services, increased control over body and sexual health, increased freedom of movement and association, better life skills, reduced support for or instance of child and forced marriage, increased participation in household/community level decision making and quality of relationship with household or community members, none of these were statistically significant. There was also a negative but non-significant impact on improved attitudes/self-image/confidence. While the general body of evidence did not allow for testing of moderation in most instances, we did find that exposure to intervention was a source of heterogeneity for women's access to and ownership of assets, credit and income such that longer exposure periods were related to smaller impacts. This is likely due to the one study with the longest

exposure period having the smallest effect. Overall, the GRADE assessments indicate a very low to low certainty in this body of evidence.

The quantitative findings are buttressed by three qualitative themes from the four qualitative linked studies. Overall, targeting is crucial for such interventions to improve the uptake and potential for long-lasting effects. By focusing specific safe spaces interventions for adolescent girls, forced marriage could potentially be reduced by the package of activities offered in such a space. Physical space as well as planned activities plays a crucial role in maintaining uptake and the perceived safety of participants and implementing partners. Lastly, awareness raising with parents (if adolescent girls are participating) or household members can prove beneficial to participation and uptake of the interventions offered in safe spaces. The summary of our quantitative findings along with the GRADE certainty of evidence ratings are presented in Table 54.

Table 54: GRADE summary of findings and certainty of evidence on safe spaces

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(AB1) Increased capacity of women to understand and use financial, banking, and business services effectively										
4	RCT – 2 QED – 2	very serious ^a	not serious	not serious	not serious	none	8679	SMD 0.05 SD higher (0.03 lower to 0.13 higher)	⊕⊕○○ LOW	Important, but not critical
(AB2) Women have increased access and ownership to assets, credit and income.										
4	RCT – 2 QED – 2	very serious ^b	serious ^c	not serious	not serious	publication bias strongly suspected	8742	SMD 0.14 SD higher (0.05 higher to 0.23 higher)	⊕○○○ VERY LOW	Important, but not critical
(AC3) Improved capacity of women entrepreneurs										
1	RCT	serious ^d	serious ^e	not serious	not serious	none	3179	Two positive effect estimates with a 95% CI of -0.05 to 0.09.	⊕⊕○○ LOW	Important, but not critical
(BA2) Women have more and better control over their bodies and sexual health.										
2	RCT – 1 QED – 1	very serious ^f	serious ^g	not serious	serious ^h	none	4226	SMD 0.1 SD higher (0.05 lower to 0.25 higher)	⊕○○○ VERY LOW	Critical
(BA3) Women have increased freedom of movement and association.										
2	QED – 2	very serious ⁱ	very serious ^j	not serious	serious ^k	none	2635	SMD 0.02 SD higher (0.07 lower to 0.1 higher)	⊕○○○ VERY LOW	Limited Importance
(BA5) Women have more positive attitude towards taking action to claim their rights.										
1	QED	very serious ^l	serious ^e	not serious	not serious	none	615	Two positive effect estimates with a 95% CI of -0.10 to 0.28.	⊕○○○ VERY LOW	Limited Importance
(BA6) Reduced percentage of women agreeing with certain reasons that justify violence against women and girls										
1	RCT	very serious ^m	serious ^e	not serious	not serious	none	732	Two positive and one negative effect estimate with a 95% CI range of -0.11 to 0.13.	⊕○○○ VERY LOW	Critical
(BA7) Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them.										
3	RCT – 1 QED – 2	very serious ⁿ	very serious ^o	not serious	serious ^p	none	6752	SMD 0.1 SD higher (0.05 lower to 0.25 higher)	⊕○○○ VERY LOW	Important, but not critical
(BA8) Reduced instances of child or forced marriage										
3	RCT – 1 QED – 2	very serious ^q	very serious ^r	not serious	serious ^s	publication bias strongly suspected strong association	6823	SMD 0.39 SD higher (0.08 lower to 0.86 higher)	⊕○○○ VERY LOW	Critical
(BB1) Increased participation in decision making by Women at the household or community level, including during crisis response										
4	RCT – 1 QED – 3	very serious ^t	very serious ^u	not serious	serious ^v	publication bias strongly suspected	6842	SMD 0.01 SD higher (0.1 lower to 0.13 higher)	⊕○○○ VERY LOW	Limited Importance
(BB2) Women participate more in their community.										
1	QED	very serious ^w	serious ^e	not serious	not serious	none	2248	SMD 0.12 SD higher (0.04 higher to 0.2 higher)	⊕○○○ VERY LOW	Limited Importance

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(CA6) Effective prevention strategies supported to end violence against women and girls										
1	QED	very serious ^x	serious ^e	not serious	not serious	none	2248	SMD 0.12 SD higher (0.04 higher to 0.2 higher)	⊕○○○ VERY LOW	Critical
(CA8) Increased community support for women's and children's human, economic and legal rights										
1	QED	very serious ^y	serious ^e	not serious	serious ^z	none	594	SMD 0.11 SD lower (0.27 lower to 0.05 higher)	⊕○○○ VERY LOW	Important, but not critical
(CB2) Communities have a more positive attitude towards women/marginalised groups.										
1	QED	very serious ^{aa}	serious ^e	not serious	not serious	none	2248	SMD 0.02 SD lower (0.1 lower to 0.06 higher)	⊕○○○ VERY LOW	Limited Importance
(CB3) Women have improved attitudes, self-image and confidence.										
4	RCT – 1 QED – 3	very serious ^{ab}	very serious ^{ac}	not serious	serious ^{ad}	none	7567	SMD 0.03 SD lower (0.14 lower to 0.08 higher)	⊕○○○ VERY LOW	Limited Importance
(CB4) Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members										
1	QED	very serious ^{ae}	serious ^e	not serious	not serious	none	2248	Two positive studies with identical Cohen's d: 0.02 (-0.06, 0.10).	⊕○○○ VERY LOW	Limited Importance
(CB6) Safer and more secure household, communities and areas/territories for women, girls, men, and boys (54258398)										
1	QED	very serious ^{af}	serious ^e	not serious	serious ^{ag}	none	615	One negative and one positive effect estimate with a 95% CI range of -0.18 to 0.21	⊕○○○ VERY LOW	Critical
(CB6) Safer and more secure household, communities and areas/territories for women, girls, men, and boys (54977131)										
1	QED	very serious ^{ah}	serious ^e	not serious	not serious	none	762	Four positive effect estimates with a 95% CI range of -0.07 to 0.43	⊕○○○ VERY LOW	Critical
(CB7) Reduced frequency and distribution of types of violence by an intimate partner										
1	RCT	serious ^{ai}	serious ^e	not serious	not serious	none	3179	One positive and one negative effect estimate with a 95% CI range of -0.05 to 0.00	⊕⊕○○ LOW	Critical
(CB8) Improved quality of relationships between women and their household and community members										
2	QED – 2	very serious ^{aj}	not serious	not serious	not serious	none	1993	SMD 0.06 SD higher (0.05 lower to 0.16 higher)	⊕⊕○○ LOW	Limited Importance

Note. CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

- Downgraded twice because all studies present risk of bias and two are of high risk of bias, including selection bias and deviation from intended intervention.
- Downgraded to very serious because none of the studies are rated as low bias. When high ROB studies are removed, summary effect changes from .14 and significant to .09 and non-significant.
- Downgraded once because point estimates vary considerably and there are CIs that do not overlap.
- Downgraded because of uncertainty related to outcome reporting bias.
- All single studies downgraded once for inconsistency.

- f. Downgraded two times because: 1. both studies present at least some concerns related to risk of bias and one is of high risk, and 2. removing either study would have an impact on the meta-analysis.
- g. Downgraded once: 1. Variation of points estimates: data on both side of the threshold, 2. overlaps of CI: systematic overlap, 3. $I^2 = 0\%$, 4. test of stat: According to the Q-test, there was no significant amount of heterogeneity in the true outcomes.
- h. Downgrade once as p value of 0.66 and Sample size > 400.
- i. Downgraded twice because: 1. both studies present at least some concerns related to risk of bias and one is of high risk, and 2. removing either study would have an impact on the meta-analysis
- j. Downgraded twice: 1. Variation of points estimates: data on both side of the threshold, 2. Overlaps of CI: no systematic overlap, 3. $I^2 = 87.07\%$, 4. test of stat: According to the Q-test, there was significant amount of heterogeneity in the true outcomes
- k. downgrade once as p value of 0.29 and Sample size > 400.
- l. Downgraded because of uncertainty across several criteria and a high risk of spill-over/contamination.
- m. Downgraded because of uncertainty across several criteria and a high risk of spill-over/contamination and reporting bias.
- n. Downgraded twice because 1. all studies in this group present risk of bias, and one study is of high risk, and 2) The high risk of bias studies make up >30% of the weight of the meta-analysis.
- o. Downgraded twice: 1. Variation of points estimates: data on both side of the threshold, 2. Overlaps of CI: no systematic overlap, 3. $I^2 = 79.74\%$, 4. test of stat: According to the Q-test, the true outcomes appear to be heterogeneous.
- p. downgrade once as p value of 0.19 and Sample size > 400.
- q. Downgraded twice because: 1. One raises some concerns related to risk of bias and 2 high risk including high risk on 3 domains 3. removing either study would have an impact on the meta-analysis.
- r. Downgraded twice: 1. Variation of points estimates: data on both side of the threshold, 2. Overlaps of CI: no systematic overlap with one study CI very far from the two others, 3. $I^2 = 98.27\%$, 4. test of stat: According to the Q-test, the true outcomes appear to be heterogeneous.
- s. downgrade once as p value of 0.10 and Sample size > 400.
- t. Downgraded twice because half of the body of evidence come from high risk of bias studies, and the rest from studies of medium or uncertain bias. Additionally, three of four of the studies are QEDs, including two high risk.
- u. Downgraded twice 1. Variation of points estimates: data on both side of the threshold, 2. Overlaps of CI: no systematic overlap with one study CI very far from the two others, 3. $I^2 = 62.29\%$, 4. test of stat: According to the Q-test, the true outcomes appear to be heterogeneous.
- v. downgrade once as p value of 0.81 and Sample size > 400.
- w. Downgraded because of uncertainty across several criteria and a high risk of spill-over/contamination and reporting bias.
- x. Downgraded because of uncertainty across several criteria and a high risk of spill-over/contamination and reporting bias.
- y. Downgraded because of uncertainty across several criteria and a high risk of spill-over/contamination.
- z. Downgraded once because of wide CIs crossing the threshold of interest.
- aa. Downgraded because of uncertainty across several criteria and a high risk of spill-over/contamination and reporting bias.
- ab. Downgraded twice because half the body of evidence comes from high risk of bias QEDs. Furthermore, removing either high RoB studies impacts the results of the meta-analysis
- ac. Downgraded twice 1. Variation of points estimates: data on both side of the threshold, 2. Overlaps of CI: no systematic overlap with one study CI very far from the two others, 3. $I^2 = 65.27\%$, 4. test of stat: According to the Q-test, the true outcomes appear to be heterogeneous.

- ad. downgrade once as p value of 0.61 and Sample size > 400.
- ae. Downgraded because of uncertainty across several criteria and a high risk of spill-over/contamination and reporting bias.
- af. Downgraded because of uncertainty across several criteria and a high risk of spill-over/contamination.
- ag. Downgraded because the point estimates are quite broad and the range of CI is widely across both sides of the threshold.
- ah. Downgraded due to uncertainty across several criteria including performance bias and outcome measurement bias.
- ai. Downgraded because of uncertainty related to outcome reporting bias.
- aj. Downgraded to very serious because both included studies are of high risk of bias.

4.6.3 Life, social and livelihood skills and capacity building

Life, social and livelihood skills and capacity building include the process of changing attitudes and behaviours by imparting knowledge and developing skills while maximising the benefits of participation, knowledge exchange and ownership (Di Pierro n.d.).

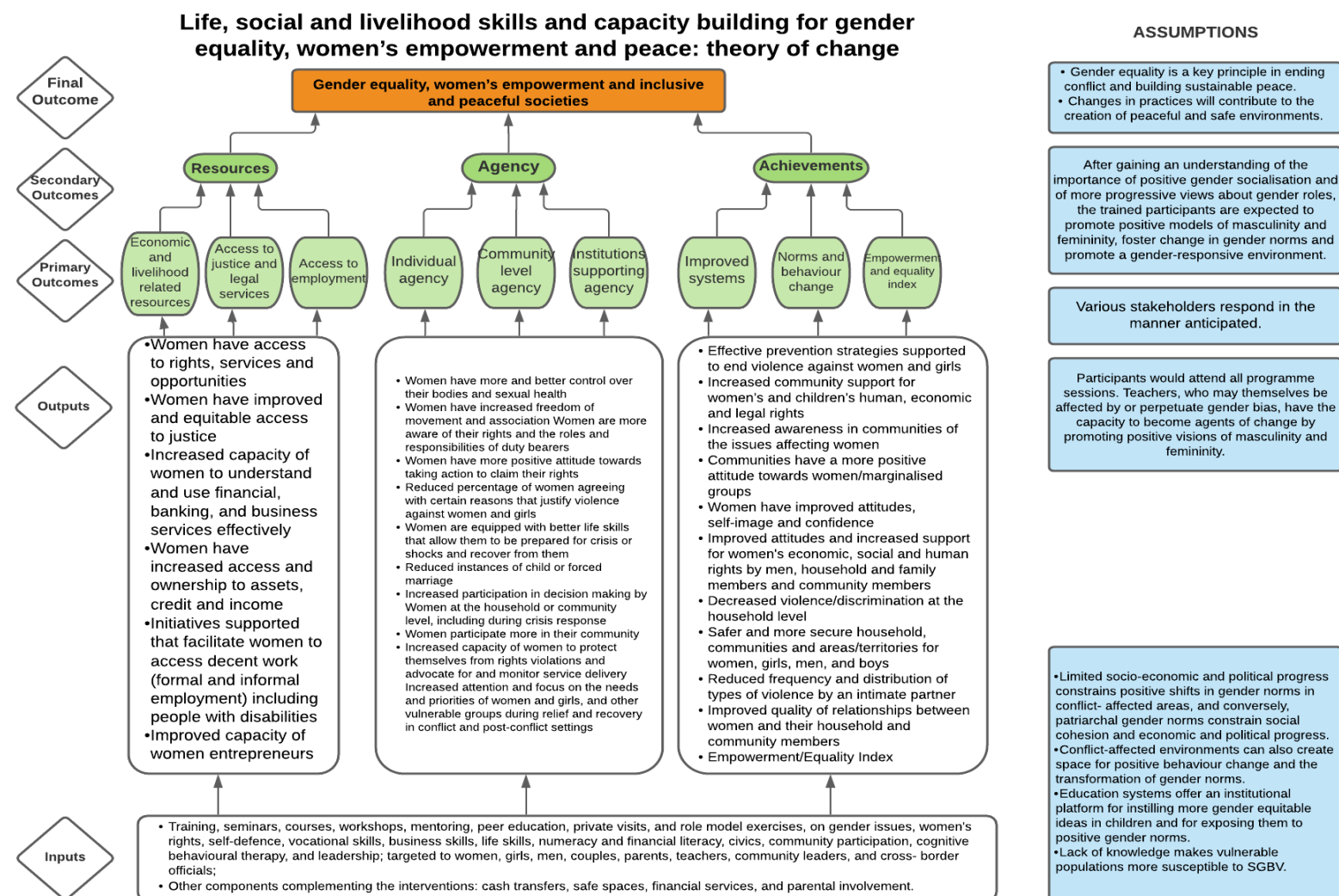
Training and capacity development encompass vocation and technical skills, cognitive and non-cognitive skills, life skills etc.

How do life, social and livelihood skills and capacity building affect gender equality, women's empowerment and peace outcomes?

Figure 145 below maps the causal chain of how life, social and livelihood skills and capacity building may improve gender equality, women's empowerment and peace outcomes.

Capacity building interventions and training in life, social and livelihood skills programmes aim to empower women through increasing their access to income generating activities through decent work and self-employment, assets, and services; the control over their bodies and sexual health; and their freedom of movement and association. Equipping women with better life skills should be translated into more participation in decision making in their households and in their communities, reduced violence against women, and improved attitudes, self-image and confidence to take action and claim their rights. In order to achieve this, both formal and informal platforms, seminars, courses, workshops, mentoring, peer education and other forms of training can be complemented with cash transfers, safe spaces, financial services and parental involvement. Assumptions supporting this causal path include the idea that lack of knowledge is a barrier for women to have more resources, agency and achievements, where training and education interventions may be effective in overcoming this barrier. Further, the educational system is assumed to be a platform where not only more equitable gender norms can be installed, but also where teachers can promote and set an example of gender equality. Equality is supposed to promote the end of conflict and build sustainable peace and safe environments.

Figure 145: Life, social and livelihood skills and capacity building for gender equality, women's empowerment and peace: theory of change



Description of included studies

We included 20 studies reported in 26 different papers that evaluated the effect of 12 identified programmes. We included more than one paper that evaluated the same programme if the author(s) undertook different analyses or reported different outcomes over several papers.

Population

Most studies targeted adolescent girls aged between 10 and 19 years old; most were interventions implemented alongside safe spaces and including parental involvement. Interventions working with adults, primarily aimed to promote egalitarian decision making within households, targeted married or cohabiting couples. Other populations targeted were teachers, small-scale coffee farmers, bank clients, micro entrepreneurs, and community members.

The included studies evaluated 12 different programmes and trials covered countries in South Asia and Sub-Saharan Africa and 13 countries including: Afghanistan, Bangladesh, India, DRC, Ethiopia, Liberia, Malawi, Mali, Sierra Leone, South Sudan, Tanzania, Uganda, Zambia.

Intervention, inputs and activities

The included studies evaluated a range of different activities and inputs including:

- Life skills sessions, mentoring, parental involvement and safe spaces (n = 8): Life skills sessions covered content ranging from decision-making and disagreement resolution, to reproductive health, puberty, gender norms, cultural pressures, healthy relationships, life goals and violence prevention, while parent/caregiver discussion groups focussed on positive relationship building, empathetic communication, non-violent discipline methods and specific developmental and cultural issues experienced by adolescent girls.
- Skills building courses (n = 7): Using peer education, participatory courses, or afterschool programmes, to train girls on topics ranging from informal negotiation, budgeting, financial literacy, business planning, and savings, to social skills, cognitive behavioural therapy skills, self-defence, leadership and advocacy skills.
- Vocational training plus resource provision (n = 1): Skill building in numeracy and chosen vocational skill (e.g. sewing) together with foundational training in modules that included: the value of women's work, ways to save money, ways to earn income and improve income-generating activities, basic business skills, ways to improve health and well-being, women's rights and prevention of SGBV, strategies to make decisions and negotiate, civic action and advocacy, social networks and safety nets. Resources provision in the form of monthly cash stipend, mechanisms to save money, and referrals to health, legal and financial services.
- Men's and couples' seminars and courses (n = 5): Delivered either at a local established ceremony for community discussion, or through workshops, covering topics such as division of roles, responsibilities, decision-making power and access to resources in their households, attitudes towards and treatment of women. With or without complementary activities such as gender dialogues with community leaders and women's self-help groups.
- Training teachers as agents of change for promoting gender norms and equality (n = 1): three training sessions plus reinforcing text messages on gender, conflict and peacebuilding.

- Training for cross-border officials on sensitising against SGBV (n = 1): professionalisation of officials to reduce harassment of traders and improve conditions for small-scale cross-border trade; empowerment of small-scale traders; and support for policy dialogue and improved coordination among local actors, with a set of joint workshops held between traders and officials.
- Civics course on community participation in governance (n = 1): A civics course providing information about the rights and responsibility of citizens in the democratic process.

Table 55: Life skills and capacity building activities design features of included studies

Study	Activity/Input	Length of treatment	Intervention frequency
Tanner - 2017	Life skills sessions, mentoring, parental involvement and safe spaces	9	Weekly
Tanner - 2017	Life skills sessions, mentoring, parental involvement and safe spaces	9	Weekly
Tanner - 2017	Life skills sessions, mentoring, parental involvement and safe spaces	9	Weekly
Buehren - 2017	Life skills and livelihoods training combined with small capital to help turn training into an income generating activity and courses on financial literacy and budgeting skills	24	Unclear
Stark - 2018	Life skills sessions, mentoring, parental involvement and safe spaces	16	Weekly
Scales - 2013	Peer education in social skills, literacy, and school learning, combined with study circles to support in-school girls, education on budgeting, financial and business planning, and savings, and an incentive to delay marriage	48	Daily
Leight - 2021	Participatory skill building courses that covered a very broad range of topics and are facilitated amongst women, men, and couples and coupled with a local ceremony	7	Daily
Lecoutere - 2017	Couple seminars and coaching through workshops and private visits, in participatory intra-household decision making, plus women's leadership training	17	Monthly
Field -	Financial literacy skills for SHG members to	8	Daily

Study	Activity/Input	Length of treatment	Intervention frequency
2010	improve access to financial services including a training module		
Sharma - 2020	Participatory skill building courses that cover a very broad range of topics and are facilitated amongst women, men, and couples and coupled with a local ceremony	7	Daily
Murray - 2020	Teaching cognitive behavioural therapy (CBT) skills (engagement, introduction/psychoeducation, safety, substance use reduction, cognitive coping and restructuring, problem solving, behavioural activation, relaxation, and exposure)	2	Weekly
Stark - 2018	Life skills sessions, mentoring, parental involvement and safe spaces	10	Weekly
Chinen - 2016	Training teachers as agents of change for promoting gender norms and equality	9	Unclear
Özler - 2020	Girl Empower life skills curriculum, caregiver discussion groups, individual savings start-up for the girls, and capacity building for local health and psychosocial service providers, combined with participation incentive payment	10	Weekly
Noble - 2019	Foundational training, skill building in numeracy and a chosen vocational skill, cash stipend, formal and informal mechanisms to save money, and referrals to health, legal, and financial services, combined with safe spaces	Unclear	Weekly
Lecoutere - 2018	Couple seminars and coaching through workshops and private visits, in participatory intra household decision making, plus women's leadership training	17	Monthly
Ashraf - 2018	Informal negotiation skills taught to girls along with role model exercises	5	Daily
Stark - 2018	Life skills sessions, mentoring, parental involvement and safe spaces	10	Weekly
Decker - 2018	Empowerment self-defence training in early recognition of boundary testing, negotiation,	1.5	Weekly

Study	Activity/Input	Length of treatment	Intervention frequency
	diffusion and distraction tactics, verbal assertiveness, and physical self-defence		
Bastian - 2018	Training on a financial product and business skills intervention	22	Weekly
Oxfam - 2014	Training on women's rights issues, leadership skills and advocacy skills, combined with awareness-raising and advocacy on women's property and literacy rights within their communities	60	Weekly
Halim - 2019	Savings groups for women plus peer-groups for men addressing gender relations and IPV prevention, with and without gender dialogues for community leaders.	8	Daily
Croke - 2020	Training for cross-border officials on sensitising against SGBV	16	Daily
Gottlieb - 2016	Civics course on community participation in governance	2	Weekly
Kumar - 2018	Women's groups for agricultural training, provision of initial inputs and nutrition BCC	48 months	Unclear

Comparison

All included studies compared treated groups to comparison groups receiving no intervention. 12 studies included multiple treatment arms.

Outcomes

The included studies reported on a number of relevant outcomes, including the following:

- Increased capacity of women to understand and use financial, banking, and business services effectively (n = 3): Women both benefit from enhanced skills and capacities in the use of financial, banking and business services. This can include training programmes, establishment of bank accounts, entrepreneurship training, business skills seminars, etc. any activity aimed to increase women's independence in managing these assets/services.
- Women have increased access and ownership to assets, credit and income (n = 8): Women are able to apply for, receive and manage assets/credit and income and have support to manage, claim and execute their assets without pressure or influence from external actors, including male family members, husbands and cultural leaders.
- Women have more and better control over their bodies and sexual health (n = 7): Women are able to make free and educated decisions about their health (not included in this study), sexuality and procreation.

- Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them (n = 4): Women develop skills, have access to resources, and make use of these skills and resources to increase their level of resilience and be equipped to face life challenges (health, education, finance, social relations, work etc.) These resources can take the form of Safe spaces, community dialogues and informal peer/community educators and increased ability of women to address life challenges (access to education, literacy, jobs, services, etc.). This would also cover activities supporting the development of self-confidence in the capacity of women to face life challenges. During conflict, Women are equipped with increased skills and knowledge for resilience to recover and adapt to shocks, including learning how to have savings, adapt employment skills, without resorting to negative coping mechanisms and the related resources.
- Women have improved attitudes, self-image and confidence (n = 6): Women feel more entitled to claim their rights and needs in community and change social norms and behaviours. They are aware of the importance of their status in society and are empowered to take this role and make use of the available resources to guarantee their rights.

The division of the immediate and secondary outcomes is reported in Table 56 below:

Table 56: Summary of secondary and immediate outcomes for LSCB and capacity building

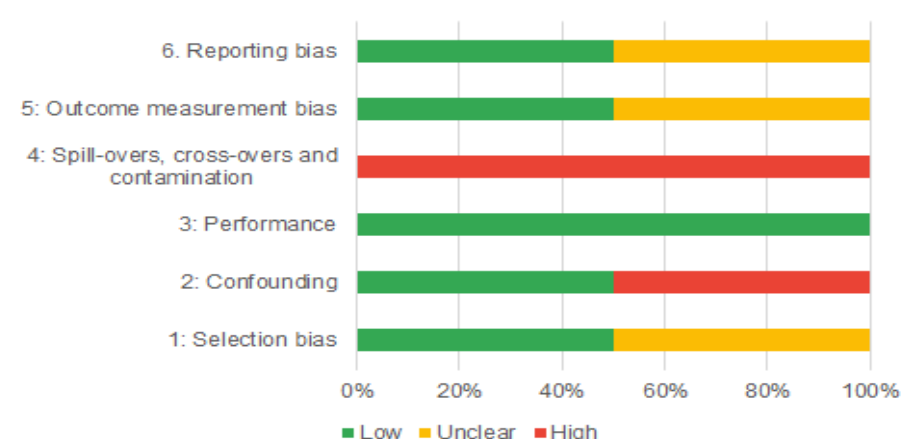
Secondary outcome category	Immediate outcome	Number of studies
Resources material, human and social resources which serve to enhance the ability to exercise choice	Access to justice and legal services	1
	Economic and livelihood related resources	9
	Access to employment	1
Agency ability to define one's goals and act upon them and operationalised decision-making;	Individual agency	15
	Community level agency	10
	Institutions supporting agency	0
Achievement ways of being and doing which can be realised by different individuals	Improved systems	1
	Norms and behaviour change	16
	Empowerment index	4

Study design

Two of our life, social and livelihood skills and capacity building studies used a QED. This included Lecoutere (2017) and Oxfam (2014). We assessed both studies as having a high risk of bias. As detailed in Figure 146 below, high risk of bias was identified in two RoB domains (confounding and spill overs). Confounding was assessed as a risk of bias

in one of the studies that did not account for cluster controls and could not reject unobserved bias (Oxfam 2014). Spill-overs were an issue when participants in comparison areas reported having participated in similar activities than those delivered by the program (Lecoutere 2017; Oxfam 2014). For other domains, we did not have enough information to justify the absence of bias in at least one study, this was the case for selection bias, outcome measurement and reporting bias. The only domain we did not identify any issue was performance, where there was no obvious issue with monitoring in any of the studies.

Figure 146: Life, social and livelihood skills and capacity building studies QED RoB assessment



Twenty-four of our life, social and livelihood skills and capacity building studies use an experimental design, including (among others) Ashraf (2018), Bastian (2018), Buehren (2017), Croke (2020), Chinen (2016), Decker (2018) and Field (2020).

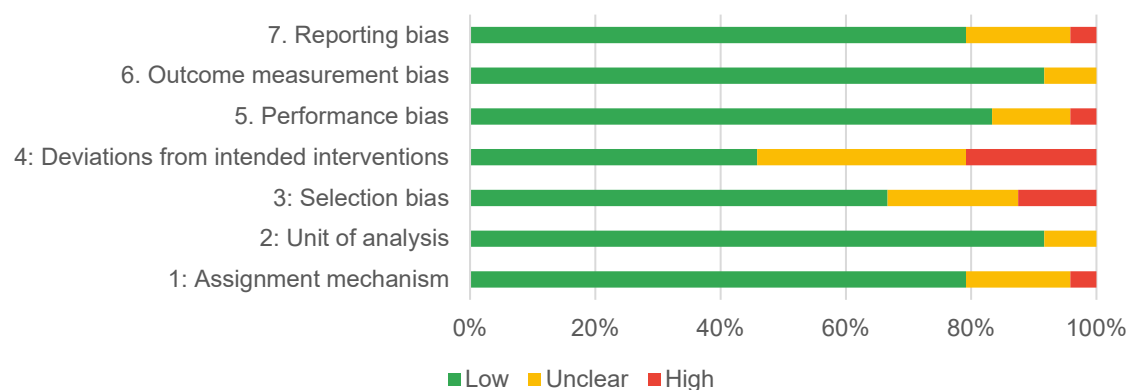
We assessed about one fourth of the studies as having a low risk of bias, however, 30 per cent had a high risk of bias and the rest had some concerns. As detailed in Figure 147 below, high risk of bias was identified in five RoB domains (assignment mechanism, selection, deviation, performance and reporting). The most common issue in the RCT studies in the capacity building interventions was deviations from the intended interventions. For example, two studies suffered from survey effects because of including information in the data collection tools that could have given the control participants an idea of what the intervention group received (Scales 2013; Murray 2020), while other studies reported risks of non-compliance of the assignment to treatment and control groups (Croke 2020; Chinen 2016; Buehren 2017).

We identified a high risk of selection bias in studies that suffered from differential attrition (Halim 2019; Murray 2020) or very high attrition even if not correlated to the intervention (Buehren 2017). We also assessed one study as having a high risk of bias on the assignment mechanism because the authors do not report balance at baseline between treatment and control groups, so that not enough information was available to confirm that randomization had worked (Gottlieb 2016).

Other limitations were related to performance and reporting biases. We assessed one study as having a high risk of motivation bias among participants, introduced by the process of monitoring, and a high risk of reporting bias because the authors do not

provide separate analysis for all treatment arms (Scales 2013). The same study does not consider clustering in the analysis of results and do not provide enough information to rule out that attrition is related to treatment, or to confirm that balance is achieved at baseline.

Figure 147: Life, social and livelihood skills and capacity building studies RCT RoB assessment



Qualitative studies, process evaluations and project documents

We identified 16 additional documents related to six (06) programmes covered by the life, social and livelihood skills and capacity building included studies.

- Creating Opportunities through Mentoring, Parental Involvement and Safe Spaces (COMPASS) programme (DRC): one qualitative and two descriptive quantitative documents
- Kishoree Kontha ("Adolescent Girls' Voices") (Bangladesh): one process evaluation
- Hanns R. Neumann Stiftung "Gender Household Approach" (Uganda): one descriptive quantitative document
- Women for Women International's (WfWI) What Works to Prevent Violence Programme (Afghanistan): four descriptive quantitative documents and one qualitative document
- Unite for a Better Life (UBL) (Ethiopia): one process evaluation, two descriptive quantitative document
- Together to End Violence Against Women (TEVAW) (Tanzania): one descriptive quantitative document
- Realigning Agriculture for Improved Nutrition (RAIN) (Zambia): two descriptive quantitative documents

The majority of these studies were marked as moderate ($n = 6$) or low ($n = 6$) empirical quality. The remainder ($n = 4$) were marked as high quality.

Synthesis of findings

The following subsection presents the results of effectiveness of life, social and livelihood skills and capacity building on gender equality, women's empowerment and peace outcomes.

Quantitative findings

Effects of capacity building on women's access to rights, services and opportunities

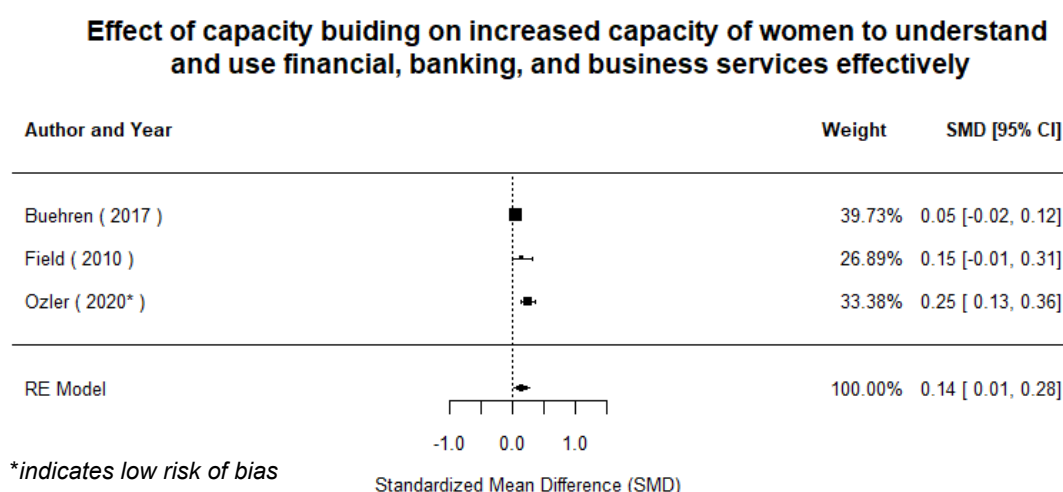
Stark and colleagues' (2018) experimental study in Ethiopia was the only study

evaluating the impact of life, social and livelihood skills and capacity building on women's access to rights, services and opportunities. There was a very small, but not statistically significant, positive point estimate ($g = 0.02$, [95% CI: -0.12 to 0.16]), and we assessed the study as having some risk of bias concerns.

Effects of capacity building on increased capacity of women to understand and use financial, banking and business services effectively

We included a total of $k = 3$ studies in the analysis. We assessed one study as low risk of bias, one as some concerns, and one as high risk of bias. The observed outcomes ranged from 0.05 to 0.25. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.14$ (95% CI: 0.01 to 0.28). Therefore, the average outcome differed significantly from zero ($z = 2.14$, $p = 0.03$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 148 CBAB1.

Figure 148: CBAB1: Forest plot showing the observed outcomes and the estimate of the random-effects model

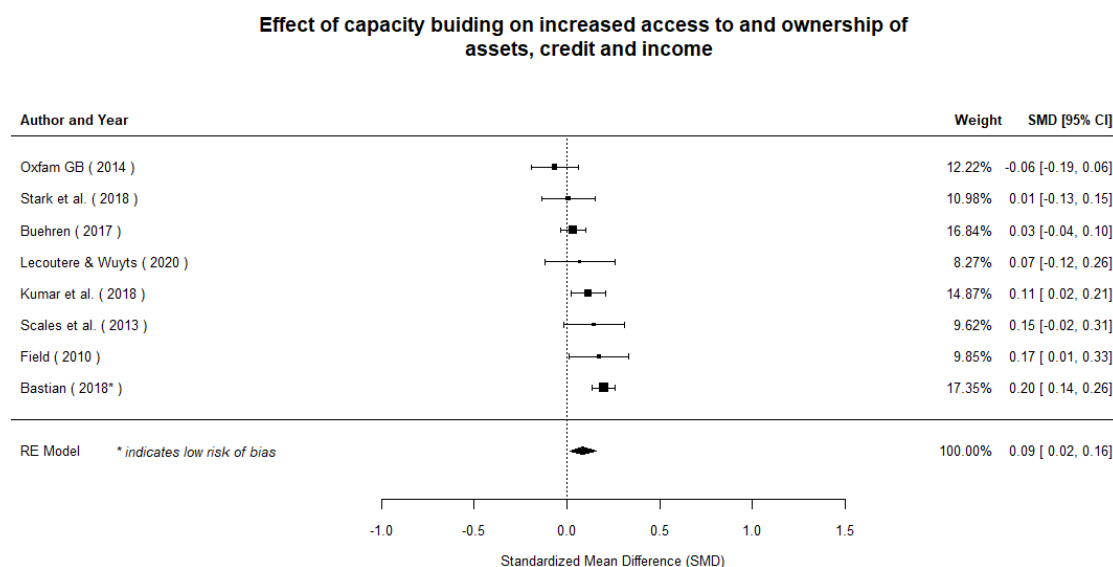


According to the Q -test, the true outcomes appear to be heterogeneous ($Q(2) = 8.53$, $p = 0.01$, $\hat{\tau}^2 = 0.01$, $I^2 = 76.54\%$). An examination of the studentised residuals revealed that one study (Buehren 2017) had a value larger than ± 2.39 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate.

Effects of capacity building on increased access to and ownership of assets, credit and income

We included a total of $k = 8$ studies in the analysis. The observed outcomes ranged from -0.06 to 0.20 . We assessed one study as low risk of bias, four as some concerns, and four as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.09$ (95% CI: 0.02 to 0.16). Therefore, the average outcome differed significantly from zero ($z = 2.42$, $p = 0.02$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 149 CBAB2.

Figure 149: CBAB2: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(7) = 22.55$, $p < 0.01$, $\hat{\tau}^2 = 0.01$, $I^2 = 68.96\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than ± 2.73 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

We again tested all moderators. High risk of bias studies were not significantly different from studies assessed as low or some concerns ($\hat{B} = -0.10$, $p = 0.07$ [95% CI: -0.21 to 0.01]). Length of exposure to the intervention ($\hat{B} = -0.003$, $p = 0.08$ [95% CI: -0.01 to 0.001]) and evaluation period ($\hat{B} = 0.001$, $p = 0.91$ [95% CI: -0.01 to 0.01]) we also not significant sources of variation. Effects from experimental studies were not different than effects from quasi-experimental studies ($\hat{B} = -0.13$, $p = 0.11$ [95% CI: -0.28 to 0.03]). Effects also did not differ by publication year ($\hat{B} = -0.002$, $p = 0.88$ [95% CI: -0.03 to 0.02]), by gender inequality score ($\hat{B} = -0.37$, $p = 0.68$ [95% CI: -2.14 to 1.40]), nor by fragile state index score ($\hat{B} = -0.01$, $p = 0.17$ [95% CI: -0.01 to 0.003]). Finally, effects were no different between studies that target women only versus studies targeting mixed genders ($\hat{B} = -0.06$, $p = 0.58$ [95% CI: -0.25 to 0.14]). There were not sufficient numbers in each group to test by country and we could not test for adjustment by covariates as only one model was unadjusted.

Effects of capacity building on initiatives that facilitate women's access to decent work
Buehren's (2017) experimental study in South Sudan was the only study evaluating the impact of life, social and livelihood skills and capacity building on initiatives that facilitate women's, including people with disabilities, access to decent work (formal and informal employment). Their report included five effects that fell into this outcome category (e.g. self-employment and farmer's self-employment). The effects ranged from very small, negative point estimates ($g = -0.01$, [95% CI: -0.08 to 0.06]) for hours worked per month to very small, positive effects ($g = 0.09$, 95% CI [0.02, 0.16]) for non-farm self-employment. We assessed the study as having high risk of bias.

Effects of capacity building on improved capacity of women entrepreneurs

Buehren's (2017) experimental study in South Sudan was also the only study evaluating the impact of life, social and livelihood skills and capacity building on improved capacity of women entrepreneurs. There was a very small, but not statistically significant, negative point estimate ($g = -0.02$, [95% CI: -0.09 to 0.05]), and we assessed the study as having high risk of bias.

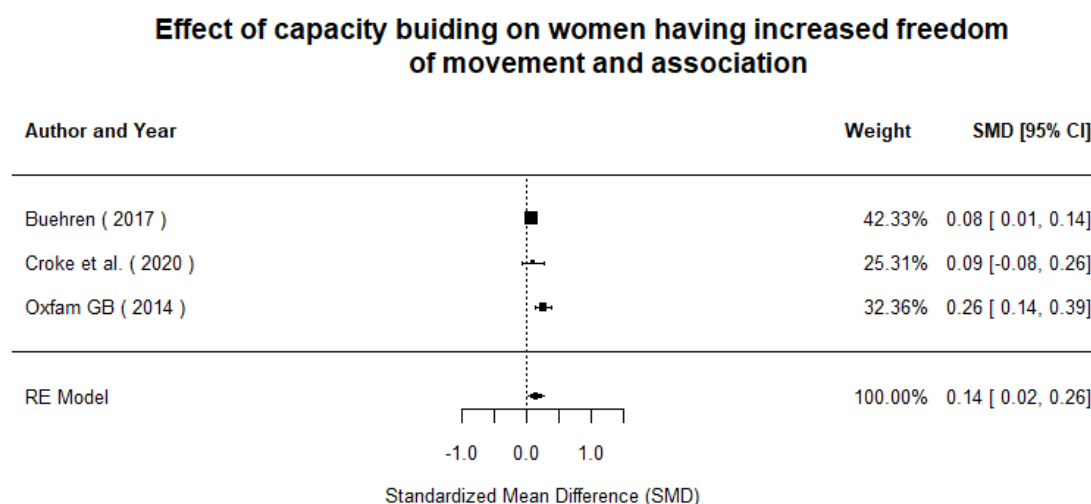
Effects of capacity building on women's success in the workplace

Bastian's (2018) experimental study in Tanzania was the only study evaluating the impact of life, social and livelihood skills and capacity building on women having improved success in the workplace. Their report included five effects that fell into this outcome category (e.g. marketing score, which is the proportion of seven business marketing practices used by the firm.). The effects ranged from very small, negative point estimates ($g = 0.06$, [95% CI: -0.00 to 0.13]) to medium, positive effects ($g = 0.35$, 95% CI [0.28 , 0.41]). We assessed the study as having low risk of bias.

Effects of capacity building on women having increased freedom of movement and association

We included a total of $k = 3$ studies in the analysis. We assessed all three studies as high risk of bias. The observed outcomes ranged from 0.08 to 0.26. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.14$ (95% CI: 0.02 to 0.26). Therefore, the average outcome differed significantly from zero ($z = 2.22$, $p = 0.03$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 150 CBBA3.

Figure 150: CBBA3: Forest plot showing the observed outcomes and the estimate of the random-effects model

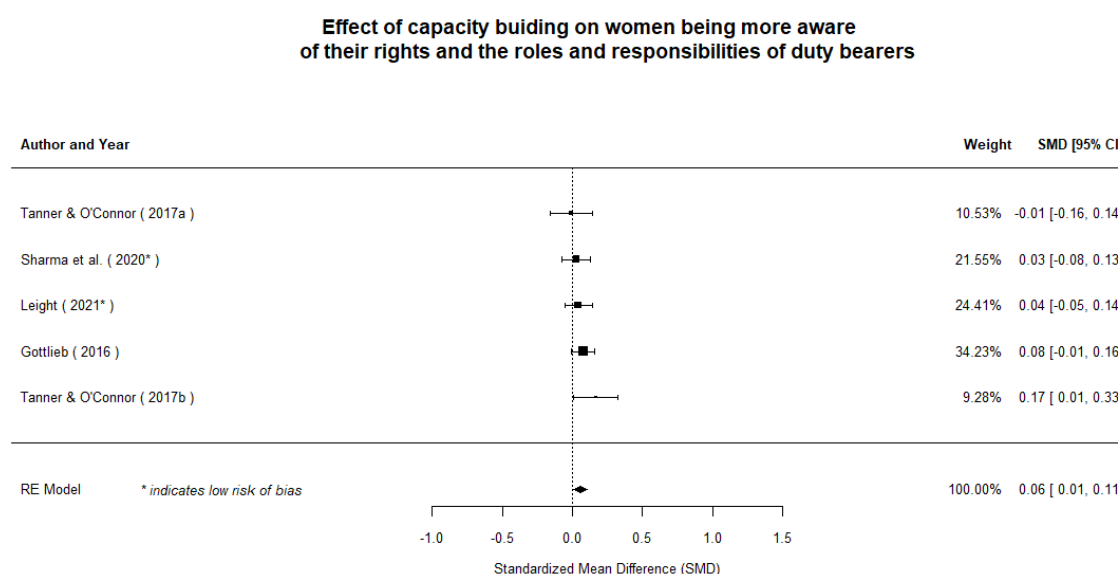


According to the Q -test, the true outcomes appear to be heterogeneous ($Q(2) = 6.63$, $p = 0.04$, $\hat{\tau}^2 = 0.01$, $I^2 = 69.82\%$). An examination of the studentised residuals revealed that one study (Oxfam GB 2014) had a value larger than ± 2.39 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be overly influential. All three studies were assessed as having a high risk of bias. With only three studies, moderation analyses and tests of publication bias were not appropriate.

Effects of capacity building on women being more aware of their rights, and the roles and responsibilities of duty bearers

We included a total of $k = 5$ studies in the analysis. We assessed two studies as low risk of bias, two as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.01 to 0.17 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.06$ (95% CI: 0.01 to 0.11). Therefore, the average outcome differed significantly from zero ($z = 2.33$, $p = 0.02$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 151 CBBA4.

Figure 151: CBBA4: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(4) = 3.17$, $p = 0.53$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$), thus we did not test for moderation. An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.58 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be overly influential.

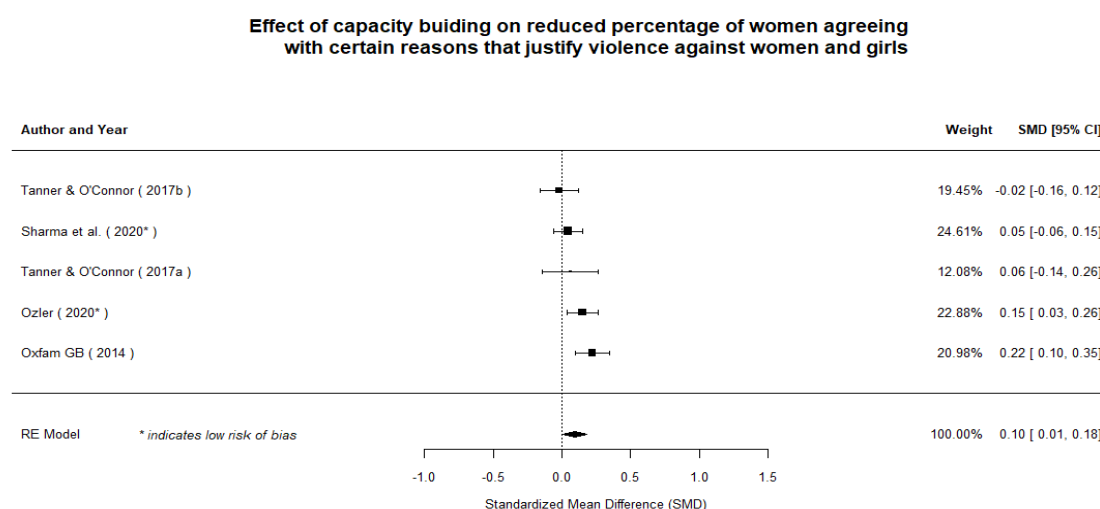
Effects of capacity building on women having more positive attitude towards taking action to claim their rights

Decker et al.'s (2018) experimental study in Malawi was the only study evaluating the impact of life, social and livelihood skills and capacity building on women having more positive attitude towards taking action to claim their rights. Their report included three effects that fell into this outcome category (e.g., sexual violence disclosure) for all schools and for primary and secondary schools separately. The effect across all schools was a very small, positive point estimate ($g = 0.03$, [95% CI: -0.03 to 0.09]), the effect across primary schools was a very small, point estimate ($g = 0.01$, [95% CI: -0.05 to 0.08]), and the effect across secondary schools was a very small, positive point estimate ($g = 0.06$, [95% CI: -0.07 to 0.18]). We assessed the study as having some risk of bias concerns.

Effects of capacity building on reduced percentage of women agreeing with certain reasons that justify violence against women and girls

We included a total of $k = 5$ studies in the analysis. We assessed two studies as low risk of bias, two as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.02 to 0.22 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.10$ (95% CI: 0.01 to 0.18). Therefore, the average outcome differed significantly from zero ($z = 2.21$, $p = 0.03$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 152 CBBA6. Tanner & O'Connor (2017a) is a study from the Democratic Republic of Congo, which Tanner & O'Connor (2017b) is a study from Ethiopia.

Figure 152: CBBA6: Forest plot showing the observed outcomes and the estimate of the random-effects model



The Q -test for heterogeneity was not significant, but some heterogeneity may still be present in the true outcomes ($Q(4) = 8.38$, $p = 0.08$, $\hat{\tau}^2 = 0.00$, $I^2 = 52.29\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.58 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

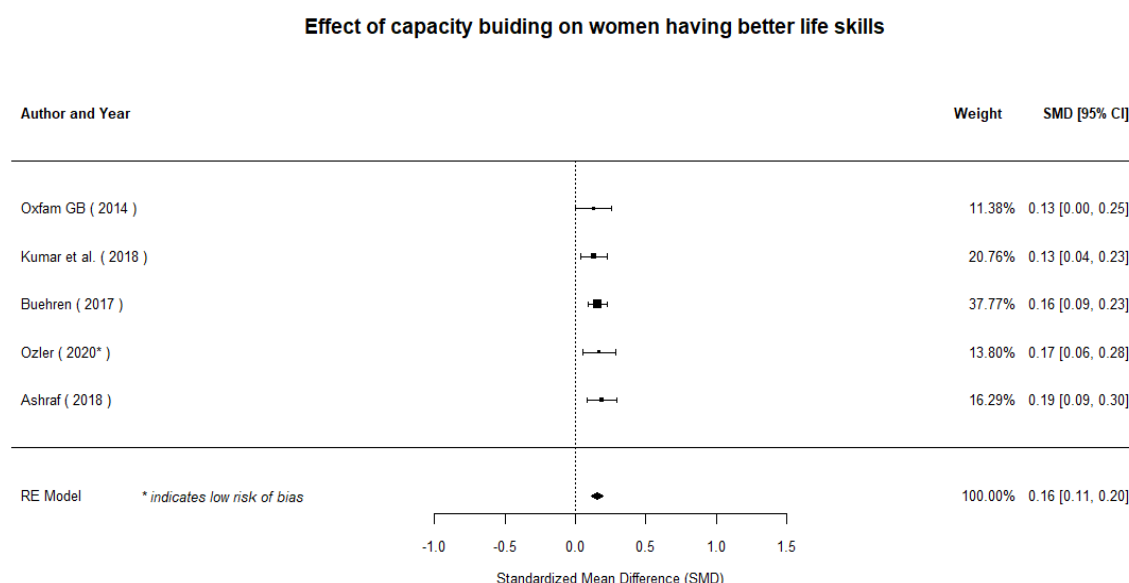
Studies with a high risk of bias were not significantly different than studies with some concerns or low risk of bias ($\hat{B} = -0.10$, $p = 0.07$ [95% CI: -0.21 to 0.01]). Exposure to intervention in months ($\hat{B} = 0.004$, $p = 0.07$ [95% CI: -0.0003 to 0.008]) and evaluation period in months ($\hat{B} = -0.002$, $p = 0.74$ [95% CI: -0.01 to 0.01]) were also not significant, nor was publication year ($\hat{B} = -0.001$, $p = 0.47$ [95% CI: -0.05 to 0.02]). The gender inequality score was a significant source of variation among effects ($\hat{B} = 1.11$, $p = 0.01$ [95% CI: 0.29 to 1.94]) such that programme effects were larger in areas with higher levels of gender inequality.

We could not test for variation in effects by study design (only 1 study used a quasi-experimental design). All models were adjusted for covariates. We also could not test by fragile state score because four of the studies had the same score. In addition, only one of the five studies targeted a mixed-gender sample and we were underpowered to test for country differences.

Effects of capacity building on women having better life skills

We included a total of $k = 5$ studies in the analysis. We assessed one study as low risk of bias, two as some concerns, and two as high risk of bias. The observed outcomes ranged from 0.13 to 0.19. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.16$ (95% CI: 0.11 to 0.20). Therefore, the average outcome differed significantly from zero ($z = 7.23$, $p < 0.01$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 153 CBBA7.

Figure 153: CBBA7: Forest plot showing the observed outcomes and the estimate of the random-effects model

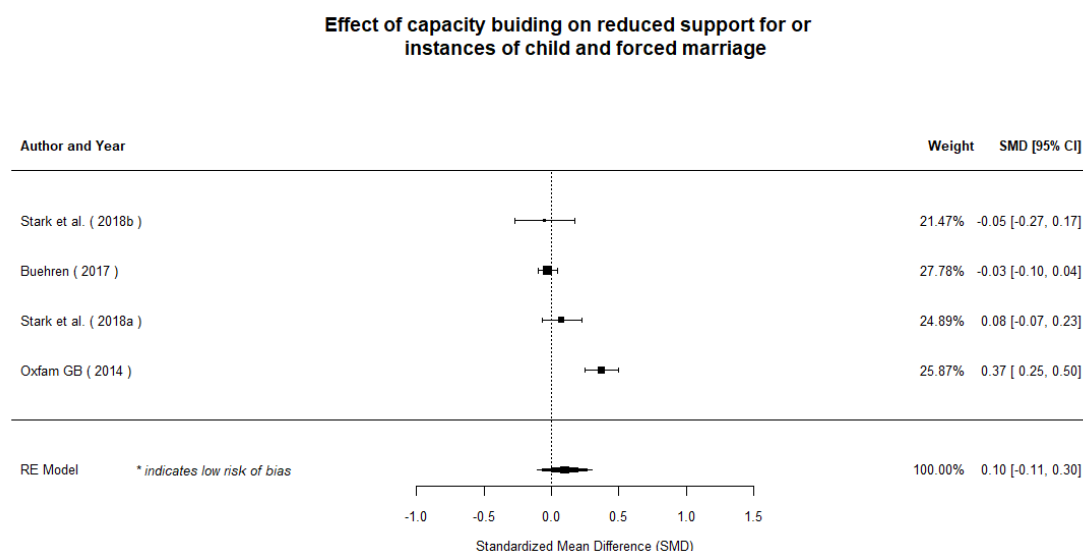


According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(4) = 0.90$, $p = 0.92$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$), thus we did not test for moderation. An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.58 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be overly influential.

Effects of capacity building on reduced support for or instances of child and forced marriage

We included a total of $k = 4$ studies in the analysis. We assessed none of the studies as low risk of bias, two as some concerns, and two as high risk of bias. The observed outcomes ranged from -0.05 to 0.37 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.10$ (95% CI: -0.11 to 0.30). Therefore, the average outcome did not differ significantly from zero ($z = 0.93$, $p = 0.35$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 154 CBBA8. Stark 2018a is a study that took place in Ethiopia, while Stark 2018b is a study that took place in the Democratic Republic of Congo. Both Stark studies were assessed as having a high risk of bias, and the other two studies were assessed as some concerns related to risk of bias.

Figure 154: CBBA8: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(3) = 30.66$, $p < 0.01$, $\hat{\tau}^2 = 0.04$, $I^2 = 90.22\%$). An examination of the studentised residuals revealed that one study (Oxfam GB 2014) had a value larger than ± 2.50 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be overly influential.

Studies with a high risk of bias were not significantly different than studies with some concerns or low risk of bias ($\hat{\beta} = 0.15$, $p = 0.50$ [95% CI: -0.29 to 0.58]). Exposure to intervention in months ($\hat{\beta} = 0.01$, $p = 0.19$ [95% CI: -0.004 to 0.02]) was also not significant. Publication year was significant such that more recent studies found smaller effects ($\hat{\beta} = -0.10$, $p = 0.01$ [95% CI: -0.17 to -0.02]), however the range of years is very small (2014 to 2018). Effects did not differ by gender inequality scores ($\hat{\beta} = 0.64$, $p = 0.81$ [95% CI: -4.52 to 5.79]), but they did differ by fragility ($\hat{\beta} = -0.02$, $p = 0.02$ [95% CI: -0.004 to -0.003]) such that programme effects were smaller in more fragile areas.

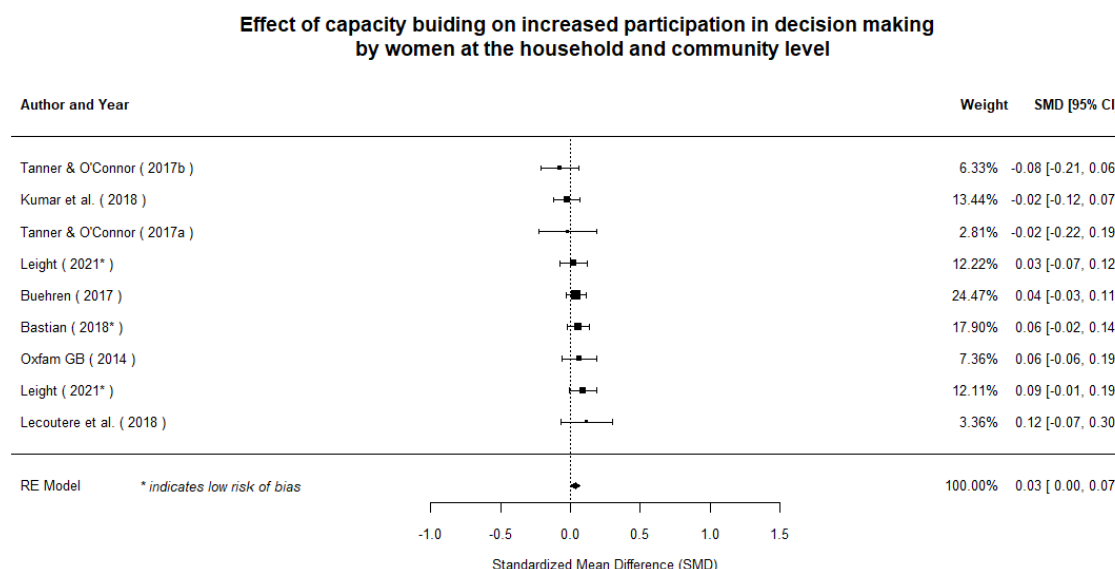
Only one study was a quasi-experimental study. We could also not test for variation by evaluation period, because three of the four studies collected data immediately after the end of the intervention. Additionally, all models were adjusted for covariates and all programmes target women only

Effects of capacity building on increased participation in decision making by women at the household or community level

We included a total of $k = 9$ studies in the analysis. We assessed two studies as low risk of bias, four as some concerns, and two as high risk of bias. The observed outcomes ranged from -0.08 to 0.12 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.03$ (95% CI: 0.001 to 0.07). Therefore, the average outcome differed significantly from zero ($z = 2.00$, $p = 0.046$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 155 CBBA1. Tanner and O'Connor (2017a) reports on a study taking place in Democratic Republic of Congo, while Tanner and O'Connor (2017b) reports on a study taking place

in Ethiopia. In addition, the two Leight (2021) effects are independent, with one effect for women ($g = 0.03$) and one for men ($g = 0.09$)

Figure 155: CBBB1: Forest plot showing the observed outcomes and the estimate of the random-effects model

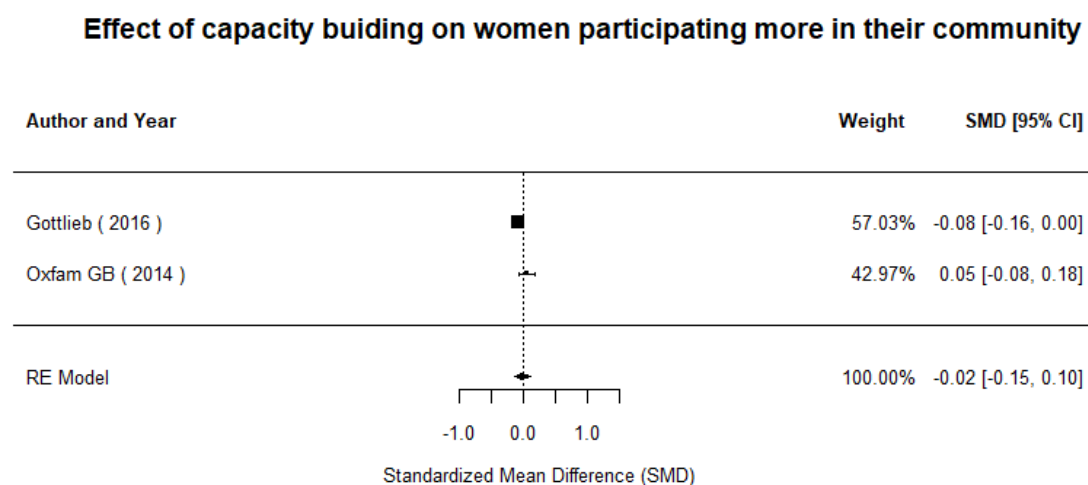


According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(8) = 7.01$, $p = 0.54$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$), thus we did not test for sources of variation. An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.77 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. With no heterogeneity, we did not test for sources of heterogeneity.

Effects of capacity building on women participating more in their community

Two studies examined the effects of capacity building on women participating more in their community, thus we included $k = 2$ studies in the analysis. We assessed both studies as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.02$ (95% CI: -0.15 to 0.10). Therefore, the average outcome did not differ significantly from zero ($z = -0.36$, $p = 0.72$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 156 CBBB2. Given the small number of studies, this result should be interpreted with caution. According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 2.82$, $p = 0.09$, $\hat{\tau}^2 = 0.01$, $I^2 = 64.49\%$). With only two studies, moderator analyses were not possible and tests of publication bias are not valid.

Figure 156: CBBB2: Forest plot showing the observed outcomes and the estimate of the random-effects model



Effects of capacity building on whether relief and recovery initiatives in conflict and post conflict situations responded to the needs of women and girls, especially vulnerable groups

Oxfam GB's (2014) quasi-experimental study in Sierra Leone was the only study evaluating the impact of life, social and livelihood skills and capacity building on whether relief and recovery initiatives in conflict and post conflict situations responded to the needs of women and girls, especially vulnerable groups. There was a very small, but not statistically significant point estimate ($g = -0.08$, [95% CI: -0.20 to 0.05]), and we assessed the study as having high risk of bias.

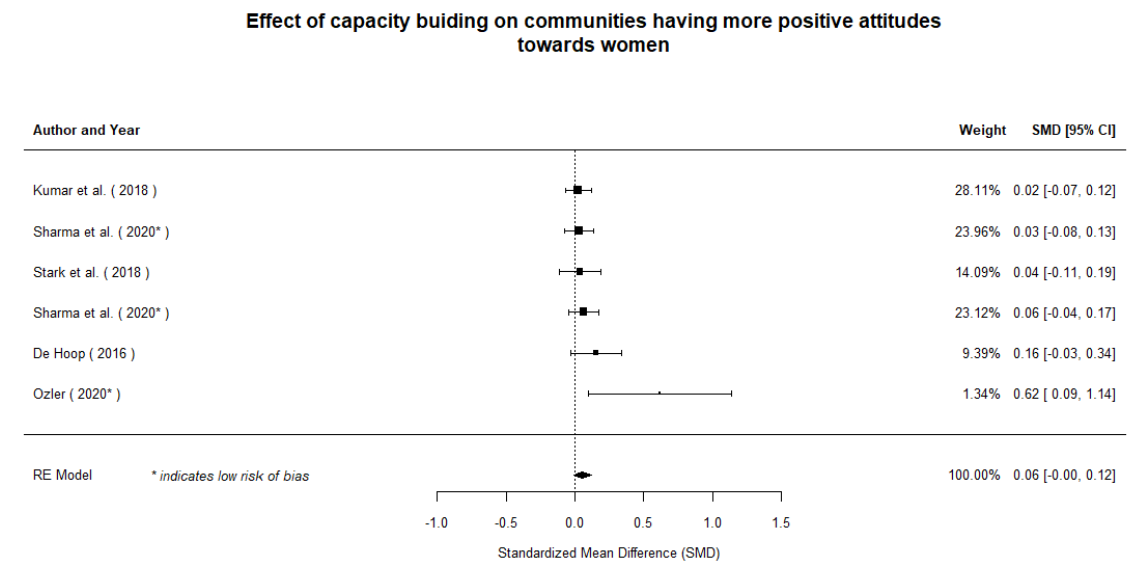
Effects of capacity building on increased community awareness of the issues affecting women

Tanner and O'Connor's (2017) experimental study in Ethiopia was the only study evaluating the impact of life, social and livelihood skills and capacity building on increased community awareness of the issues affecting women. There was a very small, but not statistically significant, point estimate ($g = -0.01$, [95% CI -0.17 to 0.15]), and we assessed the study as having some risk of bias concerns.

Effects of capacity building on communities having more positive attitudes towards women

We included a total of $k = 6$ studies in the analysis. The observed outcomes ranged from 0.02 to 0.62. We assessed two studies as low risk of bias, two as some concerns, and one as high risk of bias. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.06$ (95% CI: -0.00 to 0.12). Therefore, the average outcome did not differ significantly from zero ($z = 1.81$, $p = 0.07$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 157 CBBB2. Once again, the Sharma (2020) estimates are separate for men ($g = 0.06$) and women ($g = 0.03$).

Figure 157: CBCB2: Forest plot showing the observed outcomes and the estimate of the random-effects model

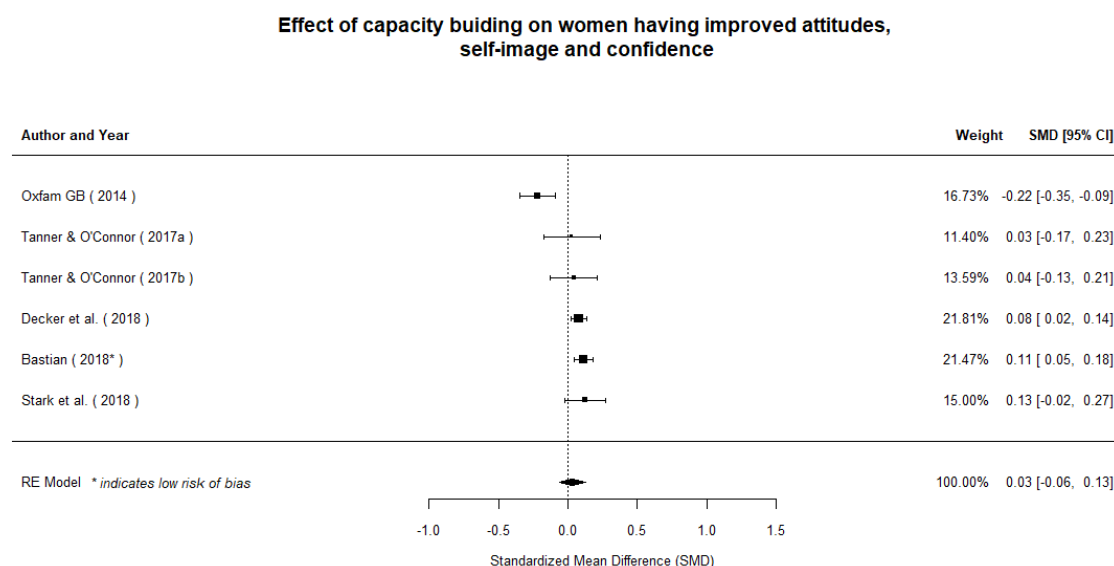


According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(5) = 6.29, p = 0.28, \hat{\tau}^2 = 0.00, I^2 = 20.52\%$), thus we did not test for moderation. An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.64 and hence there was no indication of outliers in the context of this model. According to the Cook’s distances, none of the studies could be considered to be overly influential.

Effects of capacity building on women having improved attitudes, self-image and confidence

We included a total of $k = 6$ studies in the analysis. We assessed one study as low risk of bias, four as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.22 to 0.13 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.03$ (95% CI: -0.06 to 0.13). Therefore, the average outcome did not differ significantly from zero ($z = 0.68, p = 0.49$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 158 CBCB3. One study was low risk of bias (Bastian 2018), one study was high risk of bias (OxfamGB 2014), and the remainder were assessed as some concerns related to risk of bias. Tanner & O’Connor (2017a) refers to a study from the Democratic Republic of Congo, while Tanner & O’Connor (2017b) refers to a study from Ethiopia, thus the samples are independent.

Figure 158: CBCB3: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(5) = 22.61$, $p < 0.01$, $\hat{\tau}^2 = 0.01$, $I^2 = 77.89\%$). An examination of the studentised residuals revealed that one study (Oxfam GB 2014) had a value larger than ± 2.64 and may be a potential outlier in the context of this model. According to the Cook's distances, one study (Oxfam GB 2014) could be overly influential. Indeed, sensitivity analyses leaving each study out indicated that removing Oxfam GB (2014) would increase the overall average effect ($\hat{\mu} = 0.09$ (95% CI: 0.05 to 0.13), and the effect would still be positive but would also become significant ($z = 4.46$, $p < .001$).

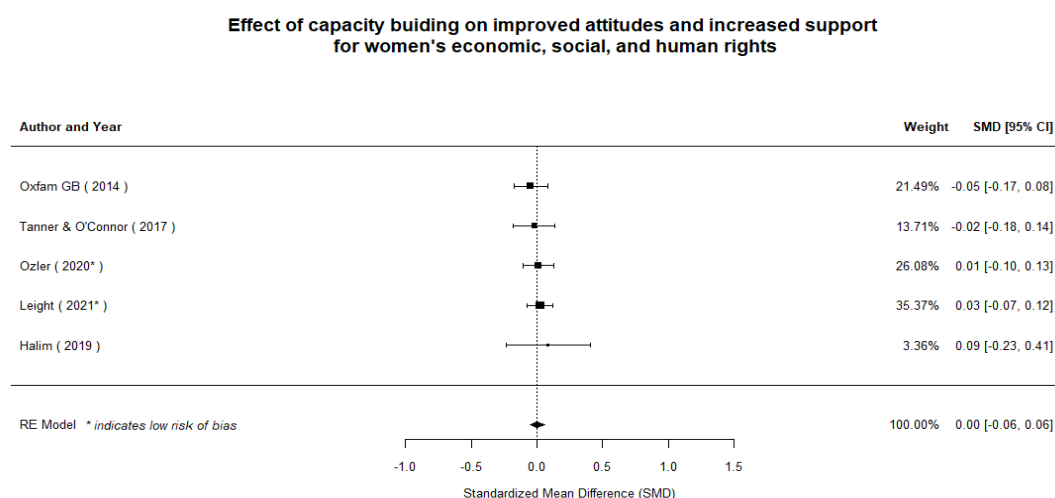
Exposure to intervention in months was significant ($\hat{\beta} = -0.01$, $p = 0.002$ [95% CI: -0.01 to -0.003]) such that longer interventions reported smaller effect (a reduction of .01 standard deviation units per month of exposure). Evaluation period in months, however, was not a significant source of variation ($\hat{\beta} = 0.01$, $p = 0.39$ [95% CI: -0.01 to 0.01]). Effects also did not differ by level of gender inequality ($\hat{\beta} = -1.40$, $p = 0.09$ [95% CI: -3.03 to 0.24]), nor by level of fragility ($\hat{\beta} = -0.002$, $p = 0.80$ [95% CI: -0.02 to 0.02]).

Only one study used a quasi-experimental design, and only one study did not adjust their model for covariates. With only one high risk of bias study, we could not examine moderation by study quality. Finally, only one study targeted mixed genders while the rest targeted women only.

Effects of capacity building on improved attitudes and increased support for women's economic, social and human rights

We included a total of $k = 5$ studies in the analysis. We assessed two studies as low risk of bias, one as some concerns, and two as high risk of bias. The observed outcomes ranged from -0.05 to 0.09 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.00$ (95% CI: -0.06 to 0.06). Therefore, the average outcome did not differ significantly from zero ($z = 0.07$, $p = 0.94$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 159 CBCB4.

Figure 159: CBCB4: Forest plot showing the observed outcomes and the estimate of the random-effects model

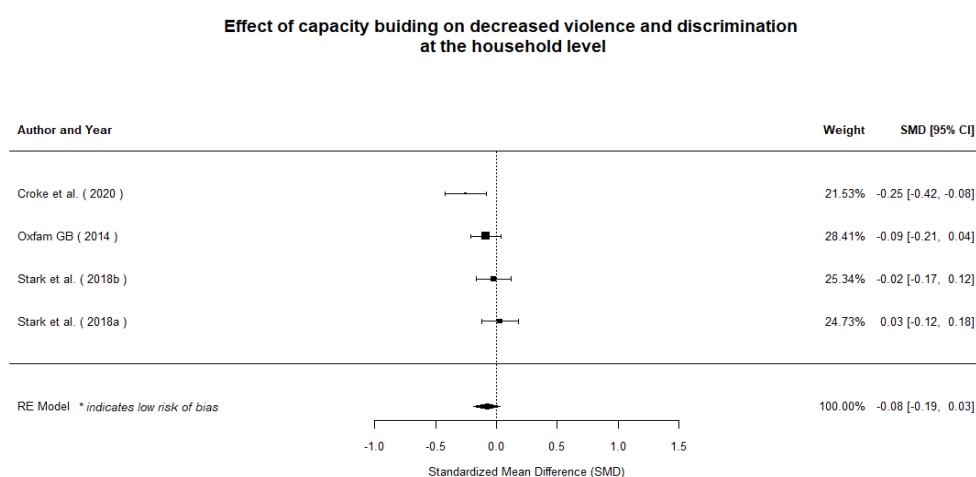


According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(4) = 1.17$, $p = 0.88$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$), thus we did not test for moderation. An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.58 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered overly influential.

Effects of capacity building on decreased violence and discrimination at the household level

We included a total of $k = 4$ studies in the analysis. We assessed none of the studies as low risk of bias, two as some concerns, and two as high risk of bias. The observed outcomes ranged from -0.25 to 0.03 . The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.08$ (95% CI: -0.19 to 0.03). Therefore, the average outcome did not differ significantly from zero ($z = -1.38$, $p = 0.17$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 160 CBCB5. Stark and colleagues (2018a) refers to a study from, while Stark and colleagues (2018b) refers to a study from Ethiopia, thus the samples are independent.

Figure 160: CBCB5: Forest plot showing the observed outcomes and the estimate of the random-effects model



The Q -test for heterogeneity was not significant, but some heterogeneity may still be present in the true outcomes ($Q(3) = 6.67, p = 0.08, \hat{\tau}^2 = 0.01, I^2 = 55.05\%$), thus we still examined potential heterogeneity. An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.50 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be overly influential.

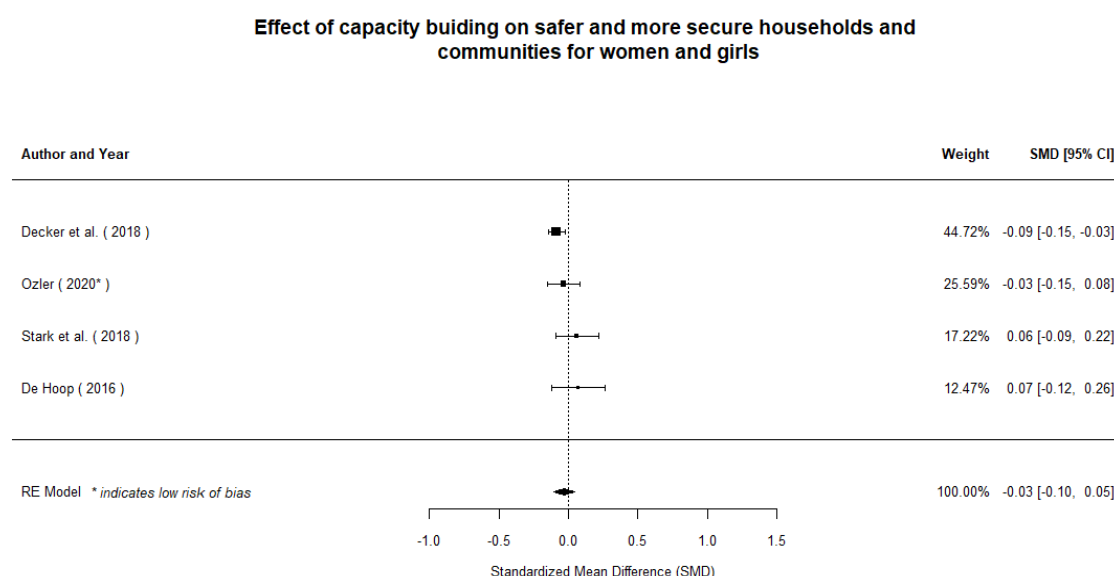
Effects from studies assessed as high risk of bias were not significantly different from studies assessed as some concerns related to risk of bias ($\hat{\beta} = -0.16, p = 0.07$ [95% CI: -0.32 to 0.01]). Exposure to intervention in months ($\hat{\beta} = -0.001, p = 0.79$ [95% CI: -0.01 to 0.01]) and publication year ($\hat{\beta} = -0.01, p = 0.65$ [95% CI: -0.08 to 0.05]) were also not significant. Finally, effects did not differ by level of gender inequality ($\hat{\beta} = -1.10, p = 0.17$ [95% CI: -2.68 to 0.48]), nor by level of fragility ($\hat{\beta} = -0.005, p = 0.54$ [95% CI: -0.02 to 0.01]).

We were not powered to test for moderation by country. In addition, only one study (Oxfam GB 2014) used a quasi-experimental design, and only one study targeted mixed genders. All studies measured the outcome immediately after the intervention was completed (thus the evaluation period was 0 for all studies), and all models were adjusted for covariates.

Effects of capacity building on safer and more secure households and communities for women and girls

We included a total of $k = 4$ studies in the analysis. We assessed one study as low risk of bias, two as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.09 to 0.07 . The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.03$ (95% CI: -0.10 to 0.05). Therefore, the average outcome did not differ significantly from zero ($z = -0.71, p = 0.48$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 161 CBCB6.

Figure 161: CBCB6: Forest plot showing the observed outcomes and the estimate of the random-effects model

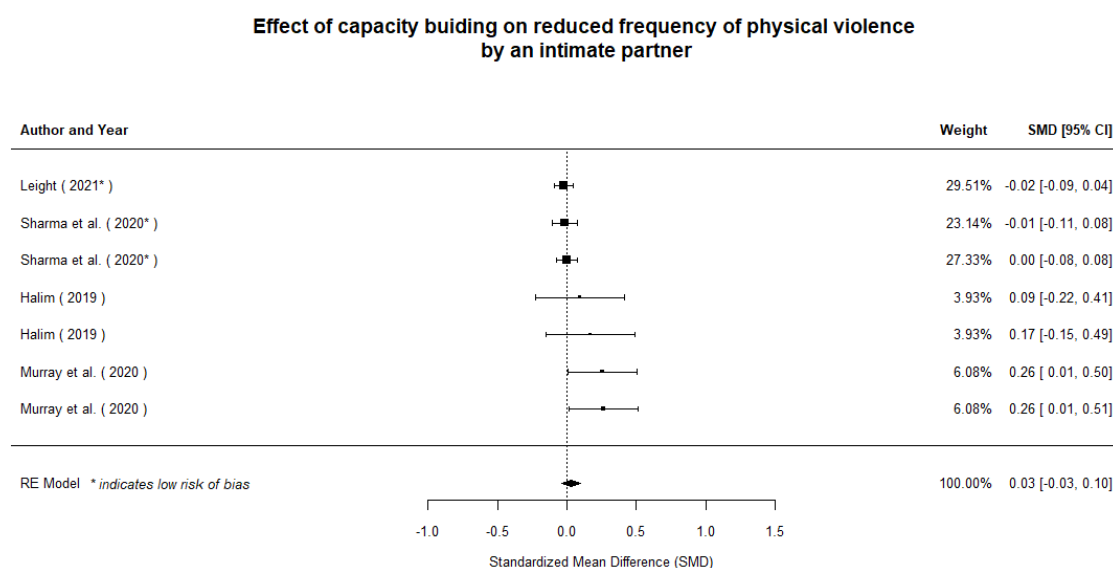


According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(3) = 5.01$, $p = 0.17$, $\hat{\tau}^2 = 0.00$, $I^2 = 40.18\%$), thus we did not test for moderation. An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.50 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be overly influential.

Effects of capacity building on reduced frequency of physical violence by an intimate partner

We included a total of $k = 7$ independent samples in the analysis. We assessed two studies as low risk of bias, none as some concerns, and two as high risk of bias. The observed outcomes ranged from -0.02 to 0.26 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.03$ (95% CI: -0.03 to 0.10). Therefore, the average outcome did not differ significantly from zero ($z = 0.93$, $p = 0.35$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 162a CBCB7pv. Three studies reported independently for men and women, thus we treated them as independent samples with separate control groups. These included Sharma and colleagues (2020; $g = -0.01$ for women and $g = 0.00$ for men), Halim (2019; $g = 0.09$ for men and $g = 0.17$ for women), and Murray and colleagues (2020; $g = .26$ for both women and men). Leight (2021) and Sharma and colleagues (2020) were assessed as low risk of bias, while the other studies were assessed as high risk of bias.

Figure 162a: CBCB7pv: Forest plot showing the observed outcomes and the estimate of the random-effects model

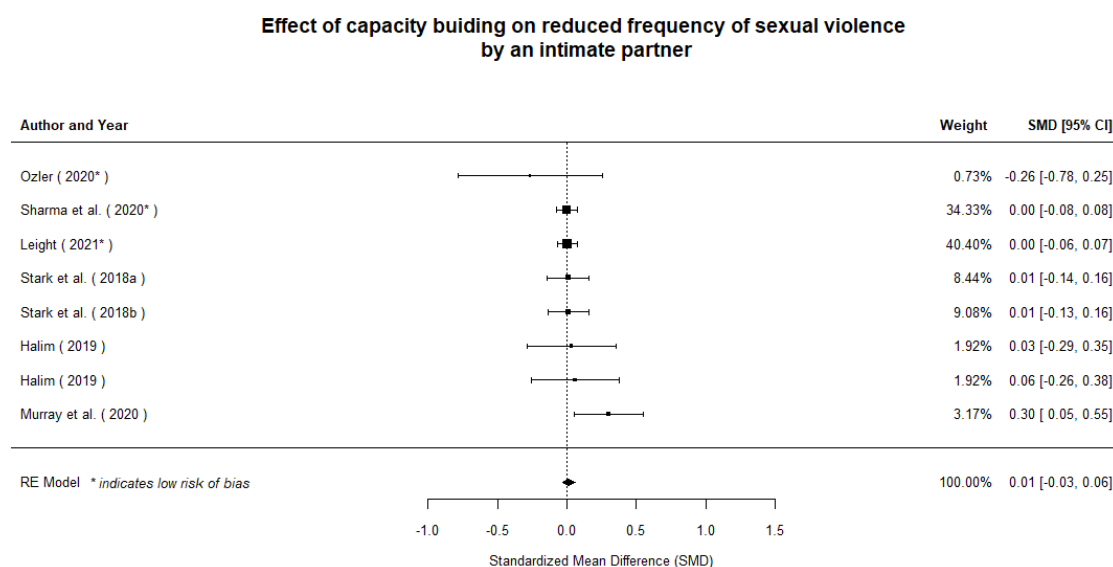


According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(6) = 10.17$, $p = 0.12$, $\hat{\tau}^2 = 0.00$, $I^2 = 40.99\%$), thus we did not test for moderation. An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.69 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be overly influential.

Effects of capacity building on reduced frequency of sexual violence by an intimate partner

We included a total of $k = 8$ studies in the analysis. The observed outcomes ranged from -0.26 to 0.30 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.01$ (95% CI: -0.03 to 0.06). Therefore, the average outcome did not differ significantly from zero ($z = 0.59$, $p = 0.56$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 162b CBCB7sv.

Figure 162b: CBCB7sv: Forest plot showing the observed outcomes and the estimate of the random-effects model

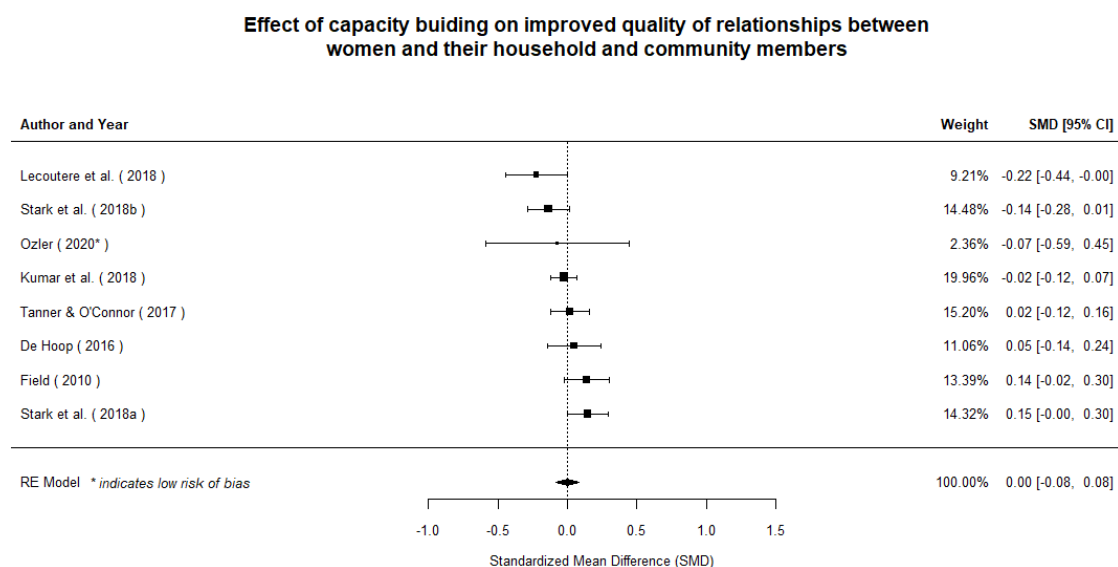


According to the Q -test, there was no significant amount of heterogeneity in the true outcomes ($Q(7) = 6.47$, $p = 0.49$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$), thus we did not test for moderation. One study (Leight 2021) had a relatively large weight compared to the rest of the studies (i.e., weight $\geq 3/k$, so a weight at least 3 times as large as having equal weights across studies). An examination of the studentized residuals revealed that none of the studies had a value larger than ± 2.73 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

Effects of capacity building on improved quality of relationships

We included a total of $k = 8$ studies in the analysis. We assessed one study as low risk of bias, six as some concerns, and one as high risk of bias. The observed outcomes ranged from -0.22 to 0.15 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.00$ (95% CI: -0.08 to 0.08). Therefore, the average outcome did not differ significantly from zero ($z = 0.03$, $p = 0.97$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 163 CBCB8. Stark and colleagues (2018a) refers to a study from, while Stark and colleagues (2018b) refers to a study from Ethiopia, thus the samples are independent.

Figure 163: CBCB8: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(7) = 14.33$, $p = 0.05$, $\hat{\tau}^2 = 0.01$, $I^2 = 51.14\%$). An examination of the studentised residuals revealed that none of the studies had a value larger than ± 2.73 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be overly influential.

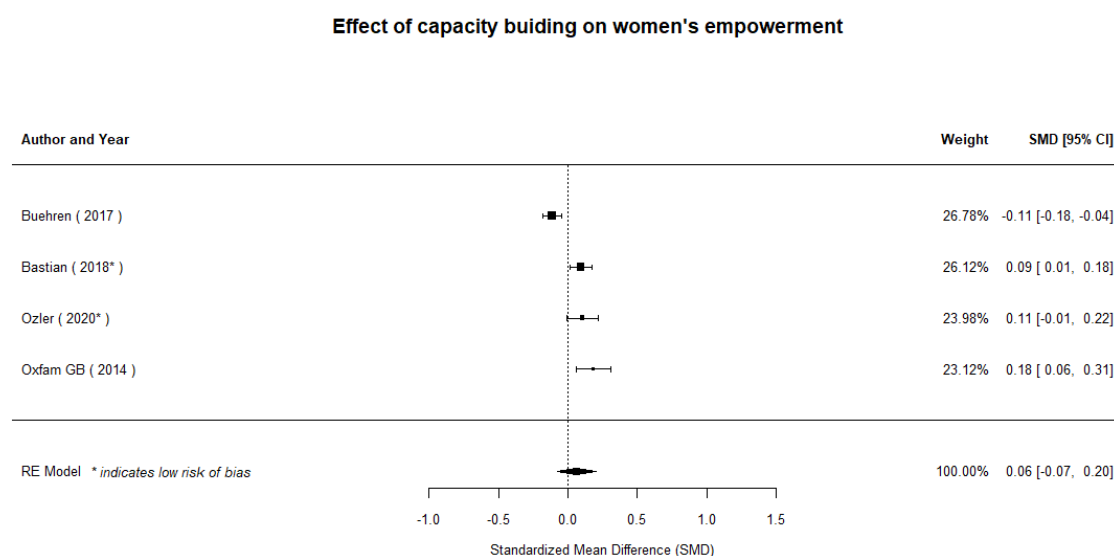
Exposure to intervention in months ($\hat{\beta} = -0.001$, $p = 0.75$ [95% CI: -0.007 to 0.005]) and evaluation period in months ($\hat{\beta} = 0.01$, $p = 0.52$ [95% CI: -0.03 to 0.06]) were not significant moderators of the effect. Nor was publication year ($\hat{\beta} = -0.02$, $p = 0.14$ [95% CI: -0.05 to 0.01]). Effects also did not differ by scores on the gender inequality index ($\hat{\beta} = -0.50$, $p = 0.62$ [95% CI: -2.45 to 1.46]) or the fragile states index ($\hat{\beta} = -0.01$, $p = 0.13$ [95% CI: -0.01 to 0.002]).

Only one study was assessed as having high risk of bias (De Hoop 2016), thus we could not conduct sensitivity analysis by study quality. Only one model was not adjusted for covariates, and only one study did not take place in Africa. All studies used an experimental design.

Effects of capacity building on women's empowerment index

We included a total of $k = 4$ studies in the analysis. We assessed two studies as low risk of bias, none as some concerns, and two as high risk of bias. The observed outcomes ranged from -0.11 to 0.18 . The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.06$ (95% CI: -0.07 to 0.20). Therefore, the average outcome did not differ significantly from zero ($z = 0.90$, $p = 0.37$). A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 164 CBCC1.

Figure 164: CBCC1: Forest plot showing the observed outcomes and the estimate of the random-effects model



According to the Q -test, the true outcomes appear to be heterogeneous ($Q(3) = 25.95$, $p < 0.01$, $\hat{\tau}^2 = 0.02$, $I^2 = 88.44\%$). An examination of the studentised residuals revealed that one study (Buehren 2017) had a value larger than ± 2.50 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be overly influential.

We were able to test several moderators in the context of this model. Studies with a high risk of bias were not significantly different than studies with low risk of bias ($\hat{\beta} = -0.04$, $p = 0.61$ [95% CI: -0.18 to 0.11]). Exposure to intervention in months ($\hat{\beta} = -0.0001$, $p = 0.99$ [95% CI: -0.01 to 0.01]) was not a significant source of heterogeneity, but evaluation period was ($\hat{\beta} = -0.02$, $p < 0.001$ [95% CI: -0.02 to 0.01]) such that smaller effects were observed as the time between the end of the intervention and data collection increased (specifically, the effect reduced by .02 standard deviation units for each month between the end of the intervention and the data collection). Publication year was not a significant source of variation among effects ($\hat{\beta} = -0.01$, $p = 0.78$ [95% CI: -0.08 to 0.06]). Finally, effects also did not differ by scores on the gender inequality index ($\hat{\beta} = 0.40$, $p = 0.43$ [95% CI: -0.60 to 1.41]) or the fragile states index ($\hat{\beta} = 0.003$, $p = 0.61$ [95% CI: -0.01 to 0.01]).

Only one study used a quasi-experimental design. In addition, all models were adjusted for covariates, all studies took place in Africa and all interventions target women only, thus none of these could be tested as potential sources of variation.

Qualitative findings

We conducted a thematic synthesis on the 16 linked qualitative studies to the included LSCB and capacity building interventions. As indicated above, this thematic synthesis aims to identify themes related to the interplay of intervention design, intervention implementation, target population, and contextual variables with intervention outcomes and effects. In total, we identified 27 descriptive themes, which we configured into six analytical themes (Online Appendix N). These six analytical themes present the synthesis results and are discussed in more detail below.

Theme 1: Contextual cultural norms determined the available space for LSCB capacity building

The qualitative synthesis identified a strong theme on the role of cultural norms across contexts. Norms on women's mobility, decision-making in the household and gendered occupation mitigated the space in which LSCB capacity building could be designed and implemented (Lecoutere and Wuyts 2020; Messersmith et al 2015; Gibbs et al 2018; Tanner and O'Conner 2017). When content of LSCB programmes contrasted existing local norms, conflict often followed and closed the space in which the programme could operate. The qualitative evidence points to possible design strategies to mitigate this issue, which include adjusting content and targeted empowerment outcomes to reduce the gap with the prevailing norms (Gibbs et al 2018; Tanner and O'Conner 2017).

The 'wall of patriarchy' is the metaphor used by Lecoutere and Wuyts (2020) to describe the barriers to women's decision-making power within the household and to their empowerment. Evaluating the Gender Household Approach intervention in Uganda, a programme introducing a Participatory Intra Household Decision Making process via family counselling and couple seminars, and couple coaching, addressing entrenched patriarchal norms required a more structural programme design and approach in order to affect change. Women participants reported positive experiences with the programme overall, but also indicated no structural shifts in power in their household. They reported that *"...if a woman's decision-making power is restricted by the morals of the man, then yes, that is difficult to change [...] Some women are better off because they cooperate more with their husbands. [...] It came as a surprise that I acquired the power to make decisions over [strategic household] needs. My husband just changed"* (Lecoutere and Wuyts 2020).

These findings were echoed by the Women for Women International's (WfWI) What Works to Prevent Violence Programme in Afghanistan. In this context, restrictive norms, particularly within religious structures, strongly constrained the programme space and potential to affect intended empowerment outcomes (Gibbs et al 2018; Huber and Zupancic 2015). Participants reported a lack of mobility and permission of male household heads to engage in income generating activities as the main barriers to their ability to apply newly gained skills in their daily lives. The qualitative evaluation concluded that *"...women's lives remained structured by the wider constraints that they lived under,"* and that the programme could have had a stronger impact if it was designed to engage family members and to address structural barriers at the household level (Huber and Zupancic 2015).

We observed related themes in the Tanzanian context in the Together to End Violence Against Women (TEVAW) programme. Community level norms undermined efforts to encourage women's reporting of gender-based violence (Messersmith et al 2015). Furthermore, in the Congolese context with the implementation of the Opportunities through Mentoring, Parental Involvement and Safe Spaces (COMPASS) programme, implementers required careful awareness raising and negotiating to convince community leaders to accept the provided life skills content (Tanner and O'Conner 2017). From a design perspective, the qualitative evidence suggested that restrictive norms and practices are difficult to change and need to be specifically targeted by dedicated programme components. The assumption that skills provision and awareness raising, for example, create spill-over effects on empowerment outcomes was consistently rejected.

More emphasis should be placed on design components that address the issue of restrictive norms and practices more directly based on a comprehensive and holistic concept of empowerment.

Theme 2: Embedding the programme design with existing cultural norms and implementing through this embeddedness enhances acceptability and the space for programme implementation in restrictive contexts

Linked to the first theme was a rich set of data on the more strategic programme orientation that life skills interventions in restrictive contexts could explore. The themes in particular emerged from the programmes in Afghanistan and DRC referenced above (Huber and Zupancic 2015; Tanner and O'Conner 2017). They explored to what extent embedding programmes within existing restrictive contexts was strategic and defensible. Programme data from both contexts indicated that life skills programmes can enhance their reach in restrictive contexts by working with local actors that share existing social norms. This can refer to religious structures (in Afghanistan), existing sub-societies (linked to FGM in DRC), and community leaders (in DRC and Tanzania).

Working with refugees living in camps on the Sudan/Ethiopia border, conflict-affected communities in eastern Congo, and displaced populations in north-west Pakistan, the COMPASS programme faced opposition from parents and community leaders in all countries. Participants and stakeholders questioned the appropriateness of the programme activities (Tanner and O'Conner 2017). Consensus was achieved and implementation granted only through extensive discussion, house-to-house visits and negotiations with community leaders, local authorities and parents. Through trust and legitimacy, the programme was able to collaborate with leaders of 'secret' societies associated with FGM, enhancing its reach to participants it would have not been able to access.

A similar theme emerged in the evaluation of the What Works to Prevent Violence Programme in Afghanistan, which pointed at the willingness to work with local religious institutions as a main facilitator to programme reach and impact (Huber and Zupancic 2015; Gibbs et al 2020). The evaluation illustrates that *"...Women's rights are discussed from the Islamic perspective. With the training of mullahs, the WfWI team seeks to frame the discussions around women's rights as they relate to Islam, highlighting particular verses from the Quran that promote women's position in society. This can be viewed as generally in line with growing best practice in Afghanistan, where Islamic perspectives on women's rights are gaining ground among various stakeholders."*

However, a careful balance was struck to manage the level of integration of programmes and to guard against co-option by these very structures. For instance, despite overall encouraging results, the COMPASS programme did not change women's own perceptions about restrictive norms. Women continued to hold attitudes that indicated their acceptance of gender inequality and SGBV. And, in the Afghanistan context, the programme resorted to paying men double the stipend it paid women, even though women were the main target beneficiaries. This led to questions around the programme's own commitment to gender equality (Huber and Zupancic 2015).

Theme 3: Ambiguous role of men in programme design

A key design question for life skills and capacity-building refers to whether to create women-only spaces or to also involve men as participants. We identified a mixed

qualitative evidence base with data from different contexts and programmes, painting a rich picture (Lecoutere and Wuyts 2020; Gibbs et al 2018; Tanner and O’Conner 2017). Overall, while women-only programme designs were welcomed by women programme participants, programme implementers and decision-makers emphasised that the lack of involvement of men in programmes impeded their impact. This appeared to be in particular the case in restrictive social settings where women are seen as unable to translate empowerment objectives in household decision-making unless men are participating in the programme.

However, in the COMPASS programme, women participants across contexts reported a strong preference for single-sex spaces. In Ethiopia, girls appreciated their protected space and even requested a further safe space which not even the life coach had access to. The safe space was seen as important in the refugee context and provided adolescent girls with a place to feel safe, learn and make friends. In addition, participants used it to report instances of SGBV and perceived strong ownership and a perception that men should not be allowed to partake. As the programmes progressed in each country, participants welcomed the inclusion of mothers and other caregivers (Tanner and O’Conner 2017).

In contrast, data from the Uganda and Tanzania programmes reported programme limitations of not involving men participants. Lecoutere and Wuyts’ (2020) argue that any attempts of breaking the ‘*Wall of Patriarchy*’ requires the support and involvement of men. The programme, however, reported a lack of structural changes despite designing combined elements such as couple seminars and coaching to change intra household decision-making; the qualitative data does illustrate this dilemma outlining different pathways to structural empowerment for women, all of which require the involvement and co-operation of the men in the household. Likewise, a key finding of the evaluation of the What Works to Prevent Violence Programme in Afghanistan pointed to its limitation of only working with individual women. This was a key explanation for the lack of structural change in women’s lives and recommended that “...*engagement with male family members may prove a more effective means of facilitating women’s social and economic empowerment*” (Huber and Zupancic 2015).

Theme 4: LSCB capacity building programmes and their design benefit from the inclusion of and coupling with TVET programmes

The qualitative evidence on programme design suggests integration of LSCB and TVET capacity building programme to be positive. The assessments of stand-alone LSCB programmes and LSCB components within TVET programmes emphasised the interconnectedness of both types of skill sets. Programme participants expressed demand and need for both types of skills in order to enhance their life capabilities and opportunities. Within LSCB programmes, participants expressed a need for practical applications and a pathway for how new skills could lead to economic opportunities. Within TVET programmes, participants cited obstacles to the application of these skills to further their life being constrained by personal and structural issues.

In Uganda, women participants of a participatory intrahousehold decision-making programme expressed a desire for agricultural skills, including farm management and cash crop cultivation, such as coffee. This knowledge was a means to economic development and contribute to shared household decision-making and resource

allocation (Lecoutere and Wuyts 2020). In Afghanistan, participants had access to a range of TVET linked programme components, such as cash crop farming, animal husbandry, and embroidery within the wider empowerment programme. Participants regarded the start-up capital and these elements as a main benefit (Gibbs et al 2018). Likewise, in the COMPASS programme, career planning and business development were the highest rated training sessions.

This general theme of wanting to apply gained skills in daily contexts reflected a desire for practical tools and applications to further one's capabilities and personal development and growth. In combination with the qualitative data on TVET programmes, which mirror this theme, this points to the direction that capacity building design can benefit from intervention packages with multiple components aimed at structural change.

Theme 5: Group and peer-to-peer design elements can target structural barriers to empowerment via social capital and collective action

The qualitative evidence indicated that life skills programmes that emphasised group and "peer-to-peer" exchanges such as mentoring, and the formation of women-only groups reported enhanced social capital and networks of women. Enhanced social capital can better target structural barriers to empowerment. This data emerged from the three-country COMPASS programme which combined mentoring with parental involvement and safe spaces (Tanner and O'Conner 2017) and the What Works to Prevent Violence Programme in Afghanistan, which created single sex networking spaces as part of the training programme (Gibbs et al 2018; Huber and Zupancic 2015).

In both programmes, themes centred on the benefit of having access to women peers for mentoring and learning from each other. This referred to both fellow programme participants and to the trainers and formal mentors provided by the programmes. The ability to learn from women in similar contexts and based on shared experiences and realities was a key benefit of the mentoring arrangements. In the COMPASS programme, mentors were carefully selected to be close in age to adolescent girl participants, from the same area, and to hold positive attitudes toward girls. In Afghanistan, women reported the programme to have been one of few opportunities to build a network outside of their families with one participant explaining that: *'My life has improved. The year that I was in this course, when we came here once a week, we met other women and made friends, it was a good change for me'* (Gibbs et al 2018).

In both programmes a transition from 'shyness' to 'confidence' was reported as women started to develop strong peer-to-peer relationships with each other and/or their mentors. In Ethiopia, this journey developed between mentor and mentees with both groups gaining in confidence as their relationship progressed, that is mentors, too, reported a growth in confidence and self-worth from the success of the relationship and seeing their mentees progress. In Pakistan, women indicated their increased confidence to partake in household discussion and the interest of other women in their learning and growth. An illustration of this process was captured in this account: *"When I go back to my home, I discuss what I learnt with people, and they praise me for being a scholar and worship what I have learnt while going to the course. It really gives great value to my talking and they stop working and listen while I am talking. All my friends, especially my sisters, have an interest in cow-keeping. (...). In fact, they showed their interest in this course and further praised me a lot by saying: You are talented and have learnt so much so far."*

Finally, in both programmes, women indicated instances on how these peer-to-peer relationships might be able to grow into collective efforts to support each other. In the COMPASS programme, participants formed dedicated social networks that were observed to prevail beyond the programme. The participants reported stronger support networks, access to trusted non-family females, and a stronger sense of companionship. In addition, mentees provided mentoring services for the next cohort of participants in Ethiopia and DRC. In Afghanistan, women indicated forming social safety nets for each other by staying in touch post the intervention and supporting each other's business ventures. The evaluation concluded that *"...the structure and approach to building social safety nets that brought women together in accepted spaces with the support of community leadership was in line with best practice in Afghanistan and appears to have facilitated opportunities for the effective development of safety nets for participants"* (Huber and Zupancic 2015).

Theme 6: A bundle of key implementation considerations apply across contexts

Across the LSCB programmes several implementation considerations were reported repeatedly as contributing factors to programme success. These implementation considerations are programmatic and granular. They are a bundle of considerations suggested across individual studies along the implementation cycle:

- Design and implement a sensitive and targeted awareness-raising and outreach campaign to attract participants into the programme. For example, the COMPASS programme managed to build trust and legitimacy with the programme's constituencies through an in-depth awareness raising campaign including house-to-house visits, community meetings, parents' meetings, and personal sessions with community and opinion leaders. This implementation factor is particularly important in restrictive social contexts.
- Identify and select a central and safe location that is easily accessible (e.g., via public transport, within walking distance of participants). For community acceptance, using an established and visible location can be a contributing factor. In Pakistan the training venue and safe space became more acceptable after it provided embroidery activities.
- The provision of childcare facilities was a highly desired aspect by programme participants and determined programme participation in Sudan/Ethiopia, DRC, Uganda and Pakistan.
- Training curricula and materials require careful adaptation to local context and translation into local languages. Participants of both the 3-country COMPASS and What Works to Prevent Violence Programme in Afghanistan raised concerns about the lack of adapted materials and content.
- A key implementation element and driver of programme satisfaction was the skilfulness and approach of programme trainers. In the COMPASS programme, participants indicated positive feedback on the use of women as trainers in the single-sex programme.
- Evaluation reports of a range of intervention outline the potential benefits of rapid programme iteration based on collected data. The COMPASS programme explicitly reported this as an applied implementation criterion. Sourcing ongoing participant feedback and consultation generated data that was used instantly to adapt programme design, for example identifying programme locations.

Discussion

The quantitative meta-analysis revealed that LSCB interventions had positive significant effects on building women's life skills and women's capacity to understand and use financial, banking, and business services effectively. In addition, positive significant but marginal effects were observed on women's participation in decision making at the household and community level; reduced agreement with reasons to justify violence against women and children; and women's awareness of their rights. The remainder of outcomes were non-significant.

Overall, the majority of community-level outcomes, such as communities having more positive attitudes towards women, and being safer for women are non-significant and the same findings are established regarding empowerment outcomes such as decreased violence and discrimination at the household level and direct measures of women's empowerment. This is in contrast to individual outcomes related to women's skill development. The team used a complementary qualitative synthesis to explore explanations and recommendations to design more effective LSCB interventions that can go beyond narrow individual outcomes on skills development and targeting broader empowerment objectives and community level changes.

- A possible explanation for the limited effects on community level outcomes and downstream empowerment effects refers to the role of social norms and restrictive contexts that determine the space in which programmes can operate. The qualitative themes discussed in detail how programmes were limited in restrictive social contexts and not able to target a full spectrum of empowerment outcomes. In particular, programme implementer had to strike a careful balance between the depth of embedding programmes within restrictive context in order to be able to operate and reach women and the adjustment of programme empowerment outcomes in order not too conflict with existing norms and practices too strongly. A reasonable observation is that the higher the level of embeddedness, the lower the scope to target radical empowerment outcomes in the short term (and in the evaluation timeframes). This does not negate the argument that this strategy translates into success in the medium- to long-term and often was the only available programme path to at least improve women's' lives on a narrower scale related to individual skill development.
- A design implication from the qualitative themes on the programme limitations by restrictive contexts and norms is the need to develop complementary programme components that address these issues directly. More holistic programme design, involving multiple components addressing root causes as well as providing livelihood support and development, would be required for this from a design perspective. Specific suggestions from the qualitative evidence base points to programme components that foster women's peer-to-peer exchange and the building of a collective. This refers to designing mentoring elements into programmes and creating (or targeting existing) women's self-help groups and clubs in which programme activities can be delivered. Such instruments seem to hold potential to build social capital, enhance women's networks, and to develop strategies for activism targeting broader empowerment outcomes.
- The qualitative evidence synthesis also identified themes around the role of men in programme implementation. While women themselves preferred single-sex programmes in which men were not involved, programme implementers and

decision-makers identified the lack of men's engagement in programmes as a key barrier to empowerment outcomes and an explanation for why the positive effects on skill development did not translate into social change. We tested these qualitative findings against quantitative moderators on TVET and capacity building interventions but this was not supported by the quantitative data because of a lack of available data and this did not lead to significant results among the capacity building data.

Summary of findings and discussion

We included 19 studies in 13 countries in South Asia and Sub-Saharan Africa that evaluated the effect of life, social and livelihood skills and capacity building. We were able to examine effects on the following outcomes: Increased capacity of women to understand and use financial, banking, and business services effectively, women have increased access and ownership to assets, credit and income, women have more and better control over their bodies and sexual health, women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them and women have improved attitudes, self-image and confidence. Our included studies report against all of the secondary outcomes (Resources, Agency and Achievement) and eight of the nine immediate outcomes for our review. Overall, the GRADE assessments generally indicate a very low to low (with one moderate) certainty in this body of evidence.

Furthermore, the 16 identified qualitative linked studies provided six key analytical themes to buttress the quantitative findings. The success of LSCB programming is often determined by contextual and societal factors that can make or break the success of a programme. This means that embedding programming within these existing norms can enhance a community's acceptance of the offered programme, but at the same time limit the types of empowerment outcomes targeted. Similar to our findings in TVET interventions, group mentorship and peer exchange offer a route to collective action and socialisation that can encourage group uptake of a LSCB programme. Again, aligning with features of safe spaces programmes, LSCB programmes can see increased success when meeting spaces are secure and reachable by participants. Table 57 presents a summary of the quantitative results alongside our ratings of the certainty of the evidence.

Table 57: GRADE summary of findings and certainty of evidence on life, social and livelihood skills and capacity building

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(AA1) Women have access to rights, services and opportunities.										
1	RCT	serious ^a	serious ^b	not serious	serious ^c	none	737	SMD 0.02 SD higher (0.12 lower to 0.16 higher)	⊕○○○ VERY LOW	Limited Importance
(AB1) Increased capacity of women to understand and use financial, banking, and business services effectively										
3	RCT – 3	serious ^d	serious ^e	not serious	not serious	none	4991	SMD 0.14 SD higher (0.01 higher to 0.28 higher)	⊕⊕○○ LOW	Important, but not critical
(AB2) Women have increased access and ownership to assets, credit and income.										
8	RCT – 6 QED – 2	very serious ^f	not serious	not serious	not serious	none	9970	SMD 0.09 SD higher (0.02 higher to 0.16 higher)	⊕⊕○○ LOW	Important, but not critical
(AC2) Initiatives supported that facilitate women to access decent work (formal and informal employment), including people with disabilities										
1	RCT	very serious ^h	serious ^b	not serious	not serious	none	3219	Two negative and three positive effect estimates with a 95% CI range of -0.08 to 0.16	⊕○○○ VERY LOW	Important, but not critical
(AC3) Improved capacity of women entrepreneurs										
1	RCT	very serious ⁱ	serious ^b	not serious	not serious	none	3219	SMD 0.02 SD higher (0.09 lower to 0.05 higher)	⊕○○○ VERY LOW	Important, but not critical
(BA1) Women have improved success in the workplace.										
1	RCT	not serious	serious ^b	not serious	serious ^j	none	3543	Five positive effect estimates with a 95% CI range of -0.00 to 0.41	⊕⊕○○ LOW	Important, but not critical
(BA2) Women have more and better control over their bodies and sexual health.										
7	RCT - 7	serious ^k	serious ^l	not serious	not serious	none	8359	SMD 0.06 SD higher (0.02 lower to 0.14 higher)	⊕⊕○○ LOW	Critical
(BA3) Women have increased freedom of movement and association.										
3	RCT – 2 QED – 1	very serious ^m	serious ⁿ	not serious	not serious	none	5233	SMD 0.14 SD higher (0.02 higher to 0.26 higher)	⊕○○○ VERY LOW	Limited Importance
(BA4) Women are more aware of their rights and the roles and responsibilities of duty bearers.										
5	RCT – 5	serious ^o	not serious	not serious	not serious	none	6565	SMD 0.06 SD higher (0.01 higher to 0.11 higher)	⊕⊕⊕○ MODERATE	Limited Importance
(BA5) Women have more positive attitude towards taking action to claim their rights.										
1	RCT	serious ^p	serious ^{b,q}	not serious	not serious	none	4278	Six positive effect estimates with a 95% CI range of -0.07 to 0.19	⊕⊕○○ LOW	Limited Importance
(BA6) Reduced percentage of women agreeing with certain reasons that justify violence against women and girls										
5	RCT – 4 QED – 1	serious ^r	serious ^s	not serious	not serious	none	4736	SMD 0.1 SD higher (0.01 higher to 0.35 higher)	⊕⊕○○ LOW	Critical
(BA7) Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them.										
4	RCT – 4 QED – 1	very serious ^t	not serious	not serious	not serious	none	6750	SMD 0.16 SD higher (0.11 higher to 0.20 higher)	⊕⊕○○ LOW	Important, but not critical

Certainty assessment							Sample size	Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
(BA8) Reduced instances of child or forced marriage										
4	RCT – 3 QED – 1	very serious ^u	serious ^v	not serious	serious ^w	none	5178	SMD 0.1 SD higher (0.11 lower to 0.3 higher)	⊕○○○ VERY LOW	Critical
(BB1) Increased participation in decision making by Women at the household or community level, including during crisis response										
9	RCT – 8 QED – 1	serious ^x	not serious	not serious	not serious	none	14914	SMD 0.03 SD higher (0.0 higher to 0.07 higher)	⊕⊕⊕○ MODERATE	Limited Importance
(BB2) Women participate more in their community.										
2	RCT – 1 QED - 1	very serious ^y	not serious	not serious	not serious	none	3203	SMD 0.02 SD lower (0.15 lower to 0.1 higher)	⊕⊕○○ LOW	Limited Importance
(CA8) Increased community support for women’s and children’s human, economic and legal rights										
1	RCT	very serious ^{ab}	serious ^b	not serious	serious ^{ac}	none	968	SMD 0.07 SD lower (0.2 lower to 0.05 higher)	⊕○○○ VERY LOW	Important, but not critical
(CB1) Increased awareness in communities of the issues affecting women										
1	RCT	serious ^z	serious ^b	not serious	serious ^{aa}	none	614	SMD 0.01 SD lower (0.17 lower to 0.15 higher)	⊕○○○ VERY LOW	Limited Importance
(CB2) Communities have a more positive attitude towards women/marginalised groups.										
4	RCT – 4	serious ^{ad}	not serious	not serious	not serious	publication bias strongly suspected	3872	SMD 0.07 SD higher (0.01 lower to 0.15 higher)	⊕⊕○○ LOW	Limited Importance
(CB3) Women have improved attitudes, self-image and confidence.										
6	RCT – 5 QED – 1	serious ^{ae}	very serious ^{af}	not serious	serious ^{ag}	publication bias strongly suspected	10481	SMD 0.03 SD higher (0.06 lower to 0.13 higher)	⊕○○○ VERY LOW	Limited Importance
(CB4) Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members										
5	RCT – 5 QED – 1	serious ^{ah}	not serious	not serious	not serious	none	5122	SMD 0.06 SD (0.0 lower to 0.12 higher)	⊕⊕⊕○ MODERATE	Limited Importance
(CB5) Decreased violence/discrimination at the household level										
4	RCT – 3 QED – 1	very serious ^{ai}	serious ^{aj}	not serious	not serious	none	3438	SMD 0.08 SD lower (0.19 lower to 0.03 higher)	⊕○○○ VERY LOW	Critical
(CB6) Safer and more secure household, communities and areas/territories for women, girls, men, and boys										
4	RCT – 4	serious ^{ak}	not serious	not serious	not serious	publication bias strongly suspected	6563	SMD 0.03 SD lower (0.1 lower to 0.15 higher)	⊕⊕○○ LOW	Critical
(CB7) Reduced frequency and distribution of types of violence by an intimate partner – Physical violence (PV)										
7	RCT – 7	serious ^{al}	serious ^{am}	not serious	not serious	publication bias strongly suspected	6964	SMD 0.03 SD higher (0.03 lower to 0.1 higher)	⊕○○○ VERY LOW	Critical
(CB7) Reduced frequency and distribution of types of violence by an intimate partner – Sexual violence (SV)										
7	RCT – 7	serious ^{an}	not serious	not serious	serious ^{ao}	publication bias strongly suspected	7809	SMD 0.02 SD higher (0.02 lower to 0.07 higher)	⊕○○○ VERY LOW	Critical
(CB8) Improved quality of relationships between women and their household and community members										

Certainty assessment							Sample size	Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Absolute (95% CI)		
7	RCT – 7	serious ^{ap}	serious ^{aq}	not serious	not serious	none	7036	SMD 0.00 SD higher (0.08 lower to 0.00 higher)	⊕⊕○○ LOW	Limited Importance
(CC1) Empowerment/Equality index										
4	RCT – 3 QED – 1	serious ^{ar}	very serious ^{as}	not serious	serious ^{at}	none	7690	SMD 0.06 SD higher (0.07 lower to 0.2 higher)	⊕○○○ VERY LOW	Important, but not critical

Note. CI: Confidence interval; SMD: Standardised mean difference

Explanations

- a. Downgraded once due to uncertainty around bias relating to deviation from intended intervention.
- b. All single studies downgraded once for inconsistency.
- c. Downgraded because of very wide CIs that cross both sides of the threshold.
- d. Downgraded once because one of the three studies is of high risk for bias, and another is unclear.
- e. Downgraded once because point estimates vary considerably, and two of the studies' CIs do not overlap.
- f. Downgraded twice because half of the body of evidence comes from high risk of bias studies with issues related to selection bias and deviations from intended bias.
- g. While some CIs do not overlap, the point estimates are generally consistent and near to the effect estimate of the meta-analysis.
- h. Downgraded to very serious because of concerns related to selection bias and deviation from intended intervention.
- i. Downgraded to very serious because of concerns related to selection bias and deviation from intended intervention.
- j. Downgraded once to reflect the wide range of point estimates for this outcome within this study (0.06 to 0.35).
- k. Downgraded once because there is uncertainty in one or more criteria across seven out of eight studies. That said, only one is rated as high risk of bias. If the high risk of bias study is removed, there is only a marginal change in the effect size.
- l. Downgraded once because there are studies on both sides of the threshold whose CI do not overlap. That said, this larger body of evidence is mostly near the null hypothesis.
- m. Downgraded twice because all three included studies are high risk of bias, including across multiple criteria.
- n. In cognizance of the high risk of bias, this group was downgraded once because point estimates vary considerably and confidence intervals do not overlap for two studies.
- o. Downgraded once because three of five studies present some risk of bias, and one is of high risk of bias. There are issues with deviation from intended intervention found in all studies.
- p. Downgraded once due to uncertainty around bias relating to selection bias.
- r. Downgraded once because three of five studies present some risk of bias, and one is of high risk of bias.
- s. Downgraded once because all but one study report positive effect sizes. The outlier study is of high risk of bias, but does have considerable CI overlap with other studies.
- t. Downgraded to very serious because all but one study presents risk of bias and half of the body of evidence is coming from high risk of bias studies. Selection bias and deviation from intended intervention are of serious concern in three of four studies.
- u. Downgraded to very serious because all studies in this group present of bias, and half of the body of evidence is from high risk of bias studies. Deviation from intended intervention is prevalent.
- v. Downgraded because one study has a point estimate that is very different than the others and has no overlap of CIs.

- w. Downgraded because of wide CI's that cross over the threshold of interest widely on both sides.
- x. Downgraded once because, despite two low risk of bias studies, six of eight present some risk, including two high risk of bias studies which both have significant issues related to deviation from intended intervention.
- y. Downgraded twice because both studies included in this group are of high risk of bias, with particular concerns for deviations from intended interventions.
- z. Downgraded once due to uncertainty around bias relating to deviation from intended intervention.
- aa. Downgraded because of very wide CIs that cross both sides of the threshold.
- ab. Downgraded to very serious because of concerns related to selection bias and deviation from intended intervention.
- ac. Downgraded because of very wide CIs that cross both sides of the threshold.
- ad. Downgraded once because two studies have risk of bias, and one has high risk. Both of the higher risk studies present reason for concern with respect to deviation from intended intervention.
- ae. Downgraded once since although five of the six studies are RCT the QED is high risk and four RCTs raise some concerns related to risk of bias.
- af. Downgraded twice because 1. effects on both side of the thresholds and CI not overlapping 2. $I^2 = 77.89\%$ and Q-test indicate the true outcomes is heterogeneous.
- ag. downgraded once because $p=0.49$ although sample size >400 .
- ah. Downgraded once because two studies present high risk of bias, with particular concerns for deviation from intended intervention.
- ai. Downgraded twice because all studies present some risk of bias and two are of high risk. All studies are either unclear or of high risk of bias for deviations from intended interventions.
- aj. While there is overlap of all studies' CIs, one study has a point estimate that varies considerably from the others.
- ak. Downgraded because two studies at rated as some concern for risk of bias, while one is a high risk of bias with considerable deviation from intended intervention.
- al. Two of four studies in this group are rated as high risk of bias with particular concern for deviation from intended intervention and performance bias.
- am. While all studies are somewhat near to the threshold, several estimates are dissimilar to the others and have large CIs.
- an. Downgraded once because although all studies are RCTs two of them are high risk of bias and two raise some concerns related to risk of bias.
- ao. Downgraded once because although CI is small $p = 0.34$.
- ap. While only one study is of high risk of, all but one have criteria of uncertain risk of bias. There is a particular risk with respect to deviation from intended intervention.
- aq. Downgraded in cognizance of the wide range of point estimates (on both sides of the threshold) and confidence intervals that do not overlap.
- ar. Downgraded once since the sample is a mix of RCT and QED and one RCT and one QED are high risk of bias while the two others are low risk.
- as. Downgraded twice since 1. Studies are on both side of the threshold and the CI are not systematically overlapping 2. $I^2 = 88.44\%$ and q-test highlights the true outcomes appear to be heterogeneous.
- at. Downgraded because although sample size >400 $p = 0.37$.

5. Summary of findings and conclusions

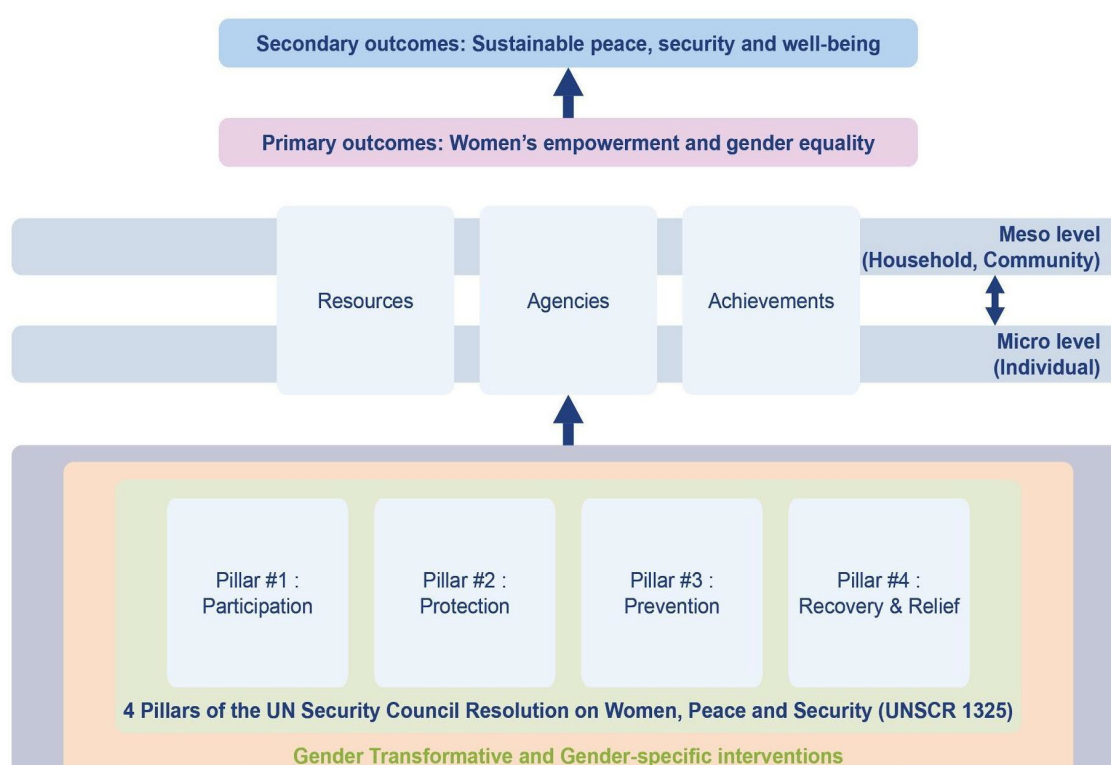
5.1 Summary of findings

5.1.1 Search results

Following the adoption of UNSCR 1325 and the development of the Women, Peace and Security Agenda, there is an increased understanding of the gendered nature of conflict. This focus on women and their empowerment in fragile and conflict affected situations is at the core of the theory of change of our systematic review.

Our ToC is based on the recognition of the importance of the status of women in FCAS: women and girls are targets of war, are particularly vulnerable in fragile contexts and are affected by both fragility and conflict in unique ways (UN, 2000). Women and girls, however, are not always passive, but can instead be powerful agents of change when they are directly involved in peacebuilding processes (Connell, 2002) and as observed in our review with interventions such as reconciliation and intergroup dialogues or community driven development. Thus, as per our ToC and observed in our findings, addressing women's needs and acting towards gender equality contributes to peace and inclusivity (Zuckerman and Greenberg, 2004). Prior evidence suggests that "countries with significant gendered inequality are more likely to prioritize violent masculinities and thus experience increased levels of organised violence" (Klugman et al., p. 3, 2021). The recognition of the link between women's empowerment in FCAS and its contribution to peaceful and inclusive societies has been the driving principle of our review (as per our ToC in Figure 165 below): this contribution is done both through the role of women as agent of peace and the indirect effect of their empowerment on changing social norms towards more peaceful societies.

Figure 165: Systematic Review Theory of Change



We have conducted a large and comprehensive systematic review on the effects of interventions to address women's empowerment and gender equality in FCAS. This review was based on the analysis of interventions under the four pillars of United Nations Security Council Resolution 1325: Participation, Protection, Prevention, Recovery and relief. We identified 104 studies evaluating the effects of 55 programmes and 14 intervention types across three dimensions of women's empowerment, namely resources, agency and achievement. Our results suggest that most interventions have an overall positive effect on beneficiaries across one or more of these empowerment domains, and that interventions tend to have their largest effects on the intervention target, and often smaller or null effects on outcomes further along the causal chain. Table 58 presents an overview of the results of the quantitative synthesis for outcomes rated as critical or important where we had sufficient evidence to conduct a synthesis. Green boxes indicate significant positive effects and yellow boxes indicate non-significant effects. White or empty boxes indicate that there were no studies that met our inclusion criteria, while grey boxes indicate only one study was included, meaning that synthesis was not possible. The full table is presented in Online Appendix Q.

There is a nexus at play: women's empowerment as a driver of peace and women as agents of peace (UN, 2000). Empowerment is a driver of peace as, by providing women with resources and agency, women's empowerment can lead to a change in social structures and norms contributing to more peaceful societies. Women can also be directly involved as peace stakeholders and actors sitting at the negotiation table and acting for peace (UN, 2000). Putting the emphasis on only one of those elements will not enable long term and sustainable change since women need to be empowered to act as agents of peace and this role of agent will contribute to their empowerment in a virtuous cycle. However, our review directly identifies only a few interventions that focus on the roles of women as agents of change, while a vast majority of interventions focus on their empowerment, demonstrating the gap in the evidence base on interventions directly focusing on women as agents of peace.

5.1.2 Overview of effects

Positive effects on targeted primary outcomes

Included interventions often have a positive effect on the primary outcomes they target (see Table 58). For example, in asset transfers (SMD = 0.34, [95% confidence interval (CI): 0.22 to 0.46]) and cash transfers (SMD = 0.22, [95% CI: 0.12 to 0.31]) we observe the largest effects on women's access to and ownership of assets, credit and income, while village savings and loan associations (SMD = 0.24 [95% CI: 0.09 to 0.39]) have their largest effects on women's capacity to use and understand financial, banking and business services effectively. Similarly, for life skills and capacity building programmes we find the largest effects on improved life skills (SMD = 0.16 [95% CI: 0.11 to 0.20]), while inclusive community-driven development has the largest effects on increased representation of women in local and sub-national civil and political processes (SMD = 0.28 [95% CI: 0.13 to 0.43]). Finally, interventions that focus on intra-group dialogue have their largest effects on household and community support for women's economic, social, and human rights (e.g. community dialogues and reconciliations (SMD = 0.18 [95% CI: 0.06 to 0.30]) and discussion groups (SMD = 0.23 [95% CI: 0.05 to 0.39])). We find a paucity of evidence around all-women police stations and community-based services, neither of which have sufficient evidence to conduct a synthesis. We do not

observe negative effects from any intervention on any of the main empowerment outcomes.

Null effects on downstream outcomes

We find that most interventions do not achieve positive and significant effects for 'downstream', behavioural outcomes. For example, despite intimate partner violence (IPV) outcomes being included in seven bodies of evidence and six meta-analyses, the non-significant results of included programmes on IPV outcomes are all very close to zero and even slightly negative in some cases. This finding suggests that the types of interventions included in this review are not affecting IPV, either because the causal chain is too complex and long for single interventions to realistically effect change, or because the programmes were not designed in such a way as to effectively target the outcome. Similarly, cash transfer programmes increase women's access to resources, but by and large they do not have an effect on norms.

Table 58: Quantitative findings

	Outcome	Asset transfers	Cash transfers	Community-based services	Inclusive community-driven development	Institutional provision of loans and savings	Quotas	Self-help groups and VSLAs	TVET	All women police stations	Sensitisation campaigns	Community dialogues and reconciliations	Discussion groups	Safe spaces	LSCB
	AB1 Increased capacity of women to understand and use financial, banking, and business services effectively	Only 1 study ⊕⊕⊕⊕				Only 1 study ⊕⊕○○		0.24 (95% CI: 0.09 to 0.39) ⊕○○○	Only 1 study ⊕⊕⊕○					0.05 (95% CI: -0.03 to 0.13) ⊕⊕○○	0.14 (95% CI: 0.01 to 0.28) ⊕⊕○○
	AB2 Women have increased access and ownership to assets, credit and income	0.34 (95% CI: 0.22 to 0.46) ⊕○○○	0.22 (95% CI: 0.12 to 0.31) ⊕⊕⊕○	Only 1 study ⊕⊕○○	Only 1 study ⊕⊕⊕○	0.21 (95% CI: 0.07 to 0.36) ⊕○○○		0.07 (95% CI: -0.00 to 0.13) ⊕○○○	0.09 (95% CI: 0.03 to 0.15) ⊕○○○		Only 1 study ⊕○○○	Only 1 study ⊕⊕⊕○	0.04 (95% CI: -0.06 to 0.15) ⊕⊕⊕○	0.14 (95% CI: 0.05 to 0.23) ⊕○○○	0.09 (95% CI: 0.02 to 0.16) ⊕⊕○○
	AB3 Women and girls have equitable access to livelihood support services		0.16 (95% CI: 0.10 to 0.22) ⊕⊕⊕⊕					Only 1 study ⊕○○○							
	AC1 More women engaged in other micro, small, and medium-sized enterprises	Only 1 study ⊕○○○	0.87 (95% CI: 0.77 to 0.97) ⊕⊕⊕○		Only 1 study ⊕○○○			Only 1 study ⊕⊕○○	Only 1 study ⊕⊕○○				Only 1 study ⊕⊕⊕○		
	AC2 Women can access decent work (formal and informal employment)	0.10 (95% CI: -0.00 to 0.20) ⊕○○○	0.18 (95% CI: -0.01 to 0.36) ⊕○○○	Only 1 study ⊕⊕○○		Only 1 study ⊕⊕○○		0.13 (95% CI: 0.04 to 0.22) ⊕⊕⊕○	0.07 (95% CI: -0.07 to 0.21) ⊕○○○		Only 1 study ⊕○○○				Only 1 study ⊕○○○
	AC3 Improved capacity of women entrepreneurs	Only 1 study ⊕⊕○○	Only 1 study ⊕⊕○○			0.01 (95% CI: -0.02 to 0.04) ⊕⊕○○		Only 1 study ⊕○○○	0.14 (95% CI: 0.09 to 0.19) ⊕⊕○○					Only 1 study ⊕⊕○○	Only 1 study ⊕○○○
	BA2 Women have more and better control over their bodies and sexual health	Only 1 study ⊕○○○	-0.07 (95% CI: 0.07 to 0.04) ⊕⊕○○		Only 1 study ⊕⊕⊕○			Only 1 study ⊕○○○	0.26 (95% CI: -0.36 to 0.88) ⊕○○○		Only 1 study ⊕⊕⊕○			0.02 (95% CI: -0.07 to 0.10) ⊕○○○	0.06 (95% CI: -0.02 to 0.14) ⊕⊕○○
	BA6 Reduced percentage of women agreeing with certain reasons that justify violence against women and girls		0.11 (95% CI: 0.02 to 0.20) ⊕⊕⊕○		-0.01 (95% CI: -0.10 to 0.08) ⊕⊕○○			-0.23 (95% CI: -0.57 to 0.12) ⊕○○○	-0.03 (95% CI: -0.13 to 0.06) ⊕⊕○○		0.02 (95% CI: -0.04 to 0.09) ⊕⊕○○			Only 1 study ⊕○○○	0.10 (95% CI: 0.01 to 0.18) ⊕⊕○○

	Outcome	Asset transfers	Cash transfers	Community-based services	Inclusive community-driven development	Institutional provision of loans and savings	Quotas	Self-help groups and VSLAs	TVET	All women police stations	Sensitisation campaigns	Community dialogues and reconciliations	Discussion groups	Safe spaces	LSCB
	BA7 Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them		0.03 (95% CI: -0.04 to 0.10) ⊕⊕⊕○		Only 1 study ⊕⊕○○			0.20 (95% CI: 0.07 to 0.32) ⊕○○○	0.04 (95% CI: -0.01 to 0.08) ⊕⊕○○		Only 1 study ⊕⊕○○			0.10 (95% CI: -0.05 to 0.25) ⊕○○○	0.16 (95% CI: 0.11 to 0.20) ⊕⊕○○
	BA8 Reduced support for or occurrence of child and forced marriage	Only 1 study ⊕⊕⊕○	0.07 (95% CI: -0.10 to 0.25) ⊕⊕⊕○				Only 1 study ⊕⊕○○	Only 1 study ⊕○○○	0.17 (95% CI: 0.004 to 0.34) ⊕⊕○○					0.39 (95% CI: -0.08 to 0.86) ⊕○○○	0.10 (95% CI: -0.11 to 0.30) ⊕○○○
	CA6 Effective prevention strategies supported to end violence against women and girls							0.14 (95% CI: 0.04 to 0.24) ⊕⊕○○	Only 1 study ⊕○○○		0.20 (95% CI: 0.02 to 0.41) ⊕⊕○○			Only 1 study ⊕○○○	
	CA8 Increased community support for women's and children's human, economic and legal rights				0.02 (95% CI: -0.09 to 0.13) ⊕○○○			Only 1 study ⊕⊕○○	0.07 (95% CI: -0.04 to 0.18) ⊕○○○					Only 1 study ⊕○○○	Only 1 study ⊕○○○
	CB5 Decreased violence/discrimination at the household level							0.16 (95% CI: 0.02 to 0.29) ⊕○○○	Only 1 study ⊕○○○		Only 1 study ⊕⊕○○				-0.08 (95% CI: -0.19 to 0.03) ⊕○○○
	CB6 Safer and more secure household, communities and areas/territories for women, girls, men, and boys		0.09 (95% CI: -0.04 to 0.22) ⊕○○○		-0.0003 (95% CI: -0.08 to 0.08) ⊕⊕○○			Only 1 study ⊕○○○	Only 1 study ⊕⊕○○					2 studies, no meta-analysis ⊕○○○	-0.03 (95% CI: -0.10 to 0.05) ⊕⊕○○
	CB7 Reduced frequency and distribution of types of violence by an intimate partner	-0.001 (95% CI: -0.07 to 0.07) ⊕⊕⊕○	0.04 (95% CI: -0.06 to 0.13) ⊕○○○						0.09 (95% CI: -0.02 to 0.20) ⊕○○○					Only 1 study ⊕⊕○○	
	CB7 Reduced frequency of physical IPV										-0.06 (95% CI: -0.18 to 0.06) ⊕○○○				0.03 (95% CI: -0.03 to 0.10) ⊕○○○

	Outcome	Asset transfers	Cash transfers	Community-based services	Inclusive community-driven development	Institutional provision of loans and savings	Quotas	Self-help groups and VSLAs	TVET	All women police stations	Sensitisation campaigns	Community dialogues and reconciliations	Discussion groups	Safe spaces	LSCB
	CB7 Reduced frequency of sexual IPV										-0.1 (95% CI: -0.27 to 0.07) ⊕○○○				0.02 (95% CI: -0.02 to 0.07) ⊕○○○
	CC1 Women's empowerment/equality index	0.22 (95% CI: -0.03 to 0.47) ⊕○○○	0.03 (95% CI: -0.10 to 0.15) ⊕○○○			0.74 (95% CI: -0.66 to 0.83) ⊕○○○		0.09 (95% CI: 0.03 to 0.15) ⊕⊕○○	0.13 (95% CI: -0.01 to 0.27) ⊕○○○		0.12 (95% CI: 0.01 to 0.24) ⊕⊕○○		Only 1 study ⊕○○○		0.06 (95% CI: -0.07 to 0.20) ⊕○○○

Note. Each coloured cell presents the effect size and 95% confidence interval as per the meta-analysis of the studies of a given intervention for the corresponding outcome (see legend below). It also presents the GRADE assessment of the certainty of the body of evidence. This table only presents outcomes considered critical or important where we were able to conduct a synthesis. The full table is presented in Online Appendix Q.

Legend
Only one study included, no synthesis
Non-significant result
Positive and significant result
• No included studies – no meta-analysis

⊕○○○: GRADE Very Low level of certainty

⊕⊕○○: GRADE Low level of certainty

⊕⊕⊕○: GRADE Medium level of certainty

⊕⊕⊕⊕: GRADE High level of certainty

5.1.3 Overview of qualitative findings and key considerations

We conducted thematic synthesis for eight intervention groups which includes over 90 linked qualitative studies. The other six groups do not have sufficient studies to conduct a qualitative synthesis. These 90 papers cover 53 unique programmes across our evidence base and provide a rich complement to our quantitative findings by generating 149 descriptive themes and 37 analytical themes. These interventions are funded by USAID, DFID/FCDO, the World Bank Group, the United Nations agencies, the Bill and Melinda Gate Foundation, the European Commission, National Governments and other foundations. They are implemented by organisations such as Oxfam, Care International, The International Initiative for Impact Evaluation (3ie), and various Universities, Academics and multilateral agencies.

Multidimensionality of effects (or lack thereof)

Many interventions have positive effects on primary outcomes, but do not have substantial effects across multiple dimensions of empowerment. Although Kabeer's (1999) framework of women's empowerment suggests that the three dimensions of empowerment--resources, agency, achievement--are inextricably linked and interdependent: resources contribute to agency, contributing to achievement, then contributing to more resources, and so on, we observe that the effects of most interventions do not demonstrate this kind of inter-dimensionality.

That said, there are some intervention types that positively and significantly affect outcomes across multiple dimensions. Cash transfers, self-help groups, and TVET programmes produce beneficial effects across all three dimensions, while asset transfers, sensitisation campaigns, and capacity building programmes show positive and significant effects across two of the three dimensions: resources and agency for asset transfers and capacity building programmes, agency and achievements for sensitisation campaigns.

We also found examples of interventions that, in addition to their effect on immediate outcomes, affect outcomes further along the causal chain, such as women's increased participation in decision making, improved attitudes towards women, and less agreement with reasons justifying sexual and gender-based violence (SGBV). These multidimensional effects are generally smaller in magnitude compared to those related to the main target of the intervention.

The importance of social norms and context

We consistently find that norms and restrictive social contexts act as a barrier to improving women's empowerment in the programmes evaluated. Social norms and contexts related to gender determine the space in which interventions operate and attempt to institute positive changes in women's lives. Within programme design and implementation, restrictive norms and social structures block pathways to change and limit intervention effects. The influence of entrenched norms and societal structures can inhibit the ability of interventions to effect women's empowerment in a multi-dimensional way despite norms' ability to affect targeted, primary outcomes such as enhanced skills and women's representation. This implies that, where the realisation of targeted intervention outcomes depends on the existence of facilitating norms and practices, interventions need to deliberately design programme components that target changing these norms and practices. This challenge is exemplified in our review, which found scant evidence on interventions' ability to change norms, and few interventions having large effects.

A challenging programme design and implementation decision is presented where interventions are taking place in highly restrictive social contexts with existing norms and practices at conflict with the empowerment outcomes targeted by interventions. The available evidence implies that programmes can only reach participants if they work through local agents and structures even though these might not share programme objectives and values.

Intervention design and implementation

A key intervention design consideration is choosing whether to design single sex interventions or to opt for a mixed-gender design approach. While qualitative evidence points to a strong preference by women participants for single-sex programme designs, implementers and decision-makers prefer a mixed-sex programme approach. This was based on the precedence that underlying structural issues and power imbalances cannot be addressed unless men participate in programmes and contribute to joint solutions and activism for change. This observation from a qualitative perspective is, nonetheless, not corroborated by our quantitative data, which show no difference in the size of the effect between single-sex and multi-sex programmes where this was testable (e.g. TVET and capacity building).

Targeting the right beneficiaries is another key consideration for the design of programmes aimed at improving women's empowerment outcomes, especially for cash transfer, self-help groups and village savings and loan associations and safe spaces interventions. Targeting practices can also act as a barrier to interventions' effectiveness when the process lacks transparency.

5.1.4 Intervention-specific findings

The following sections detail the summary of findings for each type of intervention.

Asset transfer

Asset transfer interventions yield significant and positive effects on outcomes in two of the three dimensions of the resources-agency-achievement framework. Although asset transfers interventions target outcomes across all three dimensions of women's empowerment, they only produce positive and significant results on access to assets, credit and income (resources; SMD = 0.34 [95% CI: 0.22 to 0.46], an estimated 13.3% change in the intervention group compared to the control group) and the participation in decision making (agency; SMD = 0.07 [95% CI: 0.04 to 0.11] an estimated 2.8% change in the intervention group compared to the control group). We found no evidence that asset transfer interventions improve achievement-related outcomes, such as improved self-image and confidence, the reduction of intimate partner violence and the improved quality of relationships with community and households. The examination of potential sources of variation did not provide any conclusive findings due to a lack of sufficient data.

Qualitative findings suggest that the provision of assets or tools for cultivation or livestock farming can generally support asset endowment, creditworthiness and income. If assets are utilised productively, it opens pathways to credit use options and income-generating activities. This was especially evident in multi-component asset transfer programmes, such as the Income-Generating Vulnerable Group Development (IGVGD) programme in Bangladesh, which added a microfinance facility which allowed women to use loans for a multitude of purposes (Ahmed et al. 2007). Additionally, agency is bolstered when

women are able to own and manage animals and/or crops. Women can then experience improved decision-making and social status within households and around communities.

Cash transfer

We find that cash transfers have positive effects on outcomes across the all three dimensions of the resources-agency-achievement framework. We observe improvements in access to assets, credit and income (SMD = 0.22 [95% CI: 0.12 to 0.31] an estimated 8.7% change in the intervention group compared to the control group), access to livelihood support services (SMD = 0.16 [95% CI: 0.10 to 0.22] an estimated 6.4% change in the intervention group compared to the control group), engagement in enterprises (SMD = 0.87 [95% CI: 0.77 to 0.97] an estimated 31.1% change in the intervention group compared to the control group), reduction of acceptability of violence (SMD = 0.11 [95% CI: 0.02 to 0.20] an estimated 4.4% change in the intervention group compared to the control group) and increased political representation (SMD = 0.13 [95% CI: 0.05 to 0.21] an estimated 5.2% change in the intervention group compared to the control group). However, we do not find any effects of cash transfer programmes on women's control over their bodies, freedom of movement and association, participation in decision making, better life skills, community support for women's needs, safer communities and reduced frequency of violence.

Many cash transfer programmes included additional components (e.g., a training component), which we hypothesised would be related to larger effects. However, our moderator analyses revealed no differences between the cash transfers alone and cash transfers plus an additional component on any of the women's empowerment outcomes. Similarly, we generally found no difference in the size of effects between conditional and unconditional cash transfers. The exception to this was women's participation in decision making, where we actually observed a significantly *smaller* effect of conditional transfers as compared to unconditional transfers.

Qualitative findings suggest that economic interventions should be tailored to both geographical and societal settings, including entrenched patriarchal norms. This is because gendered household and community roles often restrict women beneficiaries' ability to manage or decide how to spend money. Furthermore, implementation and community leadership are both critical pillars of influence in sensitising communities and households to participate in an intervention.

Community-based services

The effects of community-based services on women's empowerment index outcomes are unknown. The one study that is included under this intervention group only examined outcomes related to access to and ownership of assets, credit and income (SMD = 0.02, 95% CI [-0.05, 0.09] an estimated 0.8% change in the intervention group compared to the control group) and access to decent work (effects ranged from very small, negative effects (SMD = -0.02, 95% CI [-0.09, 0.05] an estimated -0.8% change in the intervention group compared to the control group) to very small, positive effects (SMD = 0.02, 95% CI [-0.05, 0.09] an estimated 0.8% change in the intervention group compared to the control group). None of the effects of this single primary study reached statistical significance. Further, all of the evidence from the community-based services intervention was found to be of low certainty, indicating that future studies of these interventions may produce different results.

Inclusive community-driven development

Inclusive community-driven development (ICDD) only yields significant and beneficial effects in the achievement dimension of the Resources-Agency-Achievement framework. There are only positive and significant results on the increased political representation of women (SMD = 0.28 [95% CI: 0.13 to 0.43] an estimated 11% change in the intervention group compared to the control group). The other outcomes targeted by ICDD programmes, such as freedom of movement and association, positive attitudes towards claiming rights, reduction of justification of violence, support from community on women's needs and the pursuit of safer communities, are not affected by the community-driven development interventions. This echoes existing synthesis that found ICDD programmes use existing social cohesion rather than building it (White et al, 2018).

Based on qualitative evidence, representation within committees in terms of gender and background diversity as well as support and integration with local government structures emerge as key contributors to intervention success. Recognising embedded power dynamics also enables women to better participate in local and sub-national political processes. However, entrenched patriarchal norms necessitate an increased buy-in by these stakeholders such as local cultural or religious leaders for appropriate uptake at the community level. Regular and sustained interaction contributed to maintained uptake and long-lasting effects at the community level.

Institutional provision of loans and savings

We find institutional provision of loans and savings only have positive effects on resources, as measured by access to or ownership of assets, credit and income (SMD = 0.21 [95% CI: 0.07 to 0.36] an estimated 8.3% change in the intervention group compared to the control group), but fail to have significant and positive effects on the increased participation of women in decision-making and the improved capacity of women entrepreneurs, and on women's empowerment index.

Quotas

We are unable to provide strong conclusions regarding quotas due to the low number of included studies using quota interventions. Women's improved attitudes towards taking action to claim their rights was the only outcome with two studies contributing effects, but the synthesis found a null effect. This must be interpreted with caution given the very small number of included studies. Further, all of the evidence from the quota interventions was found to be of low certainty, indicating that future studies of quota interventions may produce different results.

Self-help groups and village savings and loans associations

We find that SHGs and VSLAs have positive effects on outcomes across the resources-agency-achievement framework, including capacity to use and understand financial and banking services (SMD = 0.24 [95% CI: 0.09 to 0.39] an estimated 9.5% change in the intervention group compared to the control group), access to decent work (SMD = 0.13 [95% CI: 0.04 to 0.22] an estimated 5.2% change in the intervention group compared to the control group), increased freedom of movement and association (SMD = 0.18 [95% CI: 0.05 to 0.31] an estimated 7.1% change in the intervention group compared to the control group), rights awareness (SMD = 0.10 [95% CI: 0.02 to 0.18] an estimated 4% change in the intervention group compared to the control group), positive attitudes towards taking action to claim rights (SMD = 0.58 [95% CI: 0.03 to 1.14] an estimated

21.9% change in the intervention group compared to the control group), better life skills (SMD = 0.20 [95% CI: 0.07 to 0.32] an estimated 7.9% change in the intervention group compared to the control group), representation in political processes (SMD = 0.09 [95% CI: 0.04 to 0.14] an estimated 3.6% change in the intervention group compared to the control group), prevention of violence (SMD = 0.14 [95% CI: 0.04 to 0.24] an estimated 5.6% change in the intervention group compared to the control group), decreased violence at the household level (SMD = 0.16 [95% CI: 0.02 to 0.29] an estimated 6.4% change in the intervention group compared to the control group) and women's empowerment index (SMD = 0.09 [95% CI: 0.03 to 0.15] an estimated 3.6% change in the intervention group compared to the control group). This type of intervention, however, fails to affect women's access to and ownership of assets, credit and income, reduction of justification of violence, participation in decision-making, participation in the community, support from the community, more positive attitudes towards women, improved self-image and confidence and better relations with community and household members.

The qualitative evidence base supports this; social capital gains and skill acquisition seems to directly augment economic-focused outcomes rather than association and social-focused outcomes. SHGs and VSLAs are an attractive and convenient option for women who have been systematically excluded from formal economic participation. However, similar to quantitative findings, other factors such as market access and conditions minimise the collective gains for other outcomes.

Technical and vocational education and training

We find that TVET interventions have positive effects on outcomes across the three empowerment dimensions. This includes access to assets, credit and income (SMD = 0.09 [95% CI: 0.03 to 0.15] an estimated 3.6% change in the intervention group compared to the control group), improved capacities of women entrepreneurs (SMD = 0.14 [95% CI: 0.09 to 0.19] an estimated 5.6% change in the intervention group compared to the control group), reduction of forced and child marriage (SMD = 0.17 [95% CI: 0.04 to 0.34] an estimated 6.7% change in the intervention group compared to the control group), political representation of women (SMD = 0.29 [95% CI: 0.03 to 0.55] an estimated 11.4% change in the intervention group compared to the control group) and improved image and self-confidence (SMD = 0.17 [95% CI: 0.01 to 0.33] an estimated 6.7% change in the intervention group compared to the control group). However, TVET fails to have significant and positive effects on women's control over their body, increased freedom of movement and association, reduced justification of violence, improved life skills, participation in decision making, community support, reduction of violence, quality of relationships, and women's empowerment index. Though we hypothesized that TVET would affect women's access to decent work, this was not supported by the evidence, which found a null effect of TVET programmes in fragile and gender unequal states on women's access to work. This is reflected in qualitative data as an effect of social norms since, although we observe a positive effect on capacities, this does not necessarily transform into employment for women. This is often due to conflating social norms that TVET interventions need to consider in varying contexts. Unless programmes are designed to address underlying norms that hinder women's participation in such a programme, there is likely little progress to be made regarding household power relations and women's empowerment within communities. For example, providing childcare and transportation stipends to women participants can

offset the perception that they are abandoning their domestic work and offset the opportunity cost of attendance.

The qualitative evidence base indicates a need to design additional programme components within TVET in order to target outcomes beyond individual skill development and improved attitudes. Group mentorship and peer exchange offer a route to collective action and socialisation that can encourage group uptake of a life skills and capacity building programme. The most successful programmes in terms of uptake and outcomes were ones that enhanced the social aspect as well as the livelihoods aspect.

All-women police stations

We were unable to draw conclusions for all-women police stations due to the low number of included all-women police stations studies. The only included study found a significant negative impact on employment of women in law enforcement (as measured by the daily proportion of cases assigned to women (SMD = -0.45 , 95% CI [-0.58 , -0.32] an estimated 17.4% change in the intervention group compared to the control group), and a null effect on women having improved and equitable access to justice (SMD = 0.07 , 95% CI [-0.06 , 0.19] an estimated 2.8% change in the intervention group compared to the control group). The evidence from all-women police stations was found to be of low or very low certainty, indicating that future studies of all-women police stations may produce different results.

Sensitisation campaigns

Sensitisation campaigns have a beneficial effect on outcomes from two of the three dimensions of the women's empowerment framework. There is little data on resource outcomes from this type of intervention. Beneficial effects of sensitisation campaigns were found on women having more positive attitudes towards claiming their rights (SMD = 0.18 [95% CI: 0.06 to 0.30] an estimated 7.1% change in the intervention group compared to the control group), increased awareness of issues affecting women (SMD = 0.14 [95% CI: 0.02 to 0.25] an estimated 5.6% change in the intervention group compared to the control group) and women's empowerment index (SMD = 0.12 [95% CI: 0.01 to 0.24] an estimated 4.8% change in the intervention group compared to the control group). However, sensitisation campaigns fail to affect reduction of agreement with justification and occurrence of violence, improved self-image, support from households and communities, reduction of physical and sexual IPV and improved quality of relationships with household and community members.

Based on the qualitative findings, sensitisation campaigns can be limited and restricted by social factors such as gender norms and roles. Programmes that, for example, worked with constructive religious and social messaging, often through community leaders in religious institutions, were able to act as a useful channel for disseminating information. Religious leaders and institutions acted as a conduit for participants to better understand gender imbalances and how to address them. Changes that are related to altering human behaviour through learning and communication often take time to realise.

Additionally, qualitative findings reveal that training that specifically targets men and boys is useful in building awareness on IPV and conflict resolution to advocate for men directly challenging gender norms. Training men to better understand women's roles is an important exercise in encouraging men to understand women's reality in the household and community.

Community dialogues and reconciliations

We are unable to make strong conclusions related to community dialogue interventions given that only two studies were included. While the synthesis of the two studies found a positive and significant effect on an achievement-related outcome: improved attitudes and support for women's economic, social and human rights by households and communities, this result should be interpreted with caution. With only two studies, we were also unable to examine any potential sources of heterogeneity. Further, we had a low level of certainty of this body of evidence, suggesting that future studies of community dialogue interventions may produce different results.

Discussion groups

We are unable to make strong conclusions related to discussion group interventions given that only two studies were included. While the synthesis of the two studies found a positive and significant effect on an achievement-related outcome: improved attitudes and support for women's economic, social, and human rights by households and communities, this result should be interpreted with caution as we have low certainty in the body of evidence related to this effect, suggestions that future studies may produce different results. With only two studies, we were also unable to examine any potential sources of heterogeneity.

Safe spaces

Safe spaces only have a beneficial effect on the resource dimension of the empowerment framework through increased access to or ownership of assets, credit and income (SMD = 0.14 [95% CI: 0.05 to 0.23] an estimated 5.6% change in the intervention group compared to the control group). However, safe spaces fail to affect agency and achievement related outcomes in our sample of studies, including increased capacity of women to understand and use financial, banking, and business services effectively, better control over their bodies and sexual health, freedom of movement and association, support for or occurrence of child and forced marriage, participation in decision making by women at the household or community level, improved attitudes, self-image and confidence, improved quality of relationships between women and their household and community members.

Although we would expect it to increase safe communities improve skills and capacity building and reduce violence, there was a null effect of safe spaces on these outcomes in fragile and gender unequal states. Contextual findings from qualitative synthesis show that uptake and participation are buttressed by the maintenance of physical space and planned activities to ensure safety of participants and implementing partners. Including mentorship as a key aspect provides participants and their family members with confidence in dismantling harmful social norms related to the rights of women and girls.

Additionally, targeting women or girls in particular age brackets proves to be successful. For example, a programme that targeted adolescent girls of marriageable age showed success in preventing early and forced marriage in the target community by combating harmful social and religious beliefs that pressured or forced these girls into marriage.

Life, social, and livelihood skills and capacity building

Life, social, and livelihood skills and capacity building (LSCB) interventions result in positive and significant effects on outcomes within two of the three dimensions of the resources-agency-achievement framework. We find beneficial effects of LSCB

programmes on both resources and agency outcomes. Specifically, understanding and use of financial and business services (SMD = 0.14 [95% CI: 0.01 to 0.28] an estimated 5.6% change in the intervention group compared to the control group), better control over their body and sexual health (SMD = 0.06 [95% CI: 0.02 to 0.14] an estimated 2.4% change in the intervention group compared to the control group), freedom of movement and association (SMD = 0.14 [95% CI: 0.02 to 0.26] an estimated 5.6% change in the intervention group compared to the control group), awareness of rights (SMD = 0.06 [95% CI: 0.01 to 0.11] an estimated 2.4% change in the intervention group compared to the control group), reduced agreement with justification of violence (SMD = 0.10 [95% CI: 0.01 to 0.18] an estimated 4% change in the intervention group compared to the control group), equipment with life skills (SMD = 0.16 [95% CI: 0.11 to 0.20] an estimated 6.4% change in the intervention group compared to the control group), participation in decision-making (SMD = 0.03 [95% CI: 0.001 to 0.07] an estimated 1.2% change in the intervention group compared to the control group). However, LSCB interventions fail to effect access to and ownership of assets, credit, and income, reduction of child and forced marriage, participation in and support from the community, reduction of occurrence of violence, quality of women's relationships and women's self-image and confidence.

The qualitative findings suggest that the success of life skills and capacity building programming is often determined by contextual and societal factors that can make or break the success of a programme. This means that embedding programming within these existing norms can enhance a community's acceptance of the offered programme. In contrast, the more embedded a programme is, the more difficult it may be to enact radical change challenging existing norms. Similar to our findings in technical and vocational education and training interventions, group mentorship and peer exchange offer a route to collective action and socialisation that can encourage group uptake of a life skills and capacity building programme.

5.2 Overall completeness of the evidence

The scope of this review is broad, assessing the effects of covering 14 different intervention types on three dimensions of empowerment. This means that while we reviewed a relatively large evidence base, this is generally spread thinly and we identify several important gaps. Specifically, for community-based services (1 included study), quotas (2 included studies), community dialogues and reconciliations (2 included studies), all women police stations (1 included study) and discussion groups (5 included studies) we identify few or no studies reporting on empowerment-related outcomes. Further study of these intervention types is needed to provide any clear conclusions about the nature of their impact on women's empowerment, particularly in areas of conflict with high levels of gender inequality.

Considering the large number of extant interventions and impact evaluations focusing on fragile and conflict-affected states and situations (FCAS), there is a relatively low number of studies of interventions defined as gender-specific or transformative. This is particularly true for interventions primarily focusing on the United Nations Security Council Resolution recovery and relief pillar and on peacebuilding interventions in general. Although there is existing literature on peacebuilding processes, the vast majority of the available evidence is not gender-specific or gender-transformative and does not focus primarily on recovery and relief but more on participation, protection and prevention.

Our review covers evidence from 29 countries, but the geographical distribution of studies is uneven. We identified a large number of studies from Sub-Saharan Africa and Southeast Asia. This is in line with the primary focus of our review on FCAS. However, the low number of studies from the Middle East and North Africa, Latin America and Central Asia indicates that the need for evidence is especially acute for these regions.

Finally, we conducted extensive targeted searches to identify qualitative studies, process evaluations and project documents associated with included experimental and quasi-experimental studies to help us address our review questions. While we identified a number of relevant documents, the volume and quality of evidence limited the extent to which we are able to provide generalisable findings at the intervention level to eight intervention types of the 14 identified. Additionally, six of the 14 intervention groups did not have sufficient qualitative evidence to conduct a synthesis.

5.3 Quality of the evidence

The review included studies that used randomisation or other rigorous quasi-experimental study designs to answer our review questions. About 74 per cent of the included studies used a randomised controlled trial design with randomly allocated treatments to either individuals or clusters and the remaining studies used quasi-experimental designs. Among the quasi-experimental design studies, only 3% had an overall low risk of bias, while 45% had high risk of bias and 52% were assessed as some concerns. The quality of studies using RCT designs was relatively higher, with 28 per cent of the studies assessed as low risk of bias and 28 per cent of studies assessed as high risk of bias.

The results of GRADE ratings of each analysis included in this review demonstrate a varied certainty in the body of evidence. Of the analyses performed, 57% are found to be of very low certainty, 27% of low certainty, 14% of moderate certainty, and 2% of high certainty. While the reasons for low certainty vary across bodies of evidence, the most frequent reasons that certainty of evidence was downgraded are risk of bias of studies and inconsistency of effect sizes of studies included in meta-analyses. These judgments consider the risk of bias of all individual studies included in a body of evidence, and some decisions are made based on metrics from the outputs of meta-analyses themselves. See page 52 for further details on the GRADE approach.

Details on context, intervention design and implementation were sometimes missing, which made it difficult to assess what was delivered, by whom and their effects. Although most of the manuscripts reported sample characteristics, standard errors and standard deviations, some studies were missing details on the exact sample size or precision estimates. We contacted authors in order to collect missing data, but in the absence of responses, effects with missing or unclear data could not be included in the analysis.

In addition, a substantial number of studies met the geographical and thematic selection criteria but failed to meet the experimental and methodological criteria. This demonstrates that, although the evidence base is growing, there is a need for more rigorous methodological approaches using counterfactuals to measure women's empowerment and gender equality outcomes in fragile and conflict affected situations.

5.4 Agreement and disagreement with other studies

The 14 types of interventions analysed in this systematic review have all been the focus of other systematic reviews, although few focused exclusively on FCAS and none had foci on both FCAS and gender specific and transformative interventions. The frequencies of interventions studied in the literature included in this review is consistent with existing reviews: systematic reviews on interventions including cash transfers, TVET, asset transfer, and capacity building are more common, while there is less systematic review evidence on interventions such as all women police stations or community-based services (except on health-related outcomes). Our findings on the effectiveness of interventions on primary outcomes is consistent with other reviews: we observe similarly nuanced results in regards to IPV related outcomes to Spangaro's review (2013) who finds only a minority of studies successfully addressing the incidence of violence, or in the J-PAL's (2021) synthesis of randomised evaluation on managing and preventing crime, violence and conflict observing both promising and negative influence of economic intervention of IPV occurrence and promising results of the livelihood and training based interventions. However, none of those reviews are exclusively focusing on FCAS. Compared to Langer (2018) and De Koker (2014), we often observe a similar magnitude of small to medium positive effects on empowerment related outcomes although we observe that many of the systematic reviews on similar interventions and outcomes did not report effect sizes but only direction of effect. Conversely, we find the magnitude of effects relatively higher in comparison to the recent 3ie systematic review on social cohesion (Sonnenfeld et al., 2021), which assessed some similar interventions, but where the largest significant summary effect was .10 (SMD), while we find effects as large as .87 (SMD).

For all-women police stations (AWPS), we found only two other reviews. While neither conducted a meta-analysis, both reviews found no evidence in favour of AWPS, with Arango and colleagues (2014), specifying that there is a lack of sufficient evidence on AWPS in low- and middle-income countries (L&MICs). This is similar to our review regarding the literature gap that does not allow us to state the effectiveness of this type of intervention.

Two gender-specific or transformative reviews included asset transfers within their analysis, but neither separated asset transfers from other interventions such as cash transfers. In their analysis of combined interventions, del Campo and Steinert (2020) found an overall reduction in physical, sexual and emotional intimate partner violence that we did not observe in either of these intervention groups in our review.

Eight reviews included cash transfers as an intervention, but many of the reviews bundled cash transfers with asset transfers such as livestock. The results presented for cash transfers were mixed. Arango et al. (2014) found that there was not enough evidence in L&MICs to come to a definitive conclusion, whereas McQueston and colleagues (2013) found nuanced evidence on the effectiveness of cash transfers on early marriage and childbearing. Also similar to our findings, Manley and colleagues (2013) found conditional and unconditional cash transfers had similar effects and Baird and colleagues (2013) found that conditional and unconditional cash transfers had similar effects for both boys and girls.

Only one review discussed community dialogues and reconciliations. In this case, this type of intervention was grouped with others, such as sensitisation campaigns. The interventions as a whole were found to have significant and positive effects on sexual intimate partner violence (del Campo and Steinert 2020). While our included studies did not provide a measure of IPV, this is in line with our observations on improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members. However, we see that only a few studies included in the other review are gender- specific or transformative.

Community-based services were examined in four reviews, which mainly focused on health outcomes. Both studies, which presented health outcomes, found community-based services to have a significant and positive effect (Lassi et al. 2015; Malarcher et al. 2009), and those focused on women's empowerment outcomes found the same results (Arango et al. 2014; del Campo and Steinert 2014). Del Campo and Steinert (2014) grouped community-based services with asset transfers. This observation aligns with our findings and the low number of studies included under this category due to the exclusion of health from our scope.

Two reviews found evidence in favour of the effects of discussion groups, yet neither had a focus on low- and middle-income countries (L&MICs). One review included one discussion group-based intervention in an L&MICs and found a reduction in violence perpetrated by men (De Koker et al. 2014). This is in line with our findings on improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members.

Five reviews looked at the institutional provision of loans and savings (with Vaessen and colleagues (2014) conducting a meta-analysis) but found no consistent evidence for an effect of microcredit on women's control over household finances. This is in line with our findings that institutional provision of loans and credit improve both the capacities of women to use finance-related services and their access to credit, but do not increase their participation in decision-making at the household and community level.

Seven reviews assessed interventions focused on life, social and livelihood skills and capacity building (LSCB). One review, which utilised the GRADE methodology, found that life, social and livelihood skills and capacity building interventions led to an increase in women's formal wage employment, income and economic empowerment (Langer et al. 2018). Dickson and colleagues (2012) also found modest but significant effects in terms of the income generating ability of women and girls after being supported through these interventions. This is in line with our observations that LSCB led to increased capacities in the use of finance services.

Four reviews analysed safe spaces interventions with the main takeaway being that there is a lack of rigorous evidence around this intervention type in L&MICs.

Two reviews looked at self-help groups (SHGs) and village savings and loan associations (VSLAs). Brody and colleagues (2016) found an effect size through a meta-analysis, with standardised mean differences ranging from 0.06-0.41 in favour of self-help group participant's economic and political empowerment. We found that SHGs and VSLAs have, among our list of interventions, a positive impact across the three dimensions of women's empowerment.

Six reviews looked at sensitisation campaigns and found very mixed results. LaCroix and colleagues (2014) found positive and significant effects of HIV sensitisation campaigns on women's HIV knowledge and condom use. Other studies, such as Arongo and colleagues (2014) found much more mixed results with a lack of sufficient evidence being a major issue. We found the same mixed results in our review with a positive and significant effect on improved awareness and responsiveness to the demands, claims, rights and inputs of women but no significant effects on other sets of outcomes.

Finally, four reviews analysed technical and vocational education and training (TVET) interventions. Langer and colleagues (2018) found the interventions led to an increase in women's income and empowerment. However, our effect sizes from FCAS are considerably lower on income, employment and empowerment compared to similar outcomes in L&MICs in general in Chinen's (2017) review. Del Campo and Steinert (2020) also found a positive effect of the interventions on women's experience of physical and emotional intimate partner violence. Our review shows that TVET interventions mainly affect the resources-related outcomes of women's empowerment.

5.5 Discussion and conclusions

We have conducted a large and comprehensive systematic review on women's empowerment and gender equality in fragile and conflict-affected states and situations. We reviewed 104 studies evaluating the effects of interventions across 44 different outcomes. Our results suggest that most interventions have an overall positive effect on beneficiaries as compared to those not receiving these interventions. As expected, depending on which outcome we look at, different interventions produce the largest effects. Below we discuss some patterns, crosscutting observations and discussion points from our sample of included studies.

Our review finds multiple examples of interventions affecting change on primary outcomes that are closely related to the mechanism of the intervention. For example, asset transfers, cash transfers and the institutional provision of loans and savings mainly have positive and significant positive effects on economic and livelihood related outcomes. Similarly, inclusive community-driven development and community dialogues mainly have positive and significant effects on the improvement of systems. Among our 14 interventions, SHGs and VSLAs, TVET and cash transfers are the only interventions having effects across all three dimensions of women's empowerment and gender equality.

Although theory dictates that the three dimensions of empowerment are inextricably linked, we observe that the effects of most interventions are generally limited to only one dimension. In line with this, we find a relative lack of multi-component interventions that specifically target the three dimensions at the same time. Although the data of our 14 interventions does not allow us to posit on the effect of multi-component interventions, the importance of multidimensional empowerment in the theoretical literature suggests that more research on multicomponent interventions would be of interest.

As presented in Table 59, our panel of interventions have beneficial effects on different outcomes related to their categories and dimensions within the empowerment framework. These findings can be aggregated at the UNSCR 1325 Pillar level:

- Interventions classified in the **participation pillar** yield positive and significant effects across outcomes in all three dimensions of the Resources-Agency-Achievement framework. That said, only three types of interventions exhibit a positive effect across the empowerment dimensions (i.e. cash transfer, SHGs and VSLAs and TVET) and a majority of the interventions under this pillar focus on resources and agency.
- Our sample of interventions under the **prevention and protection pillars** are not large enough to make general conclusions. However, there seems to be a stronger focus on the achievement dimension of the empowerment framework.
- **Multi-pillar interventions** (i.e. discussion groups, life, social, and livelihood skills and capacity building and safe spaces) yield positive and significant effects on outcomes across a number of categories: economic and livelihood related resources, individual agency, community level agency, norms and behaviour change.
- Included interventions often yield positive and significant results on their targeted, primary outcomes, but are less effective at impacting change on more complex downstream outcomes like intimate partner violence (see Table 58). Table 59 also illustrates the absence of an intervention that will singlehandedly address all aspects of gender equality and women's empowerment.

Table 59: Overview of interventions effectiveness

<i>Empowerment Dimensions</i>	Resources			Agency			Achievement		
<i>Outcome categories</i>	Access to justice and legal services	Economic and livelihood related resources	Access to employment	Individual agency	Community level agency	Institutions supporting agency	Improved systems	Norms and behaviour change	Index
Participation		5	3	3	1		4	2	1
Asset transfer									
Cash transfer									
Community-based services									
Inclusive community-driven development									
Institutional provision of loans and savings									
Quotas									
Self-help groups and Village savings associations									
Technical and vocational education and training									
Prevention								1	
Community dialogues and reconciliations									
Protection				1				1	1
All women police stations									
Sensitisation campaigns									
Multi-pillar (participation and protection)		2		1	1			1	
Discussion groups									
Life, social, and livelihood skills and capacity building									
Safe spaces									
General	0	7	3	5	2	0	4	5	2

Legend

	Aggregation of effect of all the interventions and number of interventions with a significantly positive effect on an outcome from this category at intervention level
	Aggregation of effect at the UNSCR 1325 level and number of interventions with significant and positive effects on an outcome from this category at intervention level
	Significant and positive effect observed on an outcome from this category at intervention level

Implications for decision-making and practice

To achieve transformational effects on women's empowerment outcomes, interventions need to be carefully designed to address key barriers and dimensions of women's empowerment.

- There is a need to create and implement explicit programme components that target empowerment outcomes, if such empowerment outcomes are desired.
- For interventions to have effects on multiple dimensions of empowerment or on outcomes that are difficult to shift, they need to be designed to address the different causal pathways of each intervention type to better target the outcomes where they have a significant effect.
- Programmes that yield impacts beyond the immediate, primary outcomes that they targeted reflect a clear understanding of the barriers to empowerment within their respective contexts and tailor the design of components accordingly, including through critical gender analysis.
- Applying a gendered implementation plan may contribute to programme effectiveness. Implementation plans require careful application of a gender lens to assess whether implementation choices are gender-sensitive, from the choice of venue, to the outreach campaign, and personnel, to name a few.
- The potential of women peer-to-peer exchanges and collective-building programme components such as mentoring and women's groups are important facilitators to be considered. These types of programme components are consistently linked to the creation of social capital and networks among women, which in return provides a step towards empowerment related outcomes.

Programme design needs to be tailored to the local context and adjusted to work in restrictive contexts

- Embedding programmes in existing structures contributes to enhanced legitimacy and access to participants, facilitating programme components that target improvements in women's livelihoods and economic situations. Programmes need to balance the risks of co-option with the benefits of gaining access and the opportunity cost of women not receiving livelihood support activities on a case-by-case basis.
- A specific design consideration in these contexts is again the opportunity to apply programme components aimed at women peer-to-peer exchanges and collective-building, which hold potential to create new structures and support systems as a platform for women's empowerment.
- Open dialogue forums may support buy-in by stakeholders in areas where entrenched patriarchal norms necessitate increased buy-in by local cultural or religious leaders for appropriate uptake at the community level. Further research is necessary in different contexts and at the community level.

Targeting of the right beneficiaries is important to maximise impact

- Appropriate targeting requires understanding of who are the right groups of beneficiaries to include in programmes. Baseline assessments and contextual analysis are key, including assessments of whether male targeting is advisable (for example, the right target group for addressing norm change may, in some instances, be exclusively male)

- Targeting procedures should be tailored to minimise beneficiary selection errors and maximise trust in administrative processes, programme participation, and the achievement of aimed outcomes.
- Conflict and accusation of favouritism can be addressed proactively if targeting criteria are co-designed with stakeholders and made transparent and explicit throughout.
- The most vulnerable households and women should be the main beneficiary targets, as they are not only able to realise the most gains related to meeting their basic needs, but also benefit most from access to credit and capital to engage in economic activities. The targeting process needs to be inclusive geographically, ensuring that vulnerable groups in remote locations are selected for these programmes.
- Age is another factor that needs to be considered for targeting. For instance, safe space programmes should target girls before they reach adolescence where they are more at risk of social pressures that may drive them into forced marriages and unwanted pregnancy.

- *Prioritisation of support to women in peace processes and recovery and relief context*

Despite the broad reach of our review, many studies that exclusively focused on outcomes related to peacebuilding in fragile contexts did not have a gender component inclusive to our scope. Interventions focusing on peacebuilding need to be tailored to integrate a gender specific or transformative lens.

- *Prioritisation of gender-transformative and gender-specific interventions*
Interventions and monitoring of results should go further than disaggregating data by gender to specifically targeting gender components and outcomes. This includes tailoring interventions with an equity focus, such as aligning cash disbursement schedules with women's time use.

- *Prioritisation of measuring peace and gender-related outcomes*
The lack of evidence in this review on the recovery and relief pillar represents a need for improved resilience evaluations that are simultaneously gender-sensitive or transformative.

There are a number of intervention specific implications for policy and practice and these are provided in the Tables below.

Table 60: Participation interventions implications

Intervention	Intervention works on	Intervention does not work on
Asset transfers	<p>Resources</p> <ul style="list-style-type: none"> Women have increased access to and ownership of assets, credit and income. <p>Agency</p> <ul style="list-style-type: none"> Increased participation in decision making by women at the household or community level <p>Achievement</p> <ul style="list-style-type: none"> Effect unknown - insufficient evidence 	<p>Resources</p> <ul style="list-style-type: none"> Women can access decent work (formal and informal employment) <p>Agency</p> <ul style="list-style-type: none"> Effect unknown - insufficient evidence <p>Achievement</p> <ul style="list-style-type: none"> Women have improved attitudes, self-image and confidence Reduced frequency and distribution of types of violence by an intimate partner Improved quality of relationships between women and their household and community members Women's empowerment/equality index
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> Identify and secure multiple convenient locations for training activities (e.g., agricultural education workshops) and for asset collection to improve participation and productive use of assets. Design quality (participatory and 'hands-on'), comprehensive and inclusive training for recipients to promote understanding and successful adoption of relevant skills to carry out activities related to asset transfer such as agricultural production. Conduct quality checks for asset transfers and provide necessary supporting inputs (e.g., fences and irrigation equipment). Ensure that adequate household level social and labour support systems for women beneficiaries are available. For instance, through sensitisation or training of family members, particularly husbands. Explore multicomponent interventions that include legal support, microcredit, cash transfers, and capacity building more holistically to support the productive and sustained use of assets 	
	<p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> Traditions and norms, such as gendered divisions of labour and women's rights to land hinder the type of activities women can participate in and eventual empowerment outcomes Programmes have attempted to address this through gender sensitisation campaigns targeted at partners. Programme activities (including compulsory training) may disrupt usual women's day-to-day work obligations and affect time use preferences 	

	<p>Good practice examples:</p> <ul style="list-style-type: none"> ● Burkina Faso's Enhanced-Homestead Food Production (E-HFP) emphasises the need for additional input support to complement asset transfers meant for agricultural purposes. Specifically, whilst there was an indication of improved ability for women to use land for agricultural purposes, primary obstacles included insufficient water access, pest and animal attacks. Hence, the provision of additional inputs in the form of irrigation equipment, fences and pest control chemicals could have mitigated these challenges and enhanced the benefits realised from the assets transferred. ● Bangladesh's Targeting the Ultra Poor (TUP) illustrates the importance of ensuring adequate social and labour support systems for women beneficiaries. The livestock and poultry that women received raised women's workloads at home and in some cases prevented women from being employed outside their homes. As an alleviating factor, the women benefited from children and husband assistance with selected tasks related to the care of their livestock. ● The Bangladesh's Income-Generating Vulnerable Group Development (IGVGD) asset transfer programme included a microfinance facility which enhanced economic involvement as it allowed women to use the borrowed money for various purposes including paying for rental on land they used to start and carry out agricultural activities related to the agricultural assets they had received. 	
Intervention	Intervention works on	Intervention does not work on
Cash transfers	<p>Resources</p> <ul style="list-style-type: none"> ● Women have increased access to and ownership of assets, credit and income. ● Women and girls have equitable access to livelihood support services. ● More women engaged in other micro, small, and medium-sized enterprises. <p>Agency</p> <ul style="list-style-type: none"> ● Reduced percentage of women agreeing with certain reasons that justify violence against women and girls <p>Achievement</p> <ul style="list-style-type: none"> ● Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post-conflict restoration 	<p>Resources</p> <ul style="list-style-type: none"> ● Women can access decent work (formal and informal employment) <p>Agency</p> <ul style="list-style-type: none"> ● Women have more and better control over their bodies and sexual health ● Women have increased freedom of movement and association ● Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them ● Reduced support for or occurrence of child and forced marriage ● Increased participation in decision making by women at the household or community level ● Women participate more in their community <p>Achievement</p> <ul style="list-style-type: none"> ● Women have improved attitudes, self-image and confidence ● Improved attitudes and increased support for women's economic,

		<p>social and human rights by men, household and family members and community members</p> <ul style="list-style-type: none"> • Safer and more secure household, communities and areas/territories for women, girls, men, and boys • Reduced frequency and distribution of types of violence by an intimate partner • Improved quality of relationships between women and their household and community members • Women's empowerment/equality index
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> • Incorporating couples' relationships, social and livelihood skills, mentoring and capacity building components can support household relationships and conflict resolution skills • Targeting the most vulnerable households and women who struggle to meet their basic needs and those who do not have access to credit or capital to start businesses • Targeting family units (compared to individuals) and carrying out clear awareness-raising campaigns supports participation • When determining transfer size, considerations should be made for inflation and household size, but relatively high cash transfers may discourage participation in wage labour activities. • Mentorship, training and support from family members promotes adherence and compliance to programme requirements, especially in conditional cash transfer programmes. • We find no difference in the effects of conditional transfers compared to unconditional transfers. This should be carefully considered given the higher administrative costs of conditional transfers. 	
	<p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> • Social norms, existing gender roles and rigid restrictive beliefs can preserve men's dominance (regardless of women's income contribution) in household decision making, including participation in these programmes. 	
	<p>Good practice example:</p> <ul style="list-style-type: none"> • After noticing slow uptake and lack of awareness in communities, the implementation team from the Zomba Cash Transfer Programme (ZCTP) in Malawi adopted a sensitisation campaign to help prospective participants understand the objectives of the intervention, which improved participation rates. • In the Takaful and Karama Programme in Egypt, severely poor people used a large part of the cash transfer to take care of their basic needs including food, clothing and medical expenses following the receipt of the cash transfer. Notably, a 	

	<p>significant number of the beneficiaries were also highly in debt. Hence, the cash transfer alleviated short term needs for vulnerable beneficiaries in order to focus on more sustainable, long-term economic goals.</p> <ul style="list-style-type: none"> • The mentorship and coaching components in Kenya's Rural Entrepreneur Access Project (REAP) implemented by the BOMA Project helped ensure that participants adhered to the programme, enhanced school attendance, financial literacy and beneficiaries' knowledge. 	
Intervention	Intervention works on	Intervention does not work on
Community-based services	<p>Resources</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Agency</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Achievement</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence 	<p>Resources</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Agency</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Achievement</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
	<p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
	<p>Good practice example:</p> <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
Intervention	Intervention works on	Intervention does not work on
Inclusive community-driven development	<p>Resources</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Agency</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Achievement</p> <ul style="list-style-type: none"> • Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post-conflict restoration 	<p>Resources</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Agency</p> <ul style="list-style-type: none"> • Women have increased freedom of movement and association • Women have more positive attitude towards taking action to claim their rights • Reduced percentage of women agreeing with certain reasons that justify violence against women and girls <p>Achievement</p>

		<ul style="list-style-type: none"> • Increased attention and focus on the needs and priorities of women and girls, and other vulnerable groups during relief and recovery in conflict and post-conflict settings • Increased community support for women's and children's human, economic and legal rights • Women have improved attitudes, self-image and confidence • Safer and more secure household, communities and areas/territories for women, girls, men, and boys
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> • Composition of committees and representatives in decision-makers structures mattered largely for the legitimacies of these bodies within the communities targeted by the intervention (e.g., mandated parity) • Include capacity building initiatives for established governance bodies • Encourage collaborations between institutions to mitigate potential conflicts between multiple governance systems or bodies through regular meetings and combined working groups and balance politics and power relations between existing and new bodies. 	
	<p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> • Governance structures introduced may coexist with traditional governance authorities (e.g., local chiefs) as well as local government structures. Navigating this co-existence requires a careful balancing of politics and power relations, which limited the space for project implementation and reach in some instances. • Powerful elites with a vested interest may undermine project implementation when the legitimacy and capacity of governance bodies is a challenge. 	
	<p>Good practice example:</p> <ul style="list-style-type: none"> • The Democratic Republic of the Congo's Tuungane programme featured mandated gender parity in the composition of project committees, which tended to facilitate more women to participate in the programme. • In Afghanistan's National Solidarity Programme (NSP), collaboration between entities or institutions such as cultural or religious leaders, public and private agencies was crucial. Bottom-up initiatives inclusive of diverse groups of community members helped legitimise the intervention's committees as local institutions. • Additionally, in Afghanistan, project success hinged on timely disbursement of funds, regular meetings and follow-up on community initiatives. The timing of project implementation centred around it not coinciding with annual harvesting seasons, and the time use of community members, particularly women. 	

Intervention	Intervention works on	Intervention does not work on
Institutional provision of loans and savings	Resources <ul style="list-style-type: none"> Women have increased access to and ownership of assets, credit and income. Agency <ul style="list-style-type: none"> Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> Empowerment index 	Resources <ul style="list-style-type: none"> Improved capacity of women entrepreneurs Agency <ul style="list-style-type: none"> Increased participation in decision making by women at the household or community level Achievement <ul style="list-style-type: none"> Women's empowerment/equality index
	Intervention design suggestions to maximise benefits <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	Contextual considerations to keep in mind <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	Good practice example: <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
Intervention	Intervention works on	Intervention does not work on
Quotas	Resources <ul style="list-style-type: none"> Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> Effect unknown - insufficient evidence 	Resources <ul style="list-style-type: none"> Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> Women have more positive attitude towards taking action to claim their rights Achievement <ul style="list-style-type: none"> Effect unknown - insufficient evidence
	Intervention design suggestions to maximise benefits <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	Contextual considerations to keep in mind <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	Good practice example: <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	

Intervention	Intervention works on	Intervention does not work on
Self-help groups and village savings and loans associations	<p>Resources</p> <ul style="list-style-type: none"> Increased capacity of women to understand and use financial, banking, and business services effectively Women have increased access to and ownership of assets, credit and income Women can access decent work (formal and informal employment). <p>Agency</p> <ul style="list-style-type: none"> Women are more aware of their rights and the roles and responsibilities of duty bearers. Women have more positive attitude towards taking action to claim their rights. Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them. <p>Achievement</p> <ul style="list-style-type: none"> Effective prevention strategies supported to end violence against women and girls Decreased violence/discrimination at the household level Women's empowerment/equality index 	<p>Resources</p> <ul style="list-style-type: none"> Women have increased access to and ownership to assets, credit and income <p>Agency</p> <ul style="list-style-type: none"> Reduced percentage of women agreeing with certain reasons that justify violence against women and girls Increased participation in decision making by women at the household or community level Women participate more in their community <p>Achievement</p> <ul style="list-style-type: none"> Communities have a more positive attitude towards women/marginalised groups Women have improved attitudes, self-image and confidence Improved quality of relationships between women and their household and community members
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> Include training components (e.g., financial literacy and business skills) for participants and group leaders in self-help groups and village savings and loan associations programmes to enhance confidence in leadership by participants. To enhance participation, encourage existing groups to use traditional knowledge sharing platforms such as listening in meetings by prospective participants and motivational workshops. In programme design, ensure that it is mandatory for groups to establish clear rules of participation and communication by groups to promote adherence and willingness to participate. 	

	<p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> • Group compliance obligations and domination by elites may foster intergroup conflict, whereas restrictive legislative frameworks may stifle programme expansion, which could hinder programme impact. • Women may not receive support from their partners due to pervasive gender norms and roles. • Regularity of meetings ensures reliable participation and group success. • Meeting spaces should be in centralised locations to reduce transport burden on participants. <p>Good practice example:</p> <ul style="list-style-type: none"> • During the Saving for Change (SfC) programme in Mali, some participants accepted the programme because replicating agents were considered legitimate and this legitimacy was derived from a formal training they underwent and the certificate they received. Trainers gained status and recognition by both clients and the broader community owing to this training. • In the same programme, group support was also seen to promote participation by trust building in the system of self-help groups and village savings and loan associations first. For instance, there were groups that refused to join the association due to lock-in costs. To garner interest and encourage participation, an agent brought women from different groups to visit and observe association meetings and report back to their respective groups. 	
Intervention	Intervention works on	Intervention does not work on
Technical and vocational education and training	<p>Resources</p> <ul style="list-style-type: none"> • Women have increased access to and ownership of assets, credit and income • Improved capacity of women entrepreneurs <p>Agency</p> <ul style="list-style-type: none"> • Reduced support for or occurrence of child and forced marriage <p>Achievement</p> <ul style="list-style-type: none"> • Increased representation of women in local and sub-national civil and political processes, including during peacebuilding and post conflict restoration • Women have improved attitudes, self-image and confidence. 	<p>Resources</p> <ul style="list-style-type: none"> • Women can access decent work (formal and informal employment) <p>Agency</p> <ul style="list-style-type: none"> • Women have more and better control over their bodies and sexual health • Women have increased freedom of movement and association • Reduced percentage of women agreeing with certain reasons that justify violence against women and girls • Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them • Increased participation in decision making by women at the household or community level <p>Achievement</p> <ul style="list-style-type: none"> • Increased community support for women's and children's human, economic and legal rights

		<ul style="list-style-type: none"> • Reduced frequency and distribution of types of violence by an intimate partner • Improved quality of relationships between women and their household and community members • Women's empowerment/equality index
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> • Design and implement a tailored and sensitive awareness-raising and outreach campaign to attract participants into the programme. • Identify and select a central and safe location that is easily accessible (e.g., via public transport, within walking distance of participants). • Provide allowances for transport or childcare to encourage participation and reduce barriers of access. • Target single sex participants (women) to allow space for open engagement and enhance feelings of safety. • An integration with Life, social, and livelihood skills and capacity building (LSCB) capacity building programmes are promising as assessments of stand-alone technical and vocational education and training (TVET) programmes and TVET components within LSCB programmes emphasise the interconnectedness of both types of skill sets. • A multicomponent or multi-faceted design approach to support beyond the skill building component including mentoring, peer-to-peer exchange and 'collective-building' formats such as women's self-help groups and clubs. 	
	<p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> • Training curricula and materials require careful adaptation to local context and translation into local languages. • The skilfulness and approach of programme trainers is a key implementation element and driver of programme satisfaction. • Tailoring programming to context is crucial, particularly in contexts where women or girls are barred from certain occupations. 	
	<p>Good practice example:</p> <ul style="list-style-type: none"> • The Gender and Entrepreneurship Together (GET) Ahead for Women in Enterprise programme in Kenya was designed specifically with a gender lens throughout and aimed to address both the practical and strategic needs of low-income women in enterprise by strengthening their basic business and people management skills as well as supporting networking and relationship building. Women beneficiaries, in addition to the freedom to express themselves without the censure of men, valued the 'camaraderie', support and mentorship they received from one another. Mentoring was seen to enhance business skill transfer and development by contextualising skills in the experience of an established entrepreneur. • The Empowerment and Livelihood for Adolescents (ELA) programme in Sierra Leone referred to its multifaceted approach to simultaneously tackle multiple disadvantages young women face, related to having agency over their bodies and barriers to 	

	<p>accumulating human capital. This approach combines the provision of life skills, vocational skills and microfinance, delivered within a social club and safe space for women.</p> <ul style="list-style-type: none"> • Liberia's Economic Empowerment of Adolescent Girls and Young Women (EPAG) programme deliberately designed a pairing, small group (3-4 girls) approach between girls with the aim of improving girls' skill development as well as to promote the creation of social capital, such as friendships, mutual trust and support networks among the girls. The girl participants appreciated this design element and reported that they helped each other during and outside of the classroom activities.
--	--

Table 61: Prevention interventions implications

Intervention	Intervention works on	Intervention does not work on
Community dialogues and reconciliations	<p>Resources</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Agency</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Achievement</p> <ul style="list-style-type: none"> • Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members 	<p>Resources</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Agency</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Achievement</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
	<p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
	<p>Good practice example:</p> <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	

Table 62: Protection interventions implications

Intervention	Intervention works on	Intervention does not work on
All-women police stations	Resources <ul style="list-style-type: none"> Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> Effect unknown - insufficient evidence 	Resources <ul style="list-style-type: none"> Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> Effect unknown - insufficient evidence
	Intervention design suggestions to maximise benefits <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	Contextual considerations to keep in mind <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
	Good practice example: <ul style="list-style-type: none"> Not enough studies for qualitative synthesis 	
Intervention	Intervention works on	Intervention does not work on
Sensitisation campaigns	Resources <ul style="list-style-type: none"> Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> Women have more positive attitude towards taking action to claim their rights. Achievement <ul style="list-style-type: none"> Increased awareness in communities of the issues affecting women Women's empowerment/equality index 	Resources <ul style="list-style-type: none"> Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> Reduced percentage of women agreeing with certain reasons that justify violence against women and girls Achievement <ul style="list-style-type: none"> Effective prevention strategies supported to end violence against women and girls Women have improved attitudes, self-image and confidence Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members Reduced frequency of physical IPV Reduced frequency of sexual IPV

		<ul style="list-style-type: none"> Improved quality of relationships between women and their household and community members
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> Design more frequent programme activities such as training and counselling over an extended period of time. Adopt in-person platforms and intimate sessions that largely focus on couples, and men as perpetrators. Engage and utilise respected opinion leaders (e.g., religious leaders) to counsel, mentor and sensitise couples to IPV with focus on the tenets of healthy relationships and conflict resolution skills and knowledge. Tailor the types of community engagement to context, such as parent's evenings, dialogue groups and exchanges. Draw on religious messages to support change regarding gender equality. As seen in other intervention groups, sensitisation campaigns can be a key "add-on" to enhance the uptake and implementation of an intervention. As a result, sensitisation may be an "easy win" to increase the opportunity of trickle-down effects across programmes. 	
	<p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> Gender norms, roles and culturally developed patriarchal attitudes are often the most entrenched and difficult barrier in successful sensitisation but engaging men, particularly local opinion leaders, may counter traditional gender roles. Issues of personal image, knowledge on rights and empowerment can be addressed gradually as they depend on the nature of close relationships and personal confidence from successful programme impact. Using communal and public platforms can work to reach more participants at the same time compared to intimate settings. 	
	<p>Good practice example:</p> <ul style="list-style-type: none"> Training men to understand the value of women's roles better was an important exercise for Men's Action for Stopping Violence Against Women (MASVAW) participants in India. Particularly, activities that allowed men to imagine "the average day in a woman's life" were instrumental in provoking reflection around gender roles and norms. Opinion leaders incorporated in India's MASVAW programme expressed excitement to engage actively with other participants and supported programme staff in identifying and embedding knowledge on gender inequalities in their communities. Their involvement encouraged them to prioritise intimate partner violence prevention and wider engagement on gender issues with the community. In Rwanda, the Indashyikirwa programme's personalised settings were more conducive to engagement and listening through participatory methods of community engagement. These included community forums, to encourage openness, particularly regarding violence against women and gendered issues. Participatory training and community-based activism in the Indashyikirwa programme encouraged a culture of community engagement to activate learning and third-party intervention. In this way, participants were able to protect those that suffered from intimate partner violence. 	

	<ul style="list-style-type: none"> • A religious opinion leader in Rwanda Indashyikirwa's programme noted the importance of drawing on religious messages and principles to support change regarding gender equality as it allowed participants to understand and internalise that the genders were all equal.
--	---

Table 63: Protection interventions implications

Intervention	Intervention works on	Intervention does not work on
Discussion groups	Resources <ul style="list-style-type: none"> • Effect unknown - insufficient evidence Agency <ul style="list-style-type: none"> • Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> • Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members 	Resources <ul style="list-style-type: none"> • Women have increased access to and ownership to assets, credit and income Agency <ul style="list-style-type: none"> • Effect unknown - insufficient evidence Achievement <ul style="list-style-type: none"> • Effect unknown - insufficient evidence
	Intervention design suggestions to maximise benefits <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
	Contextual considerations to keep in mind <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
	Good practice example: <ul style="list-style-type: none"> • Not enough studies for qualitative synthesis 	
Intervention	Intervention works on	Intervention does not work on
Life, social, and livelihood skills and capacity building	Resources <ul style="list-style-type: none"> • Increased capacity of women to understand and use financial, banking, and business services effectively Agency <ul style="list-style-type: none"> • Women have increased freedom of movement and association. • Women are more aware of their rights and the roles and responsibilities of duty bearers. 	Resources <ul style="list-style-type: none"> • Women have increased access to and ownership to assets, credit and income Agency <ul style="list-style-type: none"> • Reduced support for or occurrence of child and forced marriage Women participate more in their community Achievement <ul style="list-style-type: none"> • Communities have a more positive attitude towards

	<ul style="list-style-type: none"> • Reduced percentage of women agreeing with certain reasons that justify violence against women and girls • Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them. • Increased participation in decision making by women at the household or community level <p>Achievement</p> <ul style="list-style-type: none"> • No significant effect identified in the list of studies 	<p>women/marginalised groups</p> <ul style="list-style-type: none"> • Women have improved attitudes, self-image and confidence • Improved attitudes and increased support for women's economic, social and human rights by men, household and family members and community members • Decreased violence/discrimination at the household level • Safer and more secure household, communities and areas/territories for women, girls, men, and boys • Reduced frequency of physical IPV • Reduced frequency of sexual IPV • Improved quality of relationships between women and their household and community members • Women's empowerment/equality index
	<p>Intervention design suggestions to maximise benefits</p> <ul style="list-style-type: none"> • Design and implementation of a sensitive and targeted awareness-raising and outreach campaign is key to attract participants into the programmes. • Provide participants with a central and safe location that is easily accessible and/ or provides childcare facilities promotes acceptability and programme participation. • Life, social, and livelihood skills and capacity building design benefits from the inclusion of and coupling with TVET programmes due to the interconnectedness of both types of skill sets, expressed by a demand and need for both types of skills to enhance participants life capabilities and opportunities. • Incorporate elements that address contexts and norms issues directly such as women's peer-to-peer exchange including mentoring elements, women's self-help groups and clubs • Trainer skills and approaches are an important consideration and factor for participant satisfaction • Continuous sourcing of participant feedback and consultation to allow rapid programme iteration (e.g., programme locations) <p>Contextual considerations to keep in mind</p> <ul style="list-style-type: none"> • Social norms and restrictive contexts determine the space in which programmes can operate and emphasise the need to strike balance between the depth of embedding programmes in these contexts to operate and reach women and the adjustment of programme empowerment outcomes to avoid strong conflict with existing norms and practices. • Failure to adapt the curriculum to each context may affect uptake and relevance. 	

	<ul style="list-style-type: none"> • Lack of men's engagement may be a key barrier to empowerment outcomes that may impede skill development to translate into social change. <p>Good practice example:</p> <ul style="list-style-type: none"> • The Creating Opportunities through Mentoring, Parental Involvement and Safe Spaces (COMPASS) programme faced opposition from parents and community leaders in Congo, Ethiopia and Pakistan with participants and stakeholders questioning the appropriateness of the programme activities. Only through extensive sensitive, targeted awareness-raising and outreach campaigns that included extensive discussions, house-to-house visits and negotiations with community leaders, local authorities, and parents, was consensus achieved and implementation proceeded. Through trust and legitimacy, the programme was able to collaborate with leaders of 'secret' societies associated with female genital mutilation enhancing its reach to participants it would have not been able to access. • In Afghanistan's participants What Works to Prevent Violence Programme participants had access to TVET-linked programme components such as cash crop farming, animal husbandry, and embroidery within a wider empowerment programme and women peers for mentoring including fellow programme participants, trainers and formal mentors provided by the programmes. Access to start-up capital and learning from women in similar contexts based on shared experiences and realities was reported as a key benefit of the arrangement. Women formed social safety nets for each other by staying in touch post the intervention and supporting each other's business ventures. These arrangements resulted in enhanced social capital, a crucial factor in targeting structural barriers to empowerment. 	
Intervention	Intervention works on	Intervention does not work on
Safe spaces	<p>Resources</p> <ul style="list-style-type: none"> • Women have increased access to and ownership of assets, credit and income <p>Agency</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence <p>Achievement</p> <ul style="list-style-type: none"> • Effect unknown - insufficient evidence 	<p>Resources</p> <ul style="list-style-type: none"> • Increased capacity of women to understand and use financial, banking, and business services effectively <p>Agency</p> <ul style="list-style-type: none"> • Women have more and better control over their bodies and sexual health • Women have increased freedom of movement and association • Women are equipped with better life skills that allow them to be prepared for crisis or shocks and recover from them • Reduced support for or occurrence of child and forced marriage • Increased participation in decision making by women at the household or community level

		Achievement <ul style="list-style-type: none">• Women have improved attitudes, self-image and confidence• Improved quality of relationships between women and their household and community members
	Intervention design suggestions to maximise benefits <ul style="list-style-type: none">• Target adolescent girls before they are culturally considered ready to marry to combat trends of forced marriage or teenage pregnancy.• Promote mentorship to encourage openness, collaboration, interaction and learning issues beyond what is learnt through formal education.• Secure safe, convenient meeting locations and timing of the meetings.• Provide incentives and rewards, which can encourage wide participation of the targeted population in programmes, delivered timely and without too stringent conditions.• Incorporate appropriate awareness and communication strategies within households, especially when recruiting adolescent girls to tackle negative perceptions about programme intentions and debates around children's rights.	
	Contextual considerations to keep in mind <ul style="list-style-type: none">• Changes in long-standing social and cultural norms and attitudes towards women or girls may need time to reflect.• Mentors' motivation level and capacity development in advocacy, outreach, communication and networking skills are key to winning the confidence of parents to encourage girls to participate.• Baseline analyses and participatory research before implementation support appropriate and effective safe spaces.	
	Good practice example: <ul style="list-style-type: none">• The Ishraq programme in Egypt targeted girls 13-15 years of age, which was a factor that may contribute to programme impacts where girls 15 and older were considered ready for marriage. Awareness building among participants as well as their household members and the wider community at an early stage of the intervention played a critical role as many of such outcomes were compounded by social, cultural norms and traditions.• The Adolescent Girls Empowerment Programme (AGEP) programme in Zambia highlighted the importance and centrality of mentorship as it provided a platform for girls to open up to matters that they may not be able to share with their families and this built up their self-esteem.	

Implications for research

- More research is needed regarding the effects of gender specific or transformative interventions on peace. Future syntheses and evaluations should address the need for being gender transformative in peacebuilding endeavours and evaluating their effectiveness on peace at the local and subnational levels.
- More rigorous evidence is needed regarding the effects of gender specific or transformative interventions in the Middle East, North Africa and Central and Latin America.
- Lesbian, Gay, Bisexual, Transgender, Queer and/or Questioning, Intersex, and Asexual (LGBTQIA+) groups should be specifically through gender-specific or transformative interventions.
- More research and evidence production is needed on community-based services, quotas, community dialogues and reconciliation, all women police stations and discussion group interventions
- Quasi-experimental studies can make positive contributions to the evidence base, but they need to be carefully designed methodologically to account for confounding and selection bias so that the quality of the studies can be similar to the quality of RCTs and the risks of bias are limited.
- To make meaningful knowledge contributions and to produce reliable evidence, evaluation studies need to follow rigorous methodological designs, and report their research in such a way that allows for transparency and assessment of study quality. Good impact evaluations will register a pre-analysis plan, establish an appropriate counterfactual (either through random assignment or through an appropriate quasi-experimental design) and collect baseline data from all participants. They will collect outcome data from all participants at intervals that are equivalent across the treatment and control conditions. In reporting their findings authors will account for multiple-hypothesis testing in their analysis, and report on all outcomes, regardless of significance. When reporting on these evaluations, authors should also report information related to the assignment mechanism, baseline balance among treatment and control groups, attrition and differential attrition, deviations from the intended intervention, potential spill-over effects, and potential performance bias. Ethical clearances and external validity should also be explicitly discussed.
- There is a need to standardize indicators of gender inequality, and for further research on the development of indices that allow to measure and compare gender-related contexts at the local and/or community level. This includes the development of gender focused indicators.
- A vast majority of impact evaluations measure the effect immediately after the end of an intervention. Investment in evaluations of the longer-term impact of interventions will allow for an examination of whether these types of programmes have lasting effects on the lives of women and girls.
- Researchers should consider using existing scales with documented validity and reliability for measuring empowerment. This would allow for easier comparisons across studies and contexts.
- Researchers need to report sufficient quantitative data for effect size calculation in their impact evaluations. This includes providing standard errors alongside regression coefficients, reporting exact p-values (rather than just an asterisk), and

- providing standard deviations alongside group means.
- Researchers need to report information on the methods used to generate and analyse qualitative data within impact evaluations reports. Without such information, the trustworthiness of the qualitative research becomes questionable and the information from these studies cannot be included in qualitative syntheses.
- A more fluid and widely accepted definition of fragility would result in more inclusion and evidence related to macro-level experiences of conflict. The methods utilised in this review attempt to triangulate and verify the existing definitions, but as noted in the limitations, there is still a lack of global standards. There is a need for further research on the development of indices that allow for measurement and comparison of fragility-related contexts at the local or community level.

5.6 Limitations, potential biases and deviations from the protocol

The following sections present the main limitations that may have impacted our findings and approach for analysis.

5.6.1 Limitations related to the lack of international gender, fragility and conflict indices at the individual, household and community levels

One of the main challenges in defining our inclusion and exclusion criteria was translating international indices based on national-level data into criteria for more local or macro-level analysis. The existing indices, both in regard to gender and to fragility and conflict, were not always constant and may have flaws due to the indicators they use to measure empowerment, gender equality, fragility, conflict, etc. To make the most of sometimes incomplete global indices we adopted a pragmatic approach by combining complementary indices and using the information available in the study to make specific decisions: we have combined the gender inequality index (GII), the women, peace and security (WPS) index, the World Bank list of fragile and conflict-affected states and situations (FCAS), the Organisation for Economic Co-operation and Development (OECD) list of FCAS and set thresholds with a margin of tolerance if the study context was highly relevant. Although we cover 29 countries worldwide, some countries, including South Africa, and parts of Latin America were not eligible. However, they may have had some interesting insights at the local level in areas where conflict or gender inequality are high. Additionally, countries that would not pass our threshold criteria may have interesting subnational insights, given the fact that many conflicts are either subnational or transnational. Future research and policy might consider setting more international, national and local criteria for the measurement of gender equality, women's empowerment, fragility and conflict as it will facilitate more focused research and more targeted implementation.

5.6.2 Limitations related to the absence of standardised outcome measures for gender equality and women's empowerment in fragile and conflict-affected states and situations

Similarly, we observe the relative absence of standardised indicators and outcomes in the gender and FCAS context. Thus, we constructed our set of outcomes by combining existing sets related to gender equality, women's empowerment, peace and security, and the indicators of UNSCR 1325 into a broad set of outcomes that captured each dimension of women's empowerment.

5.6.3 Limitations related to the availability of data for descriptive results and risk of bias

All descriptive results and risk of bias assessments are based on data reported in the included papers. Due to resource and time constraints, we could not contact all authors of the included studies to request additional and/or missing information about their respective study. This may result in some studies being assessed as some risk of bias concerns if they do not report information related to the assessment criteria. The risk of bias assessment may have been different had the information been reported, in which case the assessor could have determined more accurately whether the criteria was met or not met.

5.6.4 Limitations related to the availability of data for qualitative synthesis

We only included studies in the qualitative review component that were linked to included impact evaluations in the quantitative review component. This limits the sample of available evidence in the qualitative component, in particular where interventions are of small scale with little associated programme data publicly available. In addition, our review did not identify linked qualitative studies for all impact evaluations and, what is more, for some intervention categories. This naturally impedes us from being able to conduct an integrated synthesis for these interventions. A final limitation was the lack of reporting of qualitative methods and research processes within impact evaluation reports that included qualitative data, which resulted in the exclusion of potentially relevant documents.

5.6.5 Limitations related to the assessment of publication bias

It is not recommended to test for funnel plot asymmetry with fewer than 10 studies (Sterne et al., 2011), thus we were unable to test for publication bias in the majority of our analyses. This limits our ability to interpret heterogeneity, as it is still unclear whether the variation among effects is being driven by publication bias. More rigorous impact evaluations are needed in order to have a sufficient evidence base to properly assess publication bias.

5.6.6 Limitations related to quantitative study of the existing evidence base

In some instances, the number of included studies for each intervention type and outcome of interest limited our ability to make strong conclusions on the relative effectiveness of the different intervention types. This is notably the case for the following intervention groups: community-based services, quotas, community dialogues and reconciliation, all women police stations, and discussion groups. In addition, in many instances, we were unable to undertake the full moderator analyses to explore sources of variation for all outcomes due to a limited number of included studies in each intervention type, or a lack of variation among included studies.

Another challenge we faced when conducting the quantitative analysis was the lack of a reported precision estimate in some of the included primary studies. In many cases, authors presented regression coefficients without their associated standard errors, and presented just an asterisk indicating the level of significance, rather than an exact p-value. These poor reporting practices make it difficult to calculate precise effect estimates. In many cases, we had to estimate t-values based solely on the asterisk indicating significance level (as described in our published protocol and our methods section). With a maximum assigned value of 2.8 (or -2.8, depending on the direction of the effect), it is likely that we are underestimating the impact of these interventions, given that most of the effects were positive, and that in reality, t-values can certainly exceed 2.8.

5.7 Deviations from the protocol

5.7.1 Adaptation of our list of interventions

Following the extraction process, the team decided to make the list of the interventions more specific to facilitate the use of findings and evidence for policy, practice and research. The main objective of this adaptation was to turn our intervention groups into more granular and specific intervention mechanisms that are directly related to implementation practices.

The team started from the list of included studies and intervention types allocated as per the protocol, which was used as the initial intervention code. The team then looked side by side at each of the studies under each of the original intervention types to identify a specific intervention mechanism. This exercise resulted in a reduction in the number of intervention types from 27 intervention types into 14 specific sets of interventions. This allowed us to avoid working in silos and truly work across the UNSCR pillars.

By linking this new intervention typology to the previous intervention types linked to the UNSCR pillars, the team was able to identify both specific interventions under each pillar and those that are working across pillars; in our case, across the participation and protection pillars.

Table 64: Adaptation of our list of interventions

New intervention types	Protocol intervention types used for the new intervention types	UNSCR 1325 pillar
All-women police stations	Gender-sensitive policing	Protection
Asset transfers	Economic rights and entitlements Economic support asset transfers and livelihoods	Participation
Cash transfers	Cash-based approaches to support women's access to and participation in education and/or the economy	Participation
Community dialogues and reconciliations	Dialogue groups	Prevention
Community-based services	Community and leisure activities	Participation
Discussion groups	Behaviour change communication around support for women's rights and preventing sexual and gender-based violence (SGBV) Civic education and leadership	Participation Protection
Inclusive community-driven development	Voice and participation in local and subnational governance and development bodies	Participation

New intervention types	Protocol intervention types used for the new intervention types	UNSCR 1325 pillar
Institutional provision of loans and savings	Economic rights and entitlements Financial inclusion	Participation
Life, social and livelihood skills and capacity building	Behaviour change communication around support for women's rights and preventing SGBV Capacity building and technical support to subnational government officials to strengthen service provision for women and gender equality Civic education and leadership Financial inclusion Legal rights education Non-formal education Preventative protection measures	Participation Protection
Quotas	Voice and participation in local and subnational governance and development bodies	Participation
Safe Spaces	Behaviour change communication around support for women's rights and preventing SGBV	Participation Protection
Self-help groups and village savings and loan associations	Civil society associations and networks Financial inclusion	Participation
Sensitisation campaigns	Behaviour change communication around support for women's rights and preventing SGBV Preventative protection measures	Protection
Technical and vocational education and training	Technical and vocational education and training	Participation

5.7.2 Adaptation related to the use of Grading recommendations, assessment, development and evaluations

Grading of recommendations, assessment, development and evaluations (GRADE) traces its origins as a tool to rate the certainty of a body of evidence to use in the epidemiological and health sciences. As such, implementing it in the context of international development impact evaluation literature poses a number of challenges and limitations. The tool has been adapted across several domains to better fit the type of available data found in this body of literature and the realities associated with such a large review containing so many analyses. The main limitations that led to adjustments of the tool were:

Two approaches were used in our GRADE analyses that are not reflected in our protocol. First, GRADE recommends that a “large” effect size be defined as a risk ratio of 2. Because this study converted all effect sizes to Cohen’s Ds, that was not applicable. Furthermore, consensus on benchmarking using standardised mean difference in international development is not clear. We used Cochrane’s converter tool to arrive at a D of 0.383 as a “large” effect size. Second, the analyses in this review combined the results of studies that implement both quasi-experimental designs and randomised controlled trials. As such, the “Study Design” column of the Summary of Findings tables show the number of each type of study.

While some studies provided treatment and control group sample sizes, many only provided the total study sample size. GRADE tables in this review were adjusted to reflect that and only report total sample size.

5.7.3 Other deviations

There were several additional deviations to the quantitative approach. First, the team initially expected to use a combination of robust variance estimation (RVE) and traditional meta-analysis approaches. After extracting the data, we determined to use a unified strategy across all analyses. The traditional meta-analysis approach using independent effects was our preferred approach, given that our data often had small numbers of studies that would not allow for accurate or reliable RVE estimates. We also initially intended to winsorise outliers and run models with both winsorised and non-winsorised data. However, given time and resource constraints and the vast number of analyses, we opted for the outlier mechanisms built into the metafor package (Viechtbauer, 2010), including examining studentised residuals and Cook’s distances.

Finally, in our protocol we stated that “Capitalising on recent shifts towards pre-registration of studies and their associated pre-analysis plans, we also examined whether studies that were pre-registered (e.g., on platforms such as ClinicalTrials.gov, the Open Science Foundation, the American Economic Association’s trial registry, or the Registry for International Development Impact Evaluations (RIDIE)) report on all of the outcomes that were proposed in their pre-analysis plans. This additional analysis of outcome reporting bias may draw on methodologies used in previous work, such as the COMPare Trials Project (Goldacre et al. 2016).” Due to time and resource constraints, we were unable to complete this process. Instead, we noted if a PAP was reported in the study and used the methods sections of the paper to look for any outcome reporting bias.

6. Acknowledgements

6.1 Contributions of authors

Etienne Lwamba (EL), Shannon Shisler (SS), Ada Sonnenfeld (AS), Will Ridlehoover (WR), and Meital Kupfer (MK) are the core team for this review. EL is a research associate with experience in the conflict and development sector. SS is a systematic review and quantitative methods expert, with over a decade of experience in designing, managing and analysing quantitative research, including meta-analyses. AS brings relevant field experience as a practitioner in social cohesion in fragile context and co-led the development of previous systematic reviews and evidence gap maps. WR is a research associate with experience in supporting the development, data analysis and management

of systematic review projects. MK is a research associate with field experience in the sector of women's empowerment in FCAS.

Please note that this is the recommended optimal review team composition.

- Content: EL, AS, WR, MK, Nkululeko Tshabalala (NT), Promise Nduku (PN) developed the content of the review with inputs and quality assurance from SS and Laurenz Langer (LL) and with quality assurance from Bidisha Barooah (BB) and Birte Snilsveit (BS).
- Systematic review methods: EL, SS, AS, WR and MK drafted the review methods with inputs and support from the following experts: Laurenz Langer (Qualitative method expert), Bidisha Barooah (Senior Evaluation Specialist) and Sean Grant (GRADE methodology expert).
- Qualitative evidence synthesis: LL designed and oversaw the qualitative review component. NT managed and co-ordinated the research process and activities. PN, MK, NT conducted the search for qualitative evidence, screening, and critical appraisal. NT, PN, LL and a pool of consultants conducted the thematic synthesis.
- Statistical analysis: SS oversaw the statistical analysis. SS, EL, WR, MK and a pool of consultants extracted effects data, with quality assurance from SS, while the core team led outcome classification and grouping with quality assurance and inputs from SS
- Information retrieval: the core team developed the initial search strings and quality assured by John Eysers (Information specialist). EL, WR and MK led screening, grey literature review, snowballing and references checks with the support of a pool of consultants and under the oversight of AS.

6.2 Declarations of interest

There are no reported conflicts of interest on this review.

Several of the review authors are involved with the International Development Coordination Group of the Campbell Collaboration. However, the IDCG editor for this review is not involved in the review. The IDCG's independent co-chair will also independently assure this review.

6.3 Preliminary timeframe

The planned time frame for this review is as follows:

- Draft and Final Protocols: December 2020 - January 2021
- Draft and Final Report: June - October 2021

6.4 Plans for updating this review

The search of this study has been completed in February 2021 based on a search strategy developed in December 2020. The core team will explore opportunities for funding an update of this review upon submission of the final report in 2021.

Online appendixes

Online appendix A: Qualitative analysis tool

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-A-Qualitative-analysis-tool.pdf>

Online appendix B: Data extraction tools

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-B-Data-extraction-tools.pdf>

Online appendix C: Risk of bias assessment tool

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-C-Risk-of-bias-assessment-tool.pdf>

Online appendix D: Title and abstract screening procedure

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-D-Title-and-abstract-screening-procedure.pdf>

Online appendix E: List of terms for the search strategy

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-E-List-of-terms-for-the-search-strategy.pdf>

Online appendix F: Search Strategies

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-F-Search-Strategies.pdf>

Online appendix G: Full Text Screening Checklist

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-G-Full-Text-Screening-Checklist.pdf>

Online appendix H: List of Outcomes

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-H-List-of-Outcomes.pdf>

Online appendix I: Additional Background Information

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-I-Additional-Background-Information.pdf>

Online appendix J: List of excluded studies at full text screening

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-J-List-of-excluded-studies-at-full-text-screening.pdf>

Online appendix K: Descriptive Extraction Dataset

The dataset is available [here](#) and will be granted access upon request to the corresponding author.

Online appendix L: Quantitative and Risk of Bias Datasets

The dataset is available [here](#) and will be granted access upon request to the contact author.

Online appendix M: Report of qualitative coding

The dataset will be granted access upon request to the contact author.

Online appendix N: Overview of descriptive and analytical themes

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-N-Overview-of-descriptive-and-analytical-themes.pdf>

Online appendix O: Characteristics of included studies

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-O-Characteristics-of-included-studies.pdf>

Online appendix P: Studies characteristics tables per intervention

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-P-Studies-characteristics-tables-per-intervention.pdf>

Online appendix Q: Summary table of effects per intervention

<https://www.3ieimpact.org/sites/default/files/2021-10/SR47-Online-appendix-Q-Summary-table-of-effects-per-intervention.pdf>

References

Included studies

Abramsky, T., Devries, K., Kiss, L., Nakuti, J., Kyegombe, N., Starman, E., Cundill, B., Francisco, L., Kaye, D., and Musuya, T. (2014). Findings from the SASA! Study: A Cluster Randomized Controlled Trial to Assess the Impact of A Community Mobilization Intervention to Prevent Violence Against Women and Reduce HIV Risk in Kampala, Uganda. *BMC Medicine*, 12(1), 1–17.

Adoho, F., Chakravarty, S., Korkoyah, D. T., Lundberg, M., and Tasneem, A. (2014). *The Impact of An Adolescent Girls Employment Program: The EPAG Project In Liberia* (No. 6832; Policy Research Working Paper). The World Bank.

Ahmed, A. U., Quisumbing, A. R., Nasreen, M., Hoddinott, J., and Bryan, E. (2009). *Comparing Food and Cash Transfers to The Ultra Poor in Bangladesh* (No. 163; Research Monograph). International Food Policy Research Institute.

Alemu, S. H., Van Kempe, L., and Ruben, R. (2018). Women Empowerment Through Self-help Groups: The Bittersweet Fruits of Collective Apple Cultivation in Highland Ethiopia. *Journal of Human Development and Capabilities*, 19(3), 308–330.

Ambler, K., and De Brauw, A. (2017). *The Impacts of Cash Transfers on Women's Empowerment: Learning from Pakistan's BISP Program* (No. 1702; Social Protection and Labour Discussion Paper). World Bank.

Amin, S., Ahmed, J., Saha, J., Hossain, M., and Haque, E. (2016). *Delaying Child Marriage Through Community-based Skills-development Programs for Girls: Results from A Randomized Controlled Study in Rural Bangladesh*. Population Council.

Ashraf, N., Bau, N., Low, C., and McGinn, K. (2018). *Negotiating A Better Future: How Interpersonal Skills Facilitate Intergenerational Investment* (No. 18–104; Working Paper, pp. 1095–1151). Harvard Business School.

Ashraf, N., Bau, N., Low, C., and McGinn, K. (2020). Negotiating A Better Future: How Interpersonal Skills Facilitate Intergenerational Investment. *The Quarterly Journal of Economics*, 135(2), 1095–1151.

Austrian, K., and Muthengi, E. (2014). Can Economic Assets Increase Girls' Risk of Sexual Harassment? Evaluation Results from A Social, Health and Economic Asset-building Intervention for Vulnerable Adolescent Girls in Uganda. *Children and Youth Services Review*, 47, 168–175.

Austrian, K., Soler-Hampejsek, E., Behrman, J. R., Digitale, J., Hachonda, N. J., Bweupe, M., and Hewett, P. C. (2020). The Impact of The Adolescent Girls Empowerment Program (AGEP) On Short- And Long-term Social, Economic, Education and Fertility Outcomes: A Cluster Randomized Controlled Trial in Zambia. *BMC Public Health*, 20(1), 1–15.

Baird, S., McIntosh, C., and Özler, B. (2010). *Cash or Condition? Evidence from A Cash Transfer Experiment* (Impact Evaluation Series, pp. 1709–1753) [45]. World Bank.

Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., and Sulaiman, M. (2012). *Empowering Adolescent Girls: Evidence from A Randomized Control Trial in Uganda*. World Bank.

Bandiera, O., Buehren, N., Goldstein, M. P., Rasul, I., and Smurra, A. (2018). *The Economic Lives of Young Women in The Time of Ebola: Lessons from An Empowerment Program*. The World Bank.

Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I., and Sulaiman, M. (2017). Labour Markets and Poverty in Village Economies. *The Quarterly Journal of Economics*, 132(2), 811–870.

Bass, J., Murray, S., Cole, G., Bolton, P., Poulton, C., Robinette, K., Seban, J., Falb, K., and Annan, J. (2016). Economic, Social and Mental Health Impacts of An Economic Intervention for Female Sexual Violence Survivors in Eastern Democratic Republic of Congo. *Global Mental Health*, 3.

Bastian, G., Bianchi, I., Goldstein, M., and Montalvao, J. (n.d.). *Short-term Impacts of Improved Access to Mobile Savings, with and Without Business Training: Experimental Evidence from Tanzania* (No. 478; Working Paper, p. 24). Centre for Global Development.

Beaman, L., Duflo, E., Pande, R., and Topalova, P. (2012). Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India. *Science*, 335(6068), 582–586.

Beaman, L., Karlan, D., and Thuysbaert, B. (2014). *Saving for A (Not So) Rainy Day: A Randomized Evaluation of Savings Groups in Mali* (No. 20600; Working Paper). National Bureau of Economic Research.

Beath, A., Christia, F., and Enikolopov, R. (2015). The National Solidarity Programme: Assessing the Effects of Community-driven Development in Afghanistan. *International Peacekeeping*, 22(4), 302–320. <https://doi.org/10.1080/13533312.2015.1059287>

Bedoya, G., Coville, idan, Haushofer, J., Isaqzadeh, M., and Shapiro, J. (n.d.). *Reducing Poverty and Improving Gender Equality Through A Big Push: Experimental Evidence from Afghanistan*. World Bank.

Bedoya, G., Coville, A., Haushofer, J., Isaqzadeh, M. R., and Shapiro, J. (2019). *No Household Left Behind: Afghanistan Targeting the Ultra Poor Impact Evaluation* (No. 8877; Policy Research Working Paper). The World Bank.

Blattman, C., Green, E. P., Jamison, J., Lehmann, M. C., and Annan, J. (2016). The Returns to Microenterprise Support Among the Ultrapoor: A Field Experiment in Postwar Uganda. *American Economic Journal: Applied Economics*, 8(2), 35–64.

Bold, M. van den, Dillon, A., Olney, D., Ouedraogo, M., Pedehombga, A., and Quisumbing, A. (2015). Can Integrated Agriculture-nutrition Programmes Change Gender Norms on Land and Asset Ownership? Evidence from Burkina Faso. *The Journal of Development Studies*, 51(9), 1155–1174. <https://doi.org/10.1080/00220388.2015.1036036>

- Bonilla, J., Zarzur, R. C., Handa, S., Nowlin, C., Peterman, A., Ring, H., and Seidenfeld, D. (2017). Cash for Women's Empowerment? A Mixed-methods Evaluation of The Government of Zambia's Child Grant Program. *World Development*, 95, 55–72. <https://doi.org/10.1016/j.worlddev.2017.02.017>
- Bose, K. S., and Sarma, R. H. (1975). Delineation of The Intimate Details of The Backbone Conformation of Pyridine Nucleotide Coenzymes in Aqueous Solution. *Biochemical and Biophysical Research Communications*, 66(4), 1173–1179. [https://doi.org/10.1016/0006-291x\(75\)90482-9](https://doi.org/10.1016/0006-291x(75)90482-9)
- Breisinger, C., Gilligan, D., Karachiwalla, N., Kurdi, S., El-Enbaby, H., Jilani, A., and Thai, G. (2018). *Impact Evaluation Study for Egypt's Takaful And Karama Cash Transfer Program: Part 1: Quantitative Report* (No. 16; Working Paper). Intl Food Policy Res Inst.
- Brooks, W., Donovan, K., Johnson, T. R., and Oluoch-Aridi, J. (2020). *Cash Transfers as A Response To COVID-19: Experimental Evidence from Kenya*.
- Buehren, N., Chakravarty, S., Goldstein, M., Slavchevska, V., and Sulaiman, M. (2017). *Adolescent Girls' Empowerment in Conflict-affected Settings: Experimental Evidence from South Sudan*.
- Buehren, N., Goldstein, M., Gulesci, S., Sulaiman, M., and Yam, V. (2017). *Evaluation Of an Adolescent Development Program for Girls in Tanzania* (No. 7961; Policy Research Working Paper). The World Bank.
- Carney, C., and Carney, M. H. (2018). Impact of Soil Conservation Adoption on Intra-household Allocations in Zambia. *Review of Development Economics*, 22(4), 1390–1408.
- Chinen, M., and Elmeski, M. (2016). *Evaluation of The Transformative Potential of Positive Gender Socialization in Education for Peace Building*. American Institutes for Research.
- Cilliers, J., Dube, O., and Siddiqi, B. (2018). *Can the Wounds of War Be Healed? Experimental Evidence on Reconciliation in Sierra Leone | 3ie* (No. 75; 3ie Impact Evaluation Report). 3ie. <https://www.3ieimpact.org/evidence-hub/publications/impact-evaluations/can-wounds-war-be-healed-experimental-evidence>
- Clark, S., Kabiru, C. W., Laszlo, S., and Muthuri, S. (2019). The Impact of Childcare on Poor Urban Women's Economic Empowerment in Africa. *Demography*, 56(4), 1247–1272.
- Clayton, A. (2015). Women's political engagement under quota-mandated female representation: Evidence from a randomized policy experiment. *Comparative Political Studies*, 48(3), 333–369.
- Corps, M. (2015). *Improving Child and Maternal Health: Why Adolescent Girl Programming Matters. Post-intervention Evidence from Niger*. Mercy Corps.
- Croke, K., Garcia Mora, M. E., Goldstein, M., Mensah, E. R., and O'Sullivan, M. (2020). *Up Before Dawn: Experimental Evidence from A Cross-border Trader Training at The Democratic Republic of Congo? Rwanda Border* (No. 9123; Policy Research Working Paper). World Bank.

- Croke, K., Goldstein, M., and Holla, A. (2017). *Can Job Training Decrease Women's Self-defeating Biases? Experimental Evidence from Nigeria* (No. 8141; Policy Research Working Paper). World Bank.
- Das, A., Mogford, E., Singh, S. K., Barbhuiya, R. A., Chandra, S., and Wahl, R. (2012). Reviewing Responsibilities and Renewing Relationships: An Intervention with Men on Violence Against Women in India. *Culture, Health and Sexuality*, 14(6), 659–675.
- De Hoop, T., Van Kempen, L., Linssen, R., and Van Eerdewijk, A. (2010). *Women's Autonomy and Subjective Well-being in India: How Village Norms Shape the Impact of Self-help Groups* (No. 25921; MPRA Paper). MPRA.
- Decker, M. R., Wood, S. N., Hameeduddin, Z., Kennedy, S. R., Perrin, N., Tallam, C., Akumu, I., Wanjiru, I., Asira, B., Frankel, A., Omondi, B., Case, J., Clough, A., Otieno, R., Mwit, M., and Glass, N. (2020). Safety Decision-making and Planning Mobile App for Intimate Partner Violence Prevention and Response: Randomised Controlled Trial in Kenya. *BMJ Global Health*, 5(7), e002091. <https://doi.org/10.1136/bmjgh-2019-002091>
- Decker, M. R., Wood, S. N., Ndinda, E., Yenokyan, G., Sinclair, J., Maksud, N., Ross, B., Omondi, B., and Ndirangu, M. (2018). Sexual Violence Among Adolescent Girls and Young Women in Malawi: A Cluster-randomized Controlled Implementation Trial of Empowerment Self-defence Training. *BMC Public Health*, 18(1), 1–12.
- Desaib, R. M., and Joshic, S. (2013). *Collective Action and Community Development: Evidence from Women's Self-help Groups in Rural India*.
- Deininger, K., & Liu, Y. (2013). Economic and Social Impacts of an Innovative Self-Help Group Model in India. *World Development*, 43, 149–163. <https://doi.org/10.1016/j.worlddev.2012.09.019>
- Duflo, E., Banerjee, A., Glennerster, R., and Kinnan, C. G. (n.d.). *The Miracle of Microfinance? Evidence from A Randomized Evaluation* (No. 18950; NBER Working Paper Series, p. 62). NBER.
- Dunkle, K., Stern, E., Chatterji, S., and Heise, L. (2020). Effective Prevention of Intimate Partner Violence Through Couples Training: A Randomised Controlled Trial of Indashyikirwa In Rwanda. *BMJ Global Health*, 5(12), e002439. <https://doi.org/10.1136/bmjgh-2020-002439>
- Erulkar, A., and Medhin, G. (2017). Evaluation of A Safe Spaces Program for Girls in Ethiopia. *Girlhood Studies*, 10(1), 107–125. <https://doi.org/10.3167/ghs.2017.100108>
- Eze Eze, D. (2019). Microfinance Programs and Domestic Violence in Northern Cameroon; The Case of The Familial Rural Income Improvement Program. *Review of Economics of the Household*, 17(3), 947–967. <https://doi.org/10.1007/s11150-017-9393-x>
- Field, E., Jayachandran, S., and Pande, R. (2010). Do Traditional Institutions Constrain Female Entrepreneurship? A Field Experiment on Business Training in India. *American Economic Review*, 100(2), 125–129. <https://doi.org/10.1257/aer.100.2.125>
- Figueroa, M. E., Poppe, P., Carrasco, M., Pinho, M. D., Massingue, F., Tanque, M., and Kwizera, A. (2016). Effectiveness of Community Dialogue in Changing Gender and Sexual

Norms for HIV Prevention: Evaluation of The Tchova Tchova Program in Mozambique. *Journal of Health Communication*, 21(5), 554–563.
<https://doi.org/10.1080/10810730.2015.1114050>

Fuller, R. (2014). *Strengthening and Linking Women-Led Efforts to Promote Women's Property and Literacy Rights in Sierra Leone—Project Effectiveness Review—Summary Report* (p. 17). Oxfam.

Gelayay, D. (n.d.). *Did conditional cash transfers in the Productive Safety Net Program empower women in Tigray, north-east Ethiopia?* (Discussion Paper, p. 31). Institute of Development Policy.

Gibbs, A., Corboz, J., Chirwa, E., Mann, C., Karim, F., Shafiq, M., Mecagni, A., Maxwell-Jones, C., Noble, E., and Jewkes, R. (2020). The Impacts of Combined Social and Economic Empowerment Training on Intimate Partner Violence, Depression, Gender Norms and Livelihoods Among Women: An Individually Randomised Controlled Trial and Qualitative Study in Afghanistan. *BMJ Global Health*, 5(3), e001946.
<https://doi.org/10.1136/bmjgh-2019-001946>

Glass, N., Perrin, N. A., Kohli, A., Campbell, J., and Remy, M. M. (2017). Randomised Controlled Trial of A Livestock Productive Asset Transfer Programme to Improve Economic and Health Outcomes and Reduce Intimate Partner Violence in A Post conflict Setting. *BMJ Global Health*, 2(1), e000165. <https://doi.org/10.1136/bmjgh-2016-000165>

Gobin, V. J., Santos, P., and Toth, R. (2017). No Longer Trapped? Promoting Entrepreneurship Through Cash Transfers to Ultra-poor Women in Northern Kenya. *American Journal of Agricultural Economics*, 99(5), 1362–1383. <https://doi.org/10.1093/ajae/aax037>

Gottlieb, J. (2016). Why Might Information Exacerbate the Gender Gap in Civic Participation? Evidence from Mali. *World Development*, 86, 95–110.
<https://doi.org/10.1016/j.worlddev.2016.05.010>

Green, D. P., Wilke, A. M., and Cooper, J. (2020). Countering Violence Against Women by Encouraging Disclosure: A Mass Media Experiment in Rural Uganda. *Comparative Political Studies*, 53(14), 2283–2320. <https://doi.org/10.1177/0010414020912275>

Green, E. P., Blattman, C., Jamison, J., and Annan, J. (2015). Women's Entrepreneurship and Intimate Partner Violence: A Cluster Randomized Trial of Microenterprise Assistance and Partner Participation in Post-conflict Uganda (SSM-D-14-01580R1). *Social Science and Medicine*, 133, 177–188. <https://doi.org/10.1016/j.socscimed.2015.03.042>

Halim, N., Mzilangwe, E. S., Reich, N., Badi, L., Simmons, E., Servidone, M., II, N. B. H., Kawemama, P., and Messersmith, L. J. (2019). *Together to End Violence Against Women in Tanzania: Results of A Pilot Cluster Randomized Controlled Trial to Evaluate Preliminary Effectiveness of Interpersonal and Community Level Interventions to Address Intimate Partner Violence*. 1653–1668.

Handa, S., Peterman, A., Huang, C., Halpern, C., Pettifor, A., & Thirumurthy, H. (2015). Impact of the Kenya Cash Transfer for Orphans and Vulnerable Children on early pregnancy and marriage of adolescent girls. *Social Science & Medicine* (1982), 141, 36–45. <https://doi.org/10.1016/j.socscimed.2015.07.024>

Haushofer, J., and Shapiro, J. (n.d.). *The Long-term Impact of Unconditional Cash Transfers: Experimental Evidence from Kenya* (p. 64).

Heath, R., Hidrobo, M., and Roy, S. (2020). Cash Transfers, Polygamy, And Intimate Partner Violence: Experimental Evidence from Mali. *Journal of Development Economics*, 143, 102410. <https://doi.org/10.1016/j.jdeveco.2019.102410>

Hossain, M., Zimmerman, C., Kiss, L., Abramsky, T., Kone, D., Bakayoko-Topolska, M., Annan, J., Lehmann, H., and Watts, C. (2014). Working with Men to Prevent Intimate Partner Violence in A Conflict-affected Setting: A Pilot Cluster Randomized Controlled Trial in Rural Côte D'ivoire. *BMC Public Health*, 14(1), 339. <https://doi.org/10.1186/1471-2458-14-339>

Ifelunini, I. A., and Wosowei, E. C. (2012). Does Micro Finance Reduce Poverty Among Women Entrepreneurs in South-south Nigeria? Evidence from Propensity Score Matching Technique. *European Journal of Business and Management*, 4, 13.

Iqbal, T., Farooq, S., and Padda, I. U. H. (2021). Can Empowerment Be Enhanced by Putting Cash in The Hands of Poor Women? Learning from Pakistan's BISP Program. *The European Journal of Development Research*, 33(3), 760–792. <https://doi.org/10.1057/s41287-020-00320-w>

Ismayilova, L., Karimli, L., Gaveras, E., Tô-Camier, A., Sanson, J., Chaffin, J., and Nanema, R. (2018). An Integrated Approach to Increasing Women's Empowerment Status and Reducing Domestic Violence: Results of A Cluster-randomized Controlled Trial in A West African Country. *Psychology of Violence*, 8(4), 448–459. <https://doi.org/10.1037/vio0000136>

Jassal, N. (2020). Gender, Law Enforcement, And Access to Justice: Evidence from All-women Police Stations in India. *American Political Science Review*, 114(4), 1035–1054. <https://doi.org/10.1017/S0003055420000684>

Johnson, S. W. (2018). Post-conflict Reconstruction, Microfinance and Democratic Engagement. *Peace Economics, Peace Science and Public Policy*, 24(3), 20170048. <https://doi.org/10.1515/peps-2017-0048>

Karimli, L., Lecoutere, E., Wells, C. R., and Ismayilova, L. (2021). More Assets, More Decision-making Power? Mediation Model in A Cluster-randomized Controlled Trial Evaluating the Effect of The Graduation Program on Women's Empowerment in Burkina Faso. *World Development*, 137, 105159. <https://doi.org/10.1016/j.worlddev.2020.105159>

Karlan, D., Savonitto, B., Thuysbaert, B., and Udry, C. (2017). Impact of savings groups on the lives of the poor. *Proceedings of the National Academy of Sciences*, 114(12), 3079–3084.

Kumar, N., Nguyen, P. H., Harris, J., Harvey, D., Rawat, R., and Ruel, M. T. (2018). What It Takes: Evidence from A Nutrition- And Gender-sensitive Agriculture Intervention in Rural Zambia. *Journal of Development Effectiveness*, 10(3), 341–372. <https://doi.org/10.1080/19439342.2018.1478874>

- Larson, A. M., Solis, D., Duchelle, A. E., Atmadja, S., Resosudarmo, I. A. P., Dokken, T., and Komalasari, M. (2018). Gender Lessons for Climate Initiatives: A Comparative Study Of REDD+ Impacts on Subjective Wellbeing. *World Development*, 108, 86–102.
<https://doi.org/10.1016/j.worlddev.2018.02.027>
- Laudati, A., Mvukiyehe, E., and van der Windt, P. (2018). *Participatory Development in Fragile and Conflict-affected Contexts: An Impact Evaluation of the Tuungane 1 Program in the Democratic Republic of the Congo*.
- Lecoutere, E., and Wuyts, E. (2021a). Confronting the Wall of Patriarchy: Does Participatory Intrahousehold Decision Making Empower Women in Agricultural Households? *The Journal of Development Studies*, 57(6), 882–905.
<https://doi.org/10.1080/00220388.2020.1849620>
- Lecoutere, E., and Wuyts, E. (2021b). Confronting the Wall of Patriarchy: Does Participatory Intrahousehold Decision Making Empower Women in Agricultural Households? *The Journal of Development Studies*, 57(6), 882–905.
<https://doi.org/10.1080/00220388.2020.1849620>
- Leight, J., Deyessa, N., Verani, F., Tewolde, S., and Sharma, V. (2021). Community-level Spill over Effects of An Intervention to Prevent Intimate Partner Violence and HIV Transmission in Rural Ethiopia. *BMJ Global Health*, 6(1), e004075.
<https://doi.org/10.1136/bmjgh-2020-004075>
- Lombardini, S., and Bowman, K. (2015). *Women's Empowerment in Pakistan: Impact Evaluation of The Empowering Small-Scale Producers in The Dairy Sector Project*. Oxfam GB. <http://hdl.handle.net/10546/580463>
- Lombardini, S., and Yoshikawa, K. (2015). *Women's Empowerment in Uganda: Impact Evaluation of The Project 'Piloting Gender Sensitive Livelihoods in Karamoja'*. Oxfam GB. <http://hdl.handle.net/10546/592575>
- Lubega, P., Nakakawa, F., Narciso, G., Newman, C., and Kityo, C. (2017). *Inspiring women: Experimental evidence on sharing entrepreneurial skills in Uganda* (No. 2017; TEP Working Paper, p. 57).
- Lwamba, E., Ridlehoover, W., Kupfer, M., Shisler, S., Sonnenfeld, A., Langer, L., Eysers, J., Grant, S., & Barooah, B. (2021). Protocol: Strengthening women's empowerment and gender equality in fragile contexts towards peaceful and inclusive societies: A systematic review and meta-analysis. *Campbell Systematic Reviews*, 17(3), e1180.
<https://doi.org/10.1002/cl2.1180>
- Maitra, P., and Mani, S. (2017). Learning and Earning: Evidence from A Randomized Evaluation in India. *Labour Economics*, 45, 116–130.
<https://doi.org/10.1016/j.labeco.2016.11.007>
- Malaeb, B., and Uzor, E. (2017). *Multidimensional impact evaluation: A randomized control trial on conflict-affected women in Northern Uganda* (Working Paper). Centre for the Study of the Economies of Africa.

- Marniemi, J., and Parkki, M. G. (1975). Radiochemical Assay of Glutathione S-epoxide Transferase and Its Enhancement by Phenobarbital in Rat Liver In Vivo. *Biochemical Pharmacology*, 24(17), 1569–1572. [https://doi.org/10.1016/0006-2952\(75\)90080-5](https://doi.org/10.1016/0006-2952(75)90080-5)
- Mekonnen, T. (2017). *Financing rural households and its impact: Evidence from randomized field experiment data* (Working Paper). Maastricht University.
- Mukherjee, A. K., and Kundu, A. (n.d.). *Impact of Swarna Jayanti Gram Swarajgar Yojana (SGSY) on Health, Education and Women Empowerment* (No. 33258; MPRA Paper, p. 23). Munich Personal RePEc Archive.
- Müller, A., Pape, U., and Ralston, L. (2019). *Broken Promises: Evaluating an Incomplete Cash Transfer Program* (Policy Research Working Paper). The World Bank. <https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-9016>
- Murray, L. K., Kane, J. C., Glass, N., Wyk, S. S. van, Melendez, F., Paul, R., Danielson, C. K., Murray, S. M., Mayeya, J., Simenda, F., and Bolton, P. (2020). Effectiveness of The Common Elements Treatment Approach (CETA) In Reducing Intimate Partner Violence and Hazardous Alcohol Use in Zambia (VATU): A Randomized Controlled Trial. *PLOS Medicine*, 17(4), e1003056. <https://doi.org/10.1371/journal.pmed.1003056>
- Mvukiyehe, E. (2017). *Can Media Interventions Reduce Gender Gaps in Political Participation after Civil War? Evidence from a Field Experiment in Rural Liberia* (Policy Research Working Paper No. 7942). World Bank.
- Nandi, A., Agarwal, P., Chandrashekar, A., and Harper, S. (2020). Access to Affordable Daycare And Women's Economic Opportunities: Evidence from A Cluster-randomised Intervention in India. *Journal of Development Effectiveness*, 12(3), 219–239. <https://doi.org/10.1080/19439342.2020.1773898>
- Natali, L., Handa, S., Peterman, A., and Seidenfeld, D. (2016). *Making Money Work: Unconditional cash transfers allow women to save and re-invest in rural Zambia* INNOCENTI WORKING PAPERS [Innocenti Working Paper].
- Nkonya, E., Phillip, D., Mogue, T., Pender, J., and Kato, E. (2012). Impacts of Community-driven Development Programs on Income and Asset Acquisition in Africa: The Case of Nigeria. *World Development*, 40(9), 1824–1838. <https://doi.org/10.1016/j.worlddev.2012.04.028>
- Noble, E., Corboz, J., Gibbs, A., Mann, C., Mecagni, A., and Jewkes, R. (2019). *IMPACT EVALUATION OF WOMEN FOR WOMEN INTERNATIONAL'S ECONOMIC AND SOCIAL EMPOWERMENT PROGRAMME IN AFGHANISTAN: AN EVIDENCE BRIEF*.
- Olajide, D., Ikenwilo, D., Obembe, O. B., Ibeji, N., and Akindola, R. (2016). *The Impact of A Rural Microcredit Scheme Targeting Women on Household Vulnerability and Empowerment: Evidence from South West Nigeria* (SSRN Scholarly Paper ID 3167363). Social Science Research Network. <https://papers.ssrn.com/abstract=3167363>

- Olney, D. K., Bliznashka, L., Pedehombga, A., Dillon, A., Ruel, M. T., and Heckert, J. (2016). A 2-year Integrated Agriculture and Nutrition Program Targeted to Mothers of Young Children in Burkina Faso Reduces Underweight Among Mothers and Increases Their Empowerment: A Cluster-randomized Controlled Trial. *The Journal of Nutrition*, 146(5), 1109–1117. <https://doi.org/10.3945/jn.115.224261>
- Özler, B., Hallman, K., Guimond, M.-F., Kelvin, E. A., Rogers, M., and Karnley, E. (2020). Girl Empower – A Gender Transformative Mentoring and Cash Transfer Intervention to Promote Adolescent Wellbeing: Impact Findings from A Cluster-randomized Controlled Trial in Liberia. *SSM - Population Health*, 10, 100527. <https://doi.org/10.1016/j.ssmph.2019.100527>
- Pradhan, M. A. H., and Sulaiman, J. (2017). Impact of Vulnerable Group Development (VGD) Program on Improvement of Woman Headed Household Consumption Diversity in Bangladesh. *JOURNAL OF SOCIAL SCIENCE RESEARCH*, 11(1), 2292–2305. <https://doi.org/10.24297/jssr.v11i1.5746>
- R, Q., Agnes, Akhter, A., F, H., John, Audrey, P., and Shalini, R. (2020). *Designing for Empowerment Impact in Agricultural Development Projects: Experimental Evidence from The Agriculture, Nutrition, And Gender Linkages (Angel) Project in Bangladesh*. Intl Food Policy Res Inst.
- Roy, S., Hidrobo, M., Hoddinott, J., and Ahmed, A. (2019). Transfers, Behaviour Change Communication, And Intimate Partner Violence: Postprogram Evidence from Rural Bangladesh. *The Review of Economics and Statistics*, 101(5), 865–877. https://doi.org/10.1162/rest_a_00791
- Scales, P. C., Benson, P. L., Dershem, L., Fraher, K., Makonnen, R., Nazneen, S., Syvertsen, A. K., and Titus, S. (2013). Building Developmental Assets to Empower Adolescent Girls in Rural Bangladesh: Evaluation of Project Kishoree Kontha. *Journal of Research on Adolescence*, 23(1), 171–184. <https://doi.org/10.1111/j.1532-7795.2012.00805.x>
- Schensul, S. L., Singh, R., Schensul, J. J., Verma, R. K., Burleson, J. A., and Nastasi, B. K. (2015). Community Gender Norms Change as A Part of A Multilevel Approach to Sexual Health Among Married Women in Mumbai, India. *American Journal of Community Psychology*, 56(1–2), 57–68. <https://doi.org/10.1007/s10464-015-9731-1>
- Sieverding, M., and Elbadawy, A. (2016). Empowering Adolescent Girls in Socially Conservative Settings: Impacts and Lessons Learned from the Ishraq Program in Rural Upper Egypt. *Studies in Family Planning*, 47(2), 129–144.
- Smith, R. J., and Bryant, R. G. (1975). Metal Substitutions Incarbonic Anhydrase: A Halide Ion Probe Study. *Biochemical and Biophysical Research Communications*, 66(4), 1281–1286. [https://doi.org/10.1016/0006-291x\(75\)90498-2](https://doi.org/10.1016/0006-291x(75)90498-2)
- Stark, L., Asghar, K., Seff, I., Yu, G., Tesfay Gessesse, T., Ward, L., Assazenew Baysa, A., Neiman, A., and Falb, K. L. (2018). Preventing Violence Against Refugee Adolescent Girls: Findings from A Cluster Randomised Controlled Trial in Ethiopia. *BMJ Global Health*, 3(5), e000825. <https://doi.org/10.1136/bmjgh-2018-000825>

Stark, L., Seff, I., Asghar, K., Roth, D., Bakamore, T., MacRae, M., Fanton D'Andon, C., and Falb, K. L. (2018). Building Caregivers' Emotional, Parental and Social Support Skills to Prevent Violence Against Adolescent Girls: Findings from A Cluster Randomised Controlled Trial in Democratic Republic of Congo. *BMJ Global Health*, 3(5), e000824. <https://doi.org/10.1136/bmjgh-2018-000824>

Tanner, S., and O'Connor, M. (2017). *A Safe Place to Shine: Creating opportunities and raising the voice of adolescent girls in humanitarian settings*. International Rescue Committee.

Tarozzi, A., Desai, J., and Johnson, K. (2013). *On the Impact of Microcredit: Evidence from a Randomized Intervention in Rural Ethiopia* (Working Paper, p. 37).

Vaillant, J., Koussoubé, E., Roth, D., Pierotti, R., Hossain, M., and Falb, K. L. (2020). Engaging Men to Transform Inequitable Gender Attitudes and Prevent Intimate Partner Violence: A Cluster Randomised Controlled Trial in North and South Kivu, Democratic Republic of Congo. *BMJ Global Health*, 5(5), e002223. <https://doi.org/10.1136/bmjgh-2019-002223>

Vigneri, M., and Lombardini, S. (n.d.). *Women's empowerment in Mali: Impact evaluation of the educational project "Girls CAN - Promoting secondary education in West Africa"* (p. 50). Oxfam.

Weber, O., and Ahmad, A. (2014). Empowerment Through Microfinance: The Relation Between Loan Cycle and Level of Empowerment. *World Development*, 62, 75–87. <https://doi.org/10.1016/j.worlddev.2014.05.012>

World Bank, McKenzie, D., Puerto, S., International Labour Organization, Odhiambo, F., and Innovations for Poverty Action. (2019). *Unpacking the Determinants of Entrepreneurship Development and Economic Empowerment for Women in Kenya*. International Initiative for Impact Evaluation (3ie). <http://www.3ieimpact.org/evidence-hub/publications/impact-evaluations/unpacking-determinants-entrepreneurship-development>

Yaron, G., Gordon, R., Best, J., and Choudhary, S. (2018). Microfinance for The Marginalized: The Impact of The Rojiroti Approach in India. *Enterprise Development and Microfinance*, 29(1), 80–93. <https://doi.org/10.3362/1755-1986.17-00011>

Linked Studies

Baird, S., Chirwa, E., de Hoop, J., and Özler, B. (2013). *Girl Power: Cash Transfers and Adolescent Welfare. Evidence from A Cluster-randomized Experiment in Malawi* (No. 19479; p. w19479). National Bureau of Economic Research. <http://www.nber.org/papers/w19479.pdf>

Baird, S., Chirwa, E., McIntosh, C., and Özler, B. (2009). *The Short-term Impacts of A Schooling Conditional Cash Transfer Program on The Sexual Behaviour of Young Women* (Policy Research Working Paper No. 40; p. 33). World Bank. <https://openknowledge.worldbank.org/bitstream/handle/10986/4281/WPS5089.pdf>

- Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., and Sulaiman, M. (2020). Women's Empowerment in Action: Evidence from A Randomized Control Trial in Africa. *American Economic Journal: Applied Economics*, 12(1), 210–259. <https://doi.org/10.1257/app.20170416>
- Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I., and Sulaiman, M. (2013). *Can Basic Entrepreneurship Transform the Economic Lives of The Poor?* (IZA Discussion Paper No. 7386; p. 56). IZA. <https://www.econstor.eu/bitstream/10419/80628/1/745852602.pdf>
- Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2015). The Miracle of Microfinance? Evidence from a Randomized Evaluation. *American Economic Journal: Applied Economics*, 7(1), 22–53. <https://doi.org/10.1257/app.20130533>
- Beath, A., Christia, F., and Enikolopov, R. (2013a). *Randomized Impact Evaluation Of Afghanistan's National Solidarity Programme* (No. 81107). World Bank.
- Beath, A., Christia, F., and Enikolopov, R. (2013b). Empowering Women Through Development Aid: Evidence From A Field Experiment In Afghanistan. *American Political Science Review*, 107(3), 540–557. <https://doi.org/10.1017/S0003055413000270>
- Blattman, C., Fiala, N., Berlin, D., and Martinez, S. (2013). *The Economic And Social Returns To Cash Transfers: Evidence From A Ugandan Aid Program* (p. 53).
- Das, N., Yasmin, R., and Ara, J. (2013). *How Do Intrahousehold Dynamics Change When Assets Are Transferred To Women?* (IFPRI Discussion Paper No. 01317; p. 44).
- De Hoop, T., van Kempen, L., Linssen, R., & van Eerdewijk, A. (2014). Women's Autonomy and Subjective Well-Being: How Gender Norms Shape the Impact of Self-Help Groups in Odisha, India. *Feminist Economics*, 20(3), 103–135. <https://doi.org/10.1080/13545701.2014.893388>
- Gibbs, A., Corboz, J., Shafiq, M., Marofi, F., Mecagni, A., Mann, C., Karim, F., Chirwa, E., Maxwell-Jones, C., and Jewkes, R. (2018). An Individually Randomized Controlled Trial To Determine The Effectiveness Of The Women For Women International Programme In Reducing Intimate Partner Violence And Strengthening Livelihoods Amongst Women In Afghanistan: Trial Design, Methods And Baseline Findings. *BMC Public Health*, 18(1), 164. <https://doi.org/10.1186/s12889-018-5029-1>
- Green, D. P., Wilke, A., and Cooper, J. (2018). *Silence Begets Violence: A Mass Media Experiment To Prevent Violence Against Women In Rural Uganda*. Newhaven, CT: Innovations for Poverty.
- Green, E. P., Blattman, C., Jamison, J., and Annan, J. (2016). Corrigendum To "Women's Entrepreneurship And Intimate Partner Violence: A Cluster Randomized Trial Of Microenterprise Assistance And Partner Participation In Post-conflict Uganda" [Soc. Sci. Med. 133 (2015) 177–188]. *Social Science and Medicine*, 148, 139–141. <https://doi.org/10.1016/j.socscimed.2015.11.042>

- Guilbert, N., Vaillant, J., Falb, K., Mallinga, P., and Roth, D. (2016). *Preventing Intimate-partner Violence: Impact Evaluation Of Engaging Men Through Accountable Practice In Eastern DRC*. [RIDIE-STUDY-ID-576d513b31538](#)
- Haushofer, J., Ringdal, C., Shapiro, J. P., and Wang, X. Y. (2019). *Income Changes And Intimate Partner Violence: Evidence From Unconditional Cash Transfers In Kenya*. National Bureau of Economic Research.
- Haushofer, J., and Shapiro, J. (2013). *Household Response To Income Changes: Evidence From An Unconditional Cash Transfer Program In Kenya* (pp. 1–57).
- Haushofer, J., and Shapiro, J. (2016). The Short-term Impact Of Unconditional Cash Transfers To The Poor: Experimental Evidence From Kenya. *The Quarterly Journal of Economics*, 131(4), 1973–2042.
- Iqbal, T., Padda, I. U. H., and Farooq, S. (2020). Unconditional Cash Transfers And Women Empowerment: The Case Of Benazir Income Support Programme (BISP) In Pakistan. *Journal of Business and Social Review in Emerging Economies*, 6(2), 401–418.
- Karimli, L., Bose, B., and Kagotho, N. (2019). Integrated Graduation Program And Its Effect On Women And Household Economic Well-being: Findings From A Randomised Controlled Trial In Burkina Faso. *The Journal of Development Studies*, 56(7), 1277–1294. <https://doi.org/10.1080/00220388.2019.1677887>
- Lecoutere, E. (2018). *Improving intrahousehold cooperation for efficient smallholder farming. A field experiment in central Uganda* [Working Paper]. Institute of Development Policy.
- Roy, S., Ara, J., Das, N., and Quisumbing, A. R. (2013). *Asset Transfers And Intrahousehold Dynamics: Evidence From Brac's 'Targeting The Ultra Poor' program In Bangladesh*.
- Roy, S., Ara, J., Das, N., and Quisumbing, A. R. (2015). “Flypaper Effects” In Transfers Targeted To Women: Evidence From Brac’s “Targeting The Ultra Poor” Program In Bangladesh. *Journal of Development Economics*, 117, 1–19.
- Sarah, B., McIntosh, C. T., and Özler, B. (2017). *When The Money Runs Out: Do Cash Transfers Have Sustained Effects On Human Capital Accumulation?* (No. 68; WPS). Center for Effective Global Action. <https://escholarship.org/content/qt2rd3f9jv/qt2rd3f9jv.pdf>
- Sharma, V., Leight, J., Verani, F., Tewolde, S., and Deyessa, N. (2020). Effectiveness Of A Culturally Appropriate Intervention To Prevent Intimate Partner Violence And HIV Transmission Among Men, Women, And Couples In Rural Ethiopia: Findings From A Cluster-randomized Controlled Trial. *PLoS Medicine*, 17(8), e1003274.
- Stark, L., Seff, I., Assezenew, A., Eoomkham, J., Falb, K., and Ssewamala, F. M. (2018). Effects Of A Social Empowerment Intervention On Economic Vulnerability For Adolescent Refugee Girls In Ethiopia. *Journal of Adolescent Health*, 62(1), S15–S20.

Tanner, S., and O'Connor, M. (2017a). *A Safe Place To Shine: Creating Opportunities And Raising Voices Of Adolescent Girls In Humanitarian Settings*. International Rescue Committee.

Tanner, S., and O'Connor, M. (2017b). *A Safe Place To Shine: Learning From The COMPASS Programme In Ethiopia*. International Rescue Committee.

Tarozzi, A., Desai, J., & Johnson, K. (2015). The Impacts of Microcredit: Evidence from Ethiopia. *American Economic Journal: Applied Economics*, 7(1), 54–89.
<https://doi.org/10.1257/app.20130475>

Van der Windt, P. (2018). Can development aid empower women? Evidence from a field experiment in the Congo. *The Journal of Politics*, 80(3), 1039–1044.

Van der Windt, P. (2020). *Assessing the Longer Term Impact of Community-Driven Development Programs: Evidence from a Field Experiment in the Democratic Republic of Congo* (No. 9140; Research Working Paper). World Bank.

Van der Windt, P., Humphreys, M., and de la Sierra, R. S. (2018). Gender Quotas In Development Programming: Null Results From A Field Experiment In Congo. *Journal of Development Economics*, 133, 326–345.

Watts, C., Devries, K., Kiss, L., Abramsky, T., Kyegombe, N., and Michau, L. (2019). *The SASA! Study: A Cluster Randomised Trial To Assess The Impact Of A Violence And HIV Prevention Programme In Kampala, Uganda* (3ie Impact Evaluation Report No. 24). 3ie.
<https://dataverse.harvard.edu/citation?persistentId=doi:10.7910/DVN/DRXLQF>

Qualitative and process evaluations

Adoho, F., Chakravarty, S., Korkoyah, D. T., Lundberg, M., and Tasneem, A. (2014). The impact of an adolescent girls employment program: The EPAG project in Liberia. The World Bank.

Ahmed, A. U., Quisumbing, M. A. R., Hoddinott, J., Nasreen, M., and Bryan, E. (2007). Relative efficacy of food and cash transfers in improving food security and livelihoods of the ultra-poor in Bangladesh. WFP, World Food Programme.

Ahmed, A. U., Sraboni, E., and Shaba, F. K. (2014). Safety nets in Bangladesh: which form of transfer is most beneficial. Washington: International Food Policy Research Institute.

Amaral S., Bhalotraz S., and Prakash N. (2018). Gender, Crime And Punishment: Evidence From Women Police Stations In India. Essex: Economic And Social Research Council (ESRC), Research Centre On Microsocial Change (Misoc). Retrieved at: <https://ideas.repec.org/p>

Ambler, K., and De Brauw, A. (2017). The impacts of cash transfers on women's empowerment: learning from Pakistan's BISP program. World Bank.

- Ayuku, D., Embleton, L., Koech, J., Atwoli, L., Hu, L., Ayaya, S., ... and Braitstein, P. (2014). The government of Kenya cash transfer for orphaned and vulnerable children: cross-sectional comparison of household and individual characteristics of those with and without. *BMC international health and human rights*, 14(1), 1-14.
- Azadi, H., De Rudder, F., Vlassenroot, K., Nega, F., and Nyssen, J. (2017). Targeting international food aid programmes: The case of productive safety net programme in Tigray, Ethiopia. *Sustainability*, 9(10), 1716.
- Baird, S., Chirwa, E., McIntosh, C., and Özler, B. (2015). What happens once the intervention ends? The medium-term impacts of a cash transfer programme in Malawi. *3ie Impact Evaluation Report*, 27.
- Baird, S., McIntosh, C., and Ozler, B. (2009). Designing cost-effective cash transfer programs to boost schooling among young women in Sub-Saharan Africa. *World Bank Policy Research Working Paper*, (5090).
- Bandiera, O., Buehren, N., Goldstein, M. P., Rasul, I., and Smurra, A. (2019). The Economic Lives of Young Women in the Time of Ebola: Lessons from an Empowerment Program. *The World Bank*.
- Barakat, S., and Strand, A. (2006). Mid-term evaluation report of the National Solidarity Programme (NSP), Afghanistan.
- Bass, J., Bolton, P., Murray, S., Cole, G., Robinette, K., Poulton, C., ... and Annan, J. (2014). Study of effectiveness of a social-economic intervention for sexual violence survivors in Eastern DRC.
- Bass, J., Murray, S., Cole, G., Bolton, P., Poulton, C., Robinette, K., ... and Annan, J. (2016). Economic, social and mental health impacts of an economic intervention for female sexual violence survivors in Eastern Democratic Republic of Congo. *Global Mental Health*, 3.
- Beaman, L., Karlan, D., and Thuysbaert, B. (2014). Saving for a (not so) rainy day: A randomized evaluation of savings groups in Mali (No. w20600). *National Bureau of Economic Research*.
- Bedoya, G., Coville, A., Haushofer, J., Isaqzadeh, M. R., and Shapiro, J. (2019). No household left behind: Afghanistan targeting the ultra poor impact evaluation. *The World Bank*.
- Berlie, A. B. (2018). The Role of Productive Safety Net Programme (PSNP) in Improving Female-Headed Households' Food Security: The Case of Enebsie Sar Midir District of The Amhara Region, Ethiopia. *Eastern Africa Social Science Research Review*, 34(1), 109-136.
- Bermudez, L., and Matuszeski, J. (2010). Ensuring continued success: Saving for Change in older program areas of Mali. *Oxfam America*, Boston, MA.

Bhatia, J., Jareer, N., and McIntosh, R. (2018). Community-driven development in Afghanistan: A case study of the national solidarity programme in Wardak. *Asian Survey*, 58(6), 1042-1065.

Blattman, C., Fiala, N., and Martinez, S. (2013). The economic and social returns to cash transfers: evidence from a Ugandan aid program. Columbia University, Departments of Political Science and International and Public Affairs.

Blattman, C., Green, E. P., Jamison, J., Lehmann, M. C., and Annan, J. (2016). The returns to microenterprise support among the ultrapoor: A field experiment in postwar Uganda. *American economic journal: Applied economics*, 8(2), 35-64.

Blattman, C., Green, E., Annan, J., and Jamison, J. (2013). Building women's economic and social empowerment through enterprise: an experimental assessment of the women's income generating support program in Uganda.

Boesen, I. W. (2004). From subjects to citizens: Local Participation in the national solidarity programme. Kabul: Afghanistan Research and Evaluation Unit.

BOMA (2018). An Analysis of the Endline Survey for the REAP Cohort Funded Through the Bill and Melinda Gates Foundation.

Bonilla, J., Zarzur, R. C., Handa, S., Nowlin, C., Peterman, A., Ring, H., ... and Team, Z. C. G. P. E. (2017). Cash for women's empowerment? A mixed-methods evaluation of the government of Zambia's child grant program. *World Development*, 95, 55-72.

Brady, M., Assaad, R., Ibrahim, B. L., Salem, A., Salem, R., and Zibani, N. (2007). Providing new opportunities to adolescent girls in socially conservative settings: The Ishraq program in rural Upper Egypt—full report.

Buchmann, N., Field, E., Glennerster, R., Nazneen, S., Pimkina, S., and Sen, I. (2016). The effect of conditional incentives and a girls' empowerment curriculum on adolescent marriage, childbearing and education in rural Bangladesh: a community clustered randomized controlled trial. Working Paper.

Bureau of Applied Research In Anthropology and IPA (2008). Operational Evaluation of Saving for Change in Mali.

Chatterji, S., Stern, E., Dunkle, K., and Heise, L. (2020). Community activism as a strategy to reduce intimate partner violence (IPV) in rural Rwanda: Results of a community randomised trial. *Journal of global health*, 10(1).

Cherewick, M., and Glass, N. (2018). Caregiver and community insights on coping strategies used by adolescents living in conflict-affected communities. *Global public health*, 13(9), 1322-1336.

Cilliers, J., Dube, O., and Siddiqi, B. (2016). Reconciling after civil conflict increases social capital but decreases individual well-being. *Science*, 352(6287), 787-794.

Concern Worldwide (2017). Final Evaluation Report RAIN+ Project Concern Worldwide Zambia.

- Deubel, T. F., and Boyer, M. (2019). Saving for Change in Mali: From women's financial inclusion to public engagement. Oxfam Research Backgrounder series.
- Deubel, T. F., and Boyer, M. (2020). Women's Economic Empowerment and Political Accountability in Mali. IDS.
- Diwan, F., Makana, G., McKenzie, D., and Paruzzolo, S. (2014). Invitation choice structure has no impact on attendance in a female business training program in Kenya. *Plos one*, 9(10), e109873.
- Duby, Z., Zulu, C. N., and Austrian, K. (2016). Adolescent Girls Empowerment Programme in Zambia: Qualitative Evaluation Report.
- Durr-E-Nayab, and Farooq, S. (2014). Effectiveness of cash transfer programmes for household welfare in Pakistan: The case of the Benazir Income Support Programme. *The Pakistan Development Review*, 53(2), 145-174.
- ElDidi, H., El-Enbaby, H., Kassim, Y., Kurdi, S., Petesch, P., Moataz, Y., and Goessinger, K. Y. (2018). Impact evaluation study for Egypt's Takaful and Karama cash transfer program: Part 2: Qualitative Report (Vol. 15). Intl Food Policy Res Inst.
- El-Enbaby, H., Gilligan, D., Karachiwalla, N., Kassim, Y., and Kurdi, S. (2019). Cash transfers and women's control over decision-making and labor supply in Egypt (Vol. 25). Intl Food Policy Res Inst.
- Emran, M. S., Robano, V., and Smith, S. C. (2014). Assessing the frontiers of ultrapoverty reduction: evidence from challenging the frontiers of poverty reduction/targeting the ultra-poor, an innovative program in Bangladesh. *Economic Development and Cultural Change*, 62(2), 339-380.
- Falb, K. L., Tanner, S., Ward, L., Erksine, D., Noble, E., Assazenew, A., ... and Stark, L. (2016). Creating opportunities through mentorship, parental involvement, and safe spaces (COMPASS) program: multi-country study protocol to protect girls from violence in humanitarian settings. *BMC public health*, 16(1), 1-10.
- Gelayay, M. D., and Lecoutere, E. (2019). Did conditional cash transfers in the Productive Safety Net Program empower women in Tigray, north-east Ethiopia?. IOB, Institute of Development Policy, University of Antwerp.
- Gibbs, A., Corboz, J., and Jewkes, R. (2018). Factors associated with recent intimate partner violence experience amongst currently married women in Afghanistan and health impacts of IPV: a cross sectional study. *BMC Public Health*, 18(1), 1-10.
- Gibbs, A., Corboz, J., Chirwa, E., Mann, C., Karim, F., Shafiq, M., ... and Jewkes, R. (2020). The impacts of combined social and economic empowerment training on intimate partner violence, depression, gender norms and livelihoods among women: an individually randomised controlled trial and qualitative study in Afghanistan. *BMJ global health*, 5(3), e001946.

Gibbs, A., Jewkes, R., Karim, F., Marofi, F., and Corboz, J. (2018). Understanding how Afghan women utilise a gender transformative and economic empowerment intervention: a qualitative study. *Global public health*, 13(11), 1702-1712.

Glass, N., Ramazani, P., Tosha, M., Mpanano, M., and Cinyabuguma, M. (2012). A Congolese–US participatory action research partnership to rebuild the lives of rape survivors and their families in eastern Democratic Republic of Congo. *Global public health*, 7(2), 184-195.

Gobin, V. J., and Santos, P. All together now: the impact of a multifaceted approach to poverty alleviation.

Gobin, V. J., Santos, P., and Toth, R. (2016). Poverty graduation with cash transfers: a randomized evaluation. Department of Economics Discussion Paper, 23, 16.

Gobin, V. J., Santos, P., and Toth, R. (2017). No longer trapped? Promoting entrepreneurship through cash transfers to ultra-poor women in northern Kenya. *American Journal of Agricultural Economics*, 99(5), 1362-1383.

Green, E. P., Blattman, C., Jamison, J., and Annan, J. (2015). Women's entrepreneurship and intimate partner violence: A cluster randomized trial of microenterprise assistance and partner participation in post-conflict Uganda (SSM-D-14-01580R1). *Social science and medicine*, 133, 177-188.

Handa, S., Tembo, G., Natali, L., Angeles, G., and Spektor, G. (2019). In search of the holy grail: Can unconditional cash transfers graduate households out of poverty in Zambia?.

Huber, M. S. and Zupancic, M. (2015). Women for Women International's "Stronger Women, Stronger Afghanistan" Programme.

Humphreys, M., De La Sierra, R. S., and Van der Windt, P. (2012). Social and Economic Impacts of Tuungane Final Report on the Effects of a Community Driven Reconstruction Program in Eastern Democratic REPUBLIC of Congo.

IPA. (2013). Final impact evaluation of the Saving for Change program in Mali.

Ismayilova, L., Karimli, L., Gaveras, E., Tô-Camier, A., Sanson, J., Chaffin, J., and Nanema, R. (2018). An integrated approach to increasing women's empowerment status and reducing domestic violence: Results of a cluster-randomized controlled trial in a West African country. *Psychology of violence*, 8(4), 448.

Jewkes, R., Corboz, J., and Gibbs, A. (2018). Trauma exposure and IPV experienced by Afghan women: analysis of the baseline of a randomised controlled trial. *PLoS one*, 13(10), e0201974.

Karimli, L., Bose, B., and Kagotho, N. (2020). Integrated Graduation Program and its Effect on Women and Household Economic Well-being: Findings from a Randomised Controlled Trial in Burkina Faso. *The Journal of Development Studies*, 56(7), 1277-1294.

Kohli, A., Perrin, N. A., Mpanano, R. M., Mullany, L. C., Murhula, C. M., Binkurhorhwa, A. K., ... and Glass, N. (2014). Risk for family rejection and associated mental health outcomes among conflict-affected adult women living in rural eastern Democratic Republic of the Congo. *Health care for women international*, 35(7-9), 789-807.

Komorowska, D. (2016). *Citizen Voice in Afghanistan: Evaluation of National Solidarity Programme III*.

Kumar, N., Nguyen, P. H., Harris, J., Harvey, D., Rawat, R., and Ruel, M. T. (2018). What it takes: evidence from a nutrition-and gender-sensitive agriculture intervention in rural Zambia. *Journal of Development Effectiveness*, 10(3), 341-372.

Kyegombe, N., Starmann, E., Devries, K. M., Michau, L., Nakuti, J., Musuya, T., ... and Heise, L. (2014). 'SASA! is the medicine that treats violence'. Qualitative findings on how a community mobilisation intervention to prevent violence against women created change in Kampala, Uganda. *Global health action*, 7(1), 25082.

Laudati, A., Mvukiyehe, E., and van der Windt, P. (2018). Participatory Development in Fragile and Conflict-Affected Contexts: An Impact Evaluation of the Tuungane 1 Program in the Democratic Republic of the Congo.

Lecoutere, E., and Wuyts, E. (2020). Confronting the Wall of Patriarchy: Does Participatory Intrahousehold Decision Making Empower Women in Agricultural Households?. *The Journal of Development Studies*, 1-24.

Leight, J., Deyessa, N., and Sharma, V. (2021). Cost-effectiveness analysis of an intimate partner violence prevention intervention targeting men, women and couples in rural Ethiopia: evidence from the Unite for a Better Life randomised controlled trial. *BMJ open*, 11(3), e042365.

Leight, J., Deyessa, N., Verani, F., Tewolde, S., and Sharma, V. (2021). Community-Level spillover effects of an intervention to prevent intimate partner violence and HIV transmission in rural Ethiopia. *BMJ global health*, 6(1), e004075.

Mercy Corps (2015). *Improving child and maternal health: why adolescent girl programming matters post intervention: Evidence from Niger*.

Messersmith, L. J., Halim, N., Kawemama, P., Steven, E., and Reich, N. (2017). A mixed methods study to test the preliminary effect of World Education's Together to End Violence Against Women (TEVAW), a program to address intimate partner violence in northern Tanzania: Baseline report.

Mogford, E., Irby, C. A., and Das, A. (2015). *Changing Men to Change Gender: Combatting Hegemonic Masculinity through Antiviolence Activism in Northern India*. *International Journal of Sociology of the Family*, 71-93.

Nair, S. (2012). *ICTs and Economic Empowerment of Women: Evaluation of SEWA's ICT Activities*.

Olney, D. K., Behrman, J. A., Iruhiriye, E., van den Bold, M., and Pedehombga, A. (2015). Helen Keller International's enhanced homestead food production program in Burkina Faso: Results from a process evaluation.

Pierotti, R. S., Lake, M., and Lewis, C. (2018). Equality on his terms: Doing and undoing gender through men's discussion groups. *Gender and Society*, 32(4), 540-562.

Raising Voices, LSHTM and CEDOVIP (2015). Is Violence Against Women Preventable? Findings from the SASA! Study summarized for general audiences. Kampala, Uganda: Raising Voices.

Ringler, K. (2009). A Review of the Ishraq Program's Quasi-Experimental Impact Evaluation (Master's thesis, University of Minnesota, Minneapolis, United States). Retrieved at: <https://conservancy.umn.edu/handle/11299/50227>.

Saeed, M. K., and Hayat, M. A. (2020). The Impact of Social Cash Transfers on Poverty in Pakistan-A Case Study of Benazir Income Support Programme.

Sarkar, R. K., and Karim, K. M. (2018). Women's Participation in the Development Program and Household Decision Making Status among the Targeting Ultra Poor (TUP) in Rural Bangladesh. *Antrocom: Online Journal of Anthropology*, 14(1).

Seidenfeld, D., Handa, S., and Tembo, G. (2013). Social cash transfer Scheme: 24-month impact report for the child grant programme. American Institutes for Research. Retrieved at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/304213/Zambia-Child-Grant-Prog-24-Month-Impact-Report1.pdf.

Sharma, V., Leight, J., Verani, F., Tewolde, S., and Deyessa, N. (2020). Effectiveness of a culturally appropriate intervention to prevent intimate partner violence and HIV transmission among men, women, and couples in rural Ethiopia: Findings from a cluster-randomized controlled trial. *PLoS medicine*, 17(8), e1003274.

Stangl, A., Farley, K., Sievwright, K., Brady, L., and Fritz, K. (2015). Enhancing Women's Entrepreneurship in Kenya: Initial Qualitative Assessment of the ILO's GET Ahead Business Training Programme. ILO.

Starmann, E., Collumbien, M., Kyegombe, N., Devries, K., Michau, L., Musuya, T., ... and Heise, L. (2017). Exploring couples' processes of change in the context of SASA!, a violence against women and HIV prevention intervention in Uganda. *Prevention science*, 18(2), 233-244.

Stern, E., and Niyibizi, L. L. (2018). Shifting perceptions of consequences of IPV among beneficiaries of Indashyikirwa: an IPV prevention program in Rwanda. *Journal of interpersonal violence*, 33(11), 1778-1804.

Stern, E., and Nyiratunga, R. (2017). A process review of the Indashyikirwa couples curriculum to prevent intimate partner violence and support healthy, equitable relationships in Rwanda. *Social sciences*, 6(2), 63.

Stern, E., Heise, L., and Cislighi, B. (2020). Lessons learnt from engaging opinion leaders to address intimate partner violence in Rwanda. *Development in Practice*, 1-13.

Stern, E., Heise, L., Dunkle, K., and Chatterji, S. (2020). How the Indashyikirwa Intimate Partner Violence Prevention Programme in Rwanda Influenced Parenting and Violence against Children. *Journal of Family Violence*, 1-12.

Subah-Belleh Associates (2010). Economic Empowerment of Adolescent Girls and Young Women Project (EPAG) Exit Poll Report.

Tanner, S., and O'Connor, M. (2017). *A safe place to shine: creating opportunities and raising voices of adolescent girls in humanitarian settings*. New York: International Rescue Committee.

Vaillant, J., Koussoubé, E., Roth, D., Pierotti, R., Hossain, M., and Falb, K. L. (2020). Engaging men to transform inequitable gender attitudes and prevent intimate partner violence: a cluster randomised controlled trial in North and South Kivu, Democratic Republic of Congo. *BMJ global health*, 5(5), e002223.

van den Bold, M., Bliznashka, L., Ramani, G., Olney, D., Quisumbing, A., Pedehombga, A., and Ouedraogo, M. (2021). Nutrition-sensitive agriculture programme impacts on time use and associations with nutrition outcomes. *Maternal and Child Nutrition*, 17(2), e13104.

Van den Bold, M., Pedehombga, A., Ouedraogo, M., Quisumbing, A. R., and Olney, D. (2013). Can integrated agriculture-nutrition programs change gender norms on land and asset ownership? Evidence from Burkina Faso.

Vigneri, M., and Lombardini, S. (2017). Women's Empowerment in Mali: Impact evaluation of the educational project: 'Girls CAN-Promoting Secondary Education in West Africa'.

Wood, S. N. (2020). "He Tells You Your Work Is to Give Birth": Reproductive Coercion and Covert Use of Contraception Among Female Intimate Partner Violence Survivors in Nairobi, Kenya (Doctoral dissertation, Johns Hopkins University).

Zaki, E. M. (2018). Cash transfers and state-citizen relation in Egypt: Takaful and Karama in a development context (Master's thesis, American University in Cairo, Cairo, Egypt). Retrieved from:
<http://dar.aucegypt.edu/bitstream/handle/10526/5277/Cash%20Transfers%20and%20State%20citizen%20relation%20in%20Egypt.pdf?sequence=1>

Ongoing studies

A Cluster Randomised Controlled Trial To Evaluate The Impact Of Micronance And Participatory Gender Training On Intimate Partner Violence. (2019). RIDIE. <http://ridie.org>

Austrian, K., Muthengi, E., Mumah, J., Soler-Hampejsek, E., Kabiru, C. W., Abuya, B., and Maluccio, J. A. (2016). The Adolescent Girls Initiative-kenya (AGI-K): Study Protocol. *BMC Public Health*, 16(1), 210. <https://doi.org/10.1186/s12889-016-2888-1>

Bapolisi, W. A., Ferrari, G., Blampain, C., Makelele, J., Kono-Tange, L., Bisimwa, G., and Merten, S. (2020). Impact Of A Complex Gender-transformative Intervention On Maternal And Child Health Outcomes In The Eastern Democratic Republic Of Congo: Protocol Of A Longitudinal Parallel Mixed-methods Study. *BMC Public Health*, 20(1), 51. <https://doi.org/10.1186/s12889-019-8084-3>

Cherchi, L., Goldstein, M., Habyarimana, J., Montalvao, J., O'Sullivan, M., Udry, C., and Gruver, A. (2019). Empowering Women Through Equal Land Rights: Experimental Evidence From Rural Uganda (Policy Brief No. 33). Gender Innovation Lab. <https://doi.org/10.1596/31513>

Clark, C. J., Spencer, R. A., Shrestha, B., Ferguson, G., Oakes, J. M., and Gupta, J. (2017). Evaluating A Multicomponent Social Behaviour Change Communication Strategy To Reduce Intimate Partner Violence Among Married Couples: Study Protocol For A Cluster Randomized Trial In Nepal. *BMC Public Health*, 17(1), 75. <https://doi.org/10.1186/s12889-016-3909-9>

Constantine Manda, Donald Green, Dylan Groves, Rachel Jones, Bardia Rahmani, and Beatrice Montano. (2018). Mass Media Experiments to Reduce Violence Against Women in Tanzania. *Innovations for Poverty Action*.

Daruwalla, N., Machchhar, U., Pantvaidya, S., D'Souza, V., Gram, L., Copas, A., and Osrin, D. (2019). Community Interventions To Prevent Violence Against Women And Girls In Informal Settlements In Mumbai: The SNEHA-TARA Pragmatic Cluster Randomised Controlled Trial. *Trials*, 20(1), 743. <https://doi.org/10.1186/s13063-019-3817-2>

Empowering Adolescent Girls In The Sahel: Evidence From A Multi-country RCT Of The Sahel Women Empowerment And Demographic Dividend Project. (2020). RIDIE. <http://ridie.org>

Enhancing The Economic Impact Of Rural Roads On Women In Nicaragua. (2020). RIDIE. <http://ridie.org>

Falb, K. L., Tanner, S., Ward, L., Erksine, D., Noble, E., Assazenew, A., Bakomere, T., Graybill, E., Lowry, C., Mallinga, P., Neiman, A., Poulton, C., Robinette, K., Sommer, M., and Stark, L. (2016). Creating Opportunities Through Mentorship, Parental Involvement, And Safe Spaces (COMPASS) Program: Multi-country Study Protocol To Protect Girls From Violence In Humanitarian Settings. *BMC Public Health*, 16(1), 231. <https://doi.org/10.1186/s12889-016-2894-3>

Gibbs, A., Corboz, J., Shafiq, M., Marofi, F., Mecagni, A., Mann, C., Karim, F., Chirwa, E., Maxwell-Jones, C., and Jewkes, R. (2018a). An Individually Randomized Controlled Trial To Determine The Effectiveness Of The Women For Women International Programme In Reducing Intimate Partner Violence And Strengthening Livelihoods Amongst Women In Afghanistan: Trial Design, Methods And Baseline Findings. *BMC Public Health*, 18(1), 164. <https://doi.org/10.1186/s12889-018-5029-1>

Gibbs, A., Corboz, J., Shafiq, M., Marofi, F., Mecagni, A., Mann, C., Karim, F., Chirwa, E., Maxwell-Jones, C., and Jewkes, R. (2018b). An Individually Randomized Controlled Trial To Determine The Effectiveness Of The Women For Women International Programme In Reducing Intimate Partner Violence And Strengthening Livelihoods Amongst Women In Afghanistan: Trial Design, Methods And Baseline Findings. *BMC Public Health*, 18(1), 164. <https://doi.org/10.1186/s12889-018-5029-1>

Girl Empower: Studying The Impact Of Mentorship, Asset Building, Caregiver Discussion Groups, And Cash Transfers On Reducing Girl's Vulnerability To Sexual Exploitation And Abuse. (2017). RIDIE. <http://ridie.org>

Glass, N., Perrin, N., Clough, A., Desgropes, A., Kaburu, F. N., Melton, J., Rink, A., Read-Hamilton, S., and Marsh, M. (2018a). Evaluating The Communities Care Program: Best Practice For Rigorous Research To Evaluate Gender Based Violence Prevention And Response Programs In Humanitarian Settings. *Conflict and Health*, 12(1), 5. <https://doi.org/10.1186/s13031-018-0138-0>

Glass, N., Perrin, N., Clough, A., Desgropes, A., Kaburu, F. N., Melton, J., Rink, A., Read-Hamilton, S., and Marsh, M. (2018b). Evaluating The Communities Care Program: Best Practice For Rigorous Research To Evaluate Gender Based Violence Prevention And Response Programs In Humanitarian Settings. *Conflict and Health*, 12(1), 5. <https://doi.org/10.1186/s13031-018-0138-0>

Glass, N., Perrin, N., Marsh, M., Clough, A., Desgropes, A., Kaburu, F., Ross, B., and Read-Hamilton, S. (2019). Effectiveness Of The Communities Care Programme On Change In Social Norms Associated With Gender-based Violence (GBV) With Residents In Intervention Compared With Control Districts In Mogadishu, Somalia. *BMJ Open*, 9(3), e023819. <https://doi.org/10.1136/bmjopen-2018-023819>

Hallman, K., Kelvin, E., Ozler, B., Seban, J., Kuhlik, E., Alton, C., Kamara, J., and Goodman, S. (2016). Combining Mentoring Programs With Cash Transfers For Adolescent Girls In Liberia: Baseline Report (No. 7797; Policy Research Working Paper). World Bank, Washington, DC. <https://doi.org/10.1596/1813-9450-7797>

Hartmann, M. A., Datta, S., Banay, R. F., Caetano, V., Floreak, R., Appaiah, P., Sreevastha, A., Thomas, S., Selvam, S., Barnette, Q., and Srinivasan, K. (2018). Designing A Pilot Study Protocol To Test A Male Alcohol Use And Intimate Partner Violence Reduction Intervention In India: Beautiful Home. *Frontiers in Public Health*, 6, 218. <https://doi.org/10.3389/fpubh.2018.00218>

Harvey, S., Lees, S., Mshana, G., Pilger, D., Hansen, C., Kapiga, S., and Watts, C. (2018). A Cluster Randomized Controlled Trial To Assess The Impact On Intimate Partner Violence Of A 10-session Participatory Gender Training Curriculum Delivered To Women Taking Part In A Group-based Microfinance Loan Scheme In Tanzania (MAISHA CRT01): Study Protocol. *BMC Women's Health*, 18(1), 55. <https://doi.org/10.1186/s12905-018-0546-8>

Hewett, P. C., Austrian, K., Soler-Hampejsek, E., Behrman, J. R., Bozzani, F., and Jackson-Hachonda, N. A. (2017). Cluster Randomized Evaluation Of Adolescent Girls Empowerment Programme (AGEP): Study Protocol. *BMC Public Health*, 17(1), 386. <https://doi.org/10.1186/s12889-017-4280-1>

Impact Evaluation Of Care's Inspiring Married Adolescent Girls To Imagine New Empowered Futures (IMAGINE) Project. (2019). [RIDIE]. <http://ridie.org>

Jeannie Annan, Christopher Boyer, and Günther Fink. (n.d.). Modern Man Challenge: Evaluating a Men-Focused Intervention to Prevent Intimate Partner Violence in Liberia. Innovations for Poverty Action.

Ketema, T., Bastian, G., Gras, O., Abro, Z., Manchester, K., and Carranza, E. (2015). Ethiopia Women Agribusiness Leaders Network Impact Evaluation: Baseline Survey Report. World Bank. <https://doi.org/10.1596/30314>

Khobarkar, V., Ingole, D. N., and Nage, G. V. (2016). Performance Of Self-help Groups In Micro Finance. *Economic Affairs*, 61(4), 609. <https://doi.org/10.5958/0976-4666.2016.00075.9>

Krishnan, S., Subbiah, K., Chandra, P., and Srinivasan, K. (2012). Minimizing Risks And Monitoring Safety Of An Antenatal Care Intervention To Mitigate Domestic Violence Among Young Indian Women: The Dil Mil Trial. *BMC Public Health*, 12(1), 943. <https://doi.org/10.1186/1471-2458-12-943>

Lerva, B. (n.d.). Power Dynamics In Traditional Households (AEARCTR-0001806) [Data set]. American Economic Association. <https://doi.org/10.1257/rct.1806-1.0>

Leveraging Youth Clubs For Women's Empowerment In Niger | The Abdul Latif Jameel Poverty Action Lab. (2018). The Abdul Latif Jameel Poverty Action Lab (J-PAL). <https://www.povertyactionlab.org/evaluation/leveraging-youth-clubs-womens-empowerment-niger>

Mahmood, S. (2011). Microfinance And Women Entrepreneurs In Pakistan. *International Journal of Gender and Entrepreneurship*, 3(3), 265–274. <https://doi.org/10.1108/17566261111169340>

McKenzie, D., and Weber, M. (n.d.). The results of a pilot financial literacy and business planning training program for women in Uganda (No. 8; Finance and PSD Impact, p. 2). World Bank.

Patson Gondwe. (2020). Enabling Sustainable Graduation out of Poverty for the Extreme Poor in Malawi Umodzi Cohort 2: Successes and Learning Points. Concern.

Peter Carroll, Flora Myamba, Daniel Nielson, Joseph Price, Phillip Roessler, and Wayne Sandholtz. (n.d.). Fast Cash: The Impact of Up-Front Incentives on Young Women's Savings in Tanzania. Innovations for Poverty Action.

Saeed Ali, T., Karmaliani, R., Mcfarlane, J., Khuwaja, H. M. A., Somani, Y., Chirwa, E. D., and Jewkes, R. (2017). Attitude Towards Gender Roles And Violence Against Women And Girls (VAWG): Baseline Findings From An RCT Of 1752 Youths In Pakistan. *Global Health Action*, 10(1), 1342454. <https://doi.org/10.1080/16549716.2017.1342454>

Sahel Women's Empowerment And Demographic Dividend (SWEDD): Empowering Adolescent Girls In Mali. (2020). RIDIE. <http://ridie.org>

Sahel Women's Empowerment And Demographic Dividend (SWEDD): Empowering Adolescent Girls Through Safe Spaces And Accompanying Measures In Burkina Faso. (2020). RIDIE. <http://ridie.org>

Sahel Women's Empowerment And Demographic Dividend (SWEDD): Empowering Adolescent Girls Through Safe Spaces And Accompanying Measures In Cote D'ivoire. (2020). RIDIE. <http://ridie.org>

Sarnquist, C., Kang, J. L., Amuyunzu-Nyamongo, M., Oguda, G., Otieno, D., Mboya, B., Omondi, N., Kipkirui, D., and Baiocchi, M. (2019). A Protocol For A Cluster-randomized Controlled Trial Testing An Empowerment Intervention To Prevent Sexual Assault In Upper Primary School Adolescents In The Informal Settlements Of Nairobi, Kenya. *BMC Public Health*, 19(1), 834. <https://doi.org/10.1186/s12889-019-7154-x>

Sister Of Success Mentoring And Girls Groups In Liberia. (2014). RIDIE. <http://ridie.org>

Vaillant, J. (n.d.). Preventing Intimate-partner Violence: Impact Evaluation Of Engaging Men Through Accountable Practice In Eastern DRC (AEARCTR-0001218) [Data set]. American Economic Association. <https://doi.org/10.1257/rct.1218-4.0>

Women Agribusiness Leaders Network (WALN). (2018). RIDIE. <http://ridie.org>

Zupancic, Mateja. (2017). Knowledge, Attitudes And Practices On Violence And Harmful Practices Against Children In Afghanistan A Baseline Study. Save the Children.

Other references

Protocol

Aall, P., & Crocker, C. A. (2019). Building Resilience and Social Cohesion in Conflict. *Global Policy*, 10(S2), 68–75. <https://doi.org/10.1111/1758-5899.12681>

Abt, T., & Winship, C. (2016). What works in reducing community violence: A meta-review and field study for the northern triangle (p. 53). USAID. <https://www.usaid.gov/sites/default/files/USAID-2016-What-Works-in-Reducing-Community-Violence-Final-Report.pdf>

Agence Francaise de Developpement. (n.d.). Juillet 2020: L'AFD soutient 27 nouveaux projets d'ONG françaises. Retrieved November 27, 2020, from <https://www.afd.fr/fr/actualites/juillet-2020-l-afd-soutient-27-nouveaux-projets-ong-francaises>

Balarin, M., chinen, marjorie, Hoop, T., sennett, josh, & Alcázar, L. (2017). Vocational and business training to improve women's labour market outcomes in low- and middle-income countries: A systematic review. *Campbell Systematic Reviews*, 13. <https://doi.org/10.4073/csr.2017.16>

Bangpan, M., Felix, L., Chiumento, A., & Dickson, K. (2016). The impact of mental health and psychosocial support programmes for populations affected by humanitarian emergencies: A systematic review protocol. Oxfam. <https://doi.org/10.21201/2016.605150>

Barnett, M., Kim, H., O'Donnell, M., & Sitea, L. (2007). Peacebuilding: What Is in a Name? *Global Governance: A Review of Multilateralism and International Organizations*, 13(1), 35–58. <https://doi.org/10.1163/19426720-01301004>

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 50(4), 1088–1101.

Berghof Foundation. (2019). Human security. Berghof Foundation. <https://berghof-foundation.org/themes/human-security>

Bigio, J., & Vogelsteing, R. (2016). How women's participation in conflict prevention and resolution advances US Interests (p. 49). Council for Foreign Relations. https://cdn.cfr.org/sites/default/files/pdf/2016/10/Discussion_Paper_Bigio_Vogelstein_Women%20in%20CPR_OR.pdf

Björkdahl, A., & Höglund, K. (2013). Precarious peacebuilding: Friction in global–local encounters. *Peacebuilding*, 1(3), 289–299. <https://doi.org/10.1080/21647259.2013.813170>

Blair, R., Grossman, G., & Wilke, A. (2020). Community Policing in Uganda. <https://osf.io/df3jk>

Blattman, C., Green, D., Columbia University, Ortega, D., Development Bank of Latin America (CAF), Tobón, S., & University of Chicago and Innovations for Poverty Action. (2018). Hotspot interventions at scale: The effects of policing and city services on crime in Bogotá, Colombia (2018th ed.). International Initiative for Impact Evaluation (3ie). <https://doi.org/10.23846/DPW1IE88>

Blattman, C., Hartman, A., & Blair, R. (2011). Can we Teach Peace and Conflict Resolution?: Results from a randomized evaluation of the Community Empowerment Program (CEP) in Liberia: A Program to Build Peace, Human Rights, and Civic Participation. 2, 38.

Blattman, C., & Tobon, S. (n.d.). Contesting Criminal Gang Governance in Medellín: The Impacts of Intensive Municipal Governance and Community Organization on Gang Control and Governing of Neighborhoods [Data set]. American Economic Association. <https://doi.org/10.1257/rct.2622-1.0>

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to Meta-Analysis. John Wiley and Sons. <https://doi.org/10.1002/9780470743386>

Bourey, C., Bernstein, E., & Stephenson, R. (2015). Systematic review of structural interventions for intimate partner violence in low- and middle-income countries: Organizing evidence for prevention. *BMC Public Health*, 15. <https://doi.org/10.1186/s12889-015-2460-4>

Bouta, T., Frerks, G., & Bannon, I. (2004). Gender, Conflict, and Development. The World Bank. <https://doi.org/10.1596/0-8213-5968-1>

Brody, C., Hoop, T. de, Vojtkova, M., Warnock, R., Dunbar, M., Murthy, P., & Dworkin, S. L. (2017). Can self-help group programs improve women's empowerment? A systematic review. *Journal of Development Effectiveness*, 9(1), 15–40. <https://doi.org/10.1080/19439342.2016.1206607>

Bronfenbrenner, U. (1979). The ecology of human development: Experiments by nature and design. Harvard University Press.

Buvinic, M., Das Gupta, M., Casabonne, U., & Verwimp, P. (2013). Violent Conflict and Gender Inequality: An Overview. *The World Bank Research Observer*, 28(1), 110–138. <https://doi.org/10.1093/wbro/lks011>

Buvinic, M., Gupta, M. D., Casabonne, U., & Verwimp, P. (2013). Violent Conflict and Gender Inequality: An Overview. 37.

Cameron, D., Brown, A., Mishra, A., Picon, M., Esper, H., Calvo, F., & Peterson, K. (2015). Evidence for peacebuilding: An evidence gap map | 3ie (Evidence Gap Map Report 1; p. 50). International Initiative for Impact Evaluation (3ie). <https://www.3ieimpact.org/evidence-hub/publications/evidence-gap-maps/evidence-peacebuilding-evidence-gap-map>

Campbell Collaboration. (2014). Campbell Collaboration Systematic Reviews: Policies and Guidelines. The Campbell Collaboration. <https://doi.org/10.4073/cpg.2016.1>

Caprioli, M. (2000). Gendered Conflict. *Journal of Peace Research*, 37(1), 51–68. <https://doi.org/10.1177/0022343300037001003>

CASP-Qualitative-Checklist-2018.pdf. (n.d.). Retrieved June 21, 2021, from <https://casp-uk.net/wp-content/uploads/2018/01/CASP-Qualitative-Checklist-2018.pdf>

Chant, S., & Sweetman, C. (2012). Fixing women or fixing the world? ‘Smart economics’, efficiency approaches, and gender equality in development. *Gender & Development*, 20(3), 517–529. <https://doi.org/10.1080/13552074.2012.731812>

Cheldelin, S., & Mutisi, M. (2016). Deconstructing women, peace and security: A critical review of approaches to gender and empowerment. <https://doi.org/10.9770>

Chen, M., Vanek, J., & Heintz, J. (2006). Informality, Gender and Poverty: A Global Picture. *Economic and Political Weekly*, 41, 2131–2139. <https://doi.org/10.2307/4418269>

Cochrane Collaboration. (n.d.). Cochrane Handbook for Systematic Reviews of Interventions. Retrieved November 27, 2020, from <https://handbook-5-1.cochrane.org/>

Coffey, C., International Initiative for Impact Evaluation (3ie), Gallagher, E., International Initiative for Impact Evaluation (3ie), Fenton Villar, P., International Initiative for Impact Evaluation (3ie), Stevenson, J., International Initiative for Impact Evaluation (3ie), Tsoli, S., International Initiative for Impact Evaluation (3ie), Dhanasekar, S., International Initiative for Impact Evaluation (3ie), Eyers, J., & Independent consultant. (2017). State-society relations in low- and middle-income countries: An evidence gap map (2017th ed.). International Initiative for Impact Evaluation (3ie). <https://doi.org/10.23846/EGM007>

Collazos, D., Garcia, E., Mejia, D., Ortega, D., & Tobon, S. (2019). Hot Spots Policing in a High Crime Environment: An Experimental Evaluation in Medellín (SSRN Scholarly Paper ID 3316968). Social Science Research Network. <https://doi.org/10.2139/ssrn.3316968>

Committee on the Elimination of Discrimination against Women. (2013). General recommendation No. 30 on women in conflict prevention, conflict and post conflict situations [Recommendation]. <https://www.ohchr.org/Documents/HRBodies/CEDAW/GComments/CEDAW.C.CG.30.pdf>

COMPare. (n.d.). Tracking switched outcomes in clinical trials. COMPare. Retrieved January 8, 2021, from <http://compare-trials.org>

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). The handbook of research synthesis and meta-analysis 2nd edition. *The Hand. of Res. Synthesis and Meta-Analysis*, 2nd Ed., 1–615.

Cornwall, A. (2000). Missing Men? Reflections on Men, Masculinities and Gender in GAD. *IDS Bulletin*, 31(2), 18–27. <https://doi.org/10.1111/j.1759-5436.2000.mp31002003.x>

Cornwall, A. (2016). Women's Empowerment: What Works? *Journal of International Development*, 28(3), 342–359. <https://doi.org/10.1002/jid.3210>

Council on Foreign Relations. (n.d.). Including Women at the Peace Table Produces Better Outcomes. Council on Foreign Relations. Retrieved November 26, 2020, from <https://www.cfr.org/womens-participation-in-peace-processes>

Council on Foreign Relations. (2020). Including Women at the Peace Table Produces Better Outcomes. Council on Foreign Relations. <https://www.cfr.org/womens-participation-in-peace-processes>

Delkhosh, M., Ardalan, A., Rahimiforoushani, A., Keshtkar, A., Amiri Farahani, L., & Merghati Khoei, E. (2017). Interventions for Prevention of Intimate Partner Violence Against Women in Humanitarian Settings: A Protocol for a Systematic Review. *PLoS Currents*, 9. <https://doi.org/10.1371/currents.dis.f41d45fbdca13babe4ae5be0f9732e75>

Department for International Development. (n.d.-a). DevTracker Project GB-COH-1858644-114433—AAWAZ. Retrieved November 27, 2020, from <https://devtracker.fcdo.gov.uk/projects/GB-COH-1858644-114433>

Department for International Development. (n.d.-b). DevTracker Project GB-GOV-13-FUND--GCRF-AH_S004025_1—GCRF. Retrieved November 27, 2020, from https://devtracker.fcdo.gov.uk/projects/GB-GOV-13-FUND--GCRF-AH_S004025_1

Department for International Development. (n.d.-c). DFID Strategic Vision for Gender Equality: A call to action for her potential. 22.

Department for International Development. (2010). Building Peaceful States and Societies: A DFID Practice Paper (p. 60). DFID. <http://www.gsdrc.org/docs/open/con75.pdf>

DerSimonian, R., & Laird, N. (2015). Meta-Analysis in Clinical Trials Revisited. *Contemporary Clinical Trials*, 45(0 0), 139–145. <https://doi.org/10.1016/j.cct.2015.09.002>

Desmidt, S., & Davis, L. (2019). Rhetoric and real progress on the Women, Peace and Security agenda in Africa.

Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., & Sutton, A. (2005). Synthesising qualitative and quantitative evidence: A review of possible methods. *Journal of Health Services Research & Policy*, 10(1), 45–53. <https://doi.org/10.1177/135581960501000110>

Doherty, J., Sonnenfeld, A., Glandon, D., Kozakiewicz, T., Snilstveit, B., & Sabet, D. (2020). Protocol: The effects of rule of law interventions on justice outcomes: An evidence gap map. 85.

- Doocy, S., & Tappis, H. (2015). PROTOCOL: The effectiveness and efficiency of cash-based approaches in emergencies: A Systematic Review. *Campbell Systematic Reviews*, 11(1), 1–45. <https://doi.org/10.1002/CL2.150>
- Doss, C. (2013). Intrahousehold Bargaining and Resource Allocation in Developing Countries. *The World Bank Research Observer*, 28(1), 52–78.
- Dudouet, V. (2017). From Power Mediation to Dialogue Support? Assessing the European Union's Capabilities for Multi-Track Diplomacy.
- Duvendack, M., Hombrados, J. G., Palmer-Jones, R., & Waddington, H. (2012). Assessing 'what works' in international development: Meta-analysis for sophisticated dummies. *Journal of Development Effectiveness*, 4(3), 456–471. <https://doi.org/10.1080/19439342.2012.710642>
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
- Ellis, P. D. (2010). *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511761676>
- ESCAP. (1999). *Women's Empowerment in the Context of Human Security*. <https://www.un.org/womenwatch/ianwge/collaboration/finalcomm1999.htm>
- Faust, J., Grävingholt, J., & Ziaja, S. (2013). Foreign Aid and the Fragile Consensus on State Fragility. <https://doi.org/10.2139/ssrn.1489936>
- Fisher, Z., & Tipton, E. (2015). robumeta: An R-package for robust variance estimation in meta-analysis. *ArXiv:1503.02220 [Stat]*. <http://arxiv.org/abs/1503.02220>
- Galtung, J. (1996). *Peace by peaceful means: Peace and conflict, development and civilization* (pp. viii, 280). Sage Publications, Inc.
- Gibbs, A., Willan, S., Misselhorn, A., & Mangoma, J. (2012). Combined structural interventions for gender equality and livelihood security: A critical review of the evidence from southern and eastern Africa and the implications for young people. *Journal of the International AIDS Society*, 15 Suppl 1, 1–10. <https://doi.org/10.7448/IAS.15.3.17362>
- Gillespie, L.-A., Gupta, P., Landis, D., Shannon, H., & Williamson, K. (n.d.). The impact of protection interventions on unaccompanied and separated children: A systematic review. 124.
- Global Affairs Canada. (2017). Canada's feminist international assistance policy, #HerVoiceHerChoice (p. 92). Global Affairs Canada. https://www.international.gc.ca/world-monde/assets/pdfs/iap2-eng.pdf?_ga=2.237985280.167804930.1603192729-143277333.1603192729
- Global Affairs Ireland. (n.d.). A Simple Guide to Ireland's NAP on Women, Peace and Security (p. 13). Retrieved November 27, 2020, from <https://www.dfa.ie/media/dfa/ourrolepolicies/peaceandsecurity/WPS-Simple-Guide.pdf>

Goetz, A. M. (2008). Who answers to women? Gender & accountability. United Nations Development Fund for Women.

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., Schünemann, H. J., & GRADE Working Group. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ (Clinical Research Ed.)*, 336(7650), 924–926. <https://doi.org/10.1136/bmj.39489.470347.AD>

Hamilton, C., Naam, N., & Shepherd, L. J. (2020). Twenty Years of Women, Peace and Security National Action Plans: Analysis and Lessons Learned. 36.

Hammerstrøm, K., Wade, A., & Jørgensen, A. M. K. (2010a). Searching for studies: A guide to information retrieval for Campbell Systematic Reviews. *Campbell Systematic Reviews*, 1–75.

Hammerstrøm, K., Wade, A., & Jørgensen, A. M. K. (2010b). Searching for studies: A guide to information retrieval for Campbell Systematic Reviews. *Campbell Systematic Reviews*, 1–75.

Handanagic, S., Barbaric, J., Anglemeyer, A., & Bozicevic, I. (2016). Community mobilization and empowerment interventions for preventing HIV in low and middle income countries: Protocols. In *The Cochrane Library* (Vol. 2016). <https://doi.org/10.1002/14651858.CD011186.pub2>

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, 1(1), 39–65. <https://doi.org/10.1002/jrsm.5>

Hedström, J., Senarathna, T., & International Institute for Democracy and Electoral Assistance (Eds.). (2015). Women in conflict and peace. International IDEA.

Higgins, J. P., Sterne, J. A., Savovic, J., Page, M. J., Hróbjartsson, A., Boutron, I., Reeves, B., & Eldridge, S. (2016). A revised tool for assessing risk of bias in randomized trials. *Cochrane Database of Systematic Reviews*, 10(Suppl 1), 29–31.

Higgins, J. P. T., & Thomas, J. (2020). *Cochrane Handbook for Systematic Reviews of Interventions*. /handbook/current

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539–1558. <https://doi.org/10.1002/sim.1186>

Hultcrantz, M., Rind, D., Akl, E. A., Treweek, S., Mustafa, R. A., Iorio, A., Alper, B. S., Meerpohl, J. J., Murad, M. H., Ansari, M. T., Katikireddi, S. V., Östlund, P., Tranæus, S., Christensen, R., Gartlehner, G., Brozek, J., Izcovich, A., Schünemann, H., & Guyatt, G. (2017). The GRADE Working Group clarifies the construct of certainty of evidence. *Journal of Clinical Epidemiology*, 87, 4–13. <https://doi.org/10.1016/j.jclinepi.2017.05.006>

Institute for Economics and Peace. (2020). Global Peace Index 2020: Measuring Peace in a Complex World. Institute for Economics & Peace. https://visionofhumanity.org/wp-content/uploads/2020/10/GPI_2020_web.pdf

- Jong, S. de, & Kimm, S. (2017). The co-optation of feminisms: A research agenda. *International Feminist Journal of Politics*, 19(2), 185–200.
<https://doi.org/10.1080/14616742.2017.1299582>
- Kabeer, N. (1999). Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment. *Development and Change*, 30(3), 435–464.
<https://doi.org/10.1111/1467-7660.00125>
- Kabeer, N., & Waddington, H. (2015). Economic impacts of conditional cash transfer programmes: A systematic review and meta-analysis. *Journal of Development Effectiveness*, 7(3), 290–303. <https://doi.org/10.1080/19439342.2015.1068833>
- Keef, S. P., & Roberts, L. A. (2004). The meta-analysis of partial effect sizes. *British Journal of Mathematical and Statistical Psychology*, 57(1), 97–129.
<https://doi.org/10.1348/000711004849303>
- Keuleers, P. (2016). Why building peaceful societies is part of the sustainable development agenda. UNDP.
<https://www.undp.org/content/undp/en/home/blog/2016/5/18/Why-building-peaceful-societies-is-part-of-the-sustainable-development-agenda.html>
- King, E. (2013). A Critical Review of Community-Driven Development Programmes in Conflict-Affected Contexts. https://www.academia.edu/13841546/A_Critical_Review_of_Community_Driven_Development_Programmes_in_Conflict_Affected_Contexts
- King, E., Samii, C., & Snilstveit, B. (2010). Interventions to promote social cohesion in sub-Saharan Africa. *Journal of Development Effectiveness*, 2(3), 336–370.
<https://doi.org/10.1080/17449057.2010.504552>
- Kraft, J., Wilkins, K., Morales, G., Widyono, M., & Middlestadt, S. (2014). An Evidence Review of Gender-Integrated Interventions in Reproductive and Maternal-Child Health. *Journal of Health Communication*, 19, 122–141.
<https://doi.org/10.1080/10810730.2014.918216>
- Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A.-M. K., Hammerstrøm, K., & Sathe, N. (2017). Searching for studies: A guide to information retrieval for Campbell systematic reviews. *Campbell Systematic Reviews*, 13(1), 1–73.
<https://doi.org/10.4073/cmg.2016.1>
- Langenkamp, D. A. (n.d.). GIZ Gender Strategy. Gender reloaded: Vision needs Attitude – Attitude meets Action. 20.
- Lewin, S., Booth, A., Glenton, C., Munthe-Kaas, H., Rashidian, A., Wainwright, M., Bohren, M. A., Tunçalp, Ö., Colvin, C. J., Garside, R., Carlsen, B., Langlois, E. V., & Noyes, J. (2018a). Applying GRADE-CERQual to qualitative evidence synthesis findings: Introduction to the series. *Implementation Science*, 13(1), 2.
<https://doi.org/10.1186/s13012-017-0688-3>

Lewin, S., Booth, A., Glenton, C., Munthe-Kaas, H., Rashidian, A., Wainwright, M., Bohren, M. A., Tunçalp, Ö., Colvin, C. J., Garside, R., Carlsen, B., Langlois, E. V., & Noyes, J. (2018b). Applying GRADE-CERQual to qualitative evidence synthesis findings: Introduction to the series. *Implementation Science*, 13(1), 2. <https://doi.org/10.1186/s13012-017-0688-3>

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Sage Publications. <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=63274>

London School of Economics. (2019). LSE-WPS National Action Plans. LSE - Women, Peace Security. <https://www.wpsnaps.org/>

Lynch, U., McGrellis, S., Dutschke, M., Anderson, M., Arnsberger, P., & Macdonald, G. (2013). What is the evidence that the establishment or use of community accountability mechanisms and processes improves inclusive service delivery by governments, donors and NGOs to communities? Social Science Research Unit, Institute of Education, University of London. <https://pure.qub.ac.uk/en/publications/what-is-the-evidence-that-the-establishment-or-use-of-community-a>

Macdonald, G., Higgins, J. P., Ramchandani, P., Valentine, J. C., Bronger, L. P., Klein, P., O'Daniel, R., Pickering, M., Rademaker, B., Richardson, G., & Taylor, M. (2012). Cognitive-Behavioural Interventions for Children Who Have Been Sexually Abused: A Systematic Review. *Campbell Systematic Reviews*, 8(1), 1–111. <https://doi.org/10.4073/csr.2012.14>

McWilliams, M., & Kilmurray, A. (2015). From the global to the local: Grounding UNSCR 1325 on women, peace and security in post conflict policy making. *Women's Studies International Forum*, 51, 128–135. <https://doi.org/10.1016/j.wsif.2014.11.006>

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*.

Ministry of Foreign Affairs. (n.d.). Dutch National Action Plan on Resolution 1325: Taking a stand for women, peace and security. Ministry of Foreign Affairs (Netherlands). Retrieved November 27, 2020, from https://www.peacewomen.org/sites/default/files/dutch_nationalactionplan_december2007.pdf

Montgomery, P., Movsisyan, A., Grant, S. P., Macdonald, G., & Rehfuss, E. A. (2019). Considerations of complexity in rating certainty of evidence in systematic reviews: A primer on using the GRADE approach in global health. *BMJ Global Health*, 4(Suppl 1), e000848. <https://doi.org/10.1136/bmjgh-2018-000848>

Moore, L., Chersich, M. F., Steen, R., Reza-Paul, S., Dhana, A., Vuylsteke, B., Lafort, Y., & Scorgie, F. (2014). Community empowerment and involvement of female sex workers in targeted sexual and reproductive health interventions in Africa: A systematic review. *Globalization and Health*, 10(1), 47. <https://doi.org/10.1186/1744-8603-10-47>

Mundkur, A., & Shepherd, L. (2018, January 23). How (not) to make WPS count. LSE Women, Peace and Security Blog. <https://blogs.lse.ac.uk/wps/2018/01/23/how-not-to-make-wps-count/>

Munro, J. (2000). Gender and Peacebuilding. 26.

Noyes, J., Booth, A., Flemming, K., Harden, A., Harris, J., Garside, R., Hannes, K., Pantoja, T., & Thomas, J. (2020). Cochrane Handbook for Systematic Reviews of Interventions version 6.1 Chapter 21: Qualitative evidence (6.1). Cochrane Collaboration. /handbook/current/chapter-21

O'Driscoll, D. (2017). Women's participation in peacebuilding and reconciliation in Iraq. <https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/13437>

OECD. (2013). Gender and Statebuilding in Fragile and Conflict-affected States (p. 92). OECD. https://read.oecd-ilibrary.org/development/gender-and-statebuilding-in-fragile-and-conflict-affected-states_9789264202061-en

OECD. (2016a). States of Fragility 2016| OECD iLibrary. <https://www.oecd-ilibrary.org/sites/9789264267213-en/index.html?itemId=/content/publication/9789264267213-en>

OECD. (2016b). Entrepreneurship at a Glance 2016. OECD. https://doi.org/10.1787/entrepreneur_aag-2016-en

OECD. (2020). States of Fragility 2020 | OECD iLibrary. https://read.oecd-ilibrary.org/development/states-of-fragility-2020_9789264202061-en/index.html

O'Mara-Eves, A., Thomas, J., McNaught, J., Miwa, M., & Ananiadou, S. (2015a). Using text mining for study identification in systematic reviews: A systematic review of current approaches. *Systematic Reviews*, 4. <https://doi.org/10.1186/2046-4053-4-5>

O'Mara-Eves, A., Thomas, J., McNaught, J., Miwa, M., & Ananiadou, S. (2015b). Using text mining for study identification in systematic reviews: A systematic review of current approaches. *Systematic Reviews*, 4(1), 5. <https://doi.org/10.1186/2046-4053-4-5>

O'Reilly, M. (2013). Issue Brief—Women in Conflict Mediation: Why it Matters. International Peace Institute. https://www.ipinst.org/wp-content/uploads/publications/ipi_e_pub_women_in_conflict_med.pdf

Pankhurst, D. T. (2000). Women, Gender and Peacebuilding [Working Paper]. <https://bradscholars.brad.ac.uk/handle/10454/939>

Peace Women. (n.d.). Action Plan on Transformative Financing for GEWE.pdf. Retrieved December 14, 2020, from <http://www.peacewomen.org/sites/default/files/Action%20Plan%20on%20Transformative%20Financing%20for%20GEWE.pdf>

Peace Women. (2014, November 14). WPS National-Level Implementation. PeaceWomen. <https://www.peacewomen.org/member-states>

Peace Women. (2016, June 23). Women, Peace and Security Financing. PeaceWomen. <http://www.peacewomen.org/WPS-Financing>

Peace Women. (2018, March 7). UN Resolution 1325: Significant But Lacking. PeaceWomen. <http://peacewomen.org/resource/un-resolution-1325-significant-lacking>

Perrin, N., Marsh, M., Clough, A., Desgroppes, A., Yope Phanuel, C., Abdi, A., Kaburu, F., Heitmann, S., Yamashina, M., Ross, B., Read-Hamilton, S., Turner, R., Heise, L., & Glass, N. (2019). Social norms and beliefs about gender based violence scale: A measure for use with gender based violence prevention programs in low-resource and humanitarian settings. *Conflict and Health*, 13(1), 6. <https://doi.org/10.1186/s13031-019-0189-x>

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, 61(10), 991–996. <https://doi.org/10.1016/j.jclinepi.2007.11.010>

Picon, M. G., Rankin, K., International Initiative for Impact Evaluation (3ie), Ludwig, J., International Initiative for Impact Evaluation (3ie), Sabet, S. M., International Initiative for Impact Evaluation (3ie), Delaney, A., Independent consultant, Holst, A., & Independent consultant. (2017). Intimate partner violence prevention: An evidence gap map (2017th ed.). International Initiative for Impact Evaluation. <https://doi.org/10.23846/EGM008>

Pluye et al. (2009). Proposal: A mixed methods appraisal tool for systematic mixed studies reviews. 37th NAPCRG Annual Conference. <https://www.webcitation.org/5tTRTc9yJ>

Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2016). Estimating the Difference Between Published and Unpublished Effect Sizes: A Meta-Review. *Review of Educational Research*, 86(1), 207–236. <https://doi.org/10.3102/0034654315582067>

Purgato, M., Gastaldon, C., Papola, D., van Ommeren, M., Barbui, C., & Tol, W. A. (2018). Psychological therapies for the treatment of mental disorders in low- and middle-income countries affected by humanitarian crises. *The Cochrane Database of Systematic Reviews*, 7, CD011849. <https://doi.org/10.1002/14651858.CD011849.pub2>

Purgato, M., Gross, A. L., Betancourt, T., Bolton, P., Bonetto, C., Gastaldon, C., Gordon, J., O'Callaghan, P., Papola, D., Peltonen, K., Punamaki, R.-L., Richards, J., Staples, J. K., Unterhitzenberger, J., van Ommeren, M., de Jong, J., Jordans, M. J. D., Tol, W. A., & Barbui, C. (2018). Focused psychosocial interventions for children in low-resource humanitarian settings: A systematic review and individual participant data meta-analysis. *The Lancet. Global Health*, 6(4), e390–e400. [https://doi.org/10.1016/S2214-109X\(18\)30046-9](https://doi.org/10.1016/S2214-109X(18)30046-9)

R Core Team. (2020). R: What is R? <https://www.r-project.org/about.html>

R Project. (n.d.). R: The R Project for Statistical Computing. Retrieved January 8, 2021, from <https://www.r-project.org/>

Red nacional de mujeres. (n.d.). Coalición 1325—Coalición 1325. Retrieved November 27, 2020, from <https://www.rednacionaldemujeres.org/index.php/publicaciones/coalicion-1325>

Rihoux, B. (2006). Qualitative Comparative Analysis (QCA) and Related Systematic Comparative Methods. *International Sociology - INT SOCIOL*, 21, 679–706. <https://doi.org/10.1177/0268580906067836>

Rivas, C., Ramsay, J., Sadowski, L., Davidson, L. L., Dunne, D., Eldridge, S., Hegarty, K., Taft, A., & Feder, G. (2015). Advocacy interventions to reduce or eliminate violence and promote the physical and psychosocial well-being of women who experience intimate partner abuse. *The Cochrane Database of Systematic Reviews*, 12, CD005043. <https://doi.org/10.1002/14651858.CD005043.pub3>

Rue, L. D. L., Polanin, J. R., Espelage, D. L., & Pigott, T. D. (2013). PROTOCOL: School-based Interventions to Reduce Dating and Sexual Violence: A Systematic Review. *Campbell Systematic Reviews*, 9(1), 1–43. <https://doi.org/10.1002/CL2.106>

Ruta pacifica de las mujeres. (2016). Plan Estrategico 2016—2020. Ruta pacifica de las mujeres. https://rutapacifica.org.co/wp/wp-content/uploads/2018/06/PlanEstrategico2016_2020.pdf

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, 8(4), 448–467. <https://doi.org/10.1037/1082-989X.8.4.448>

Santesso, N., Carrasco-Labra, A., Langendam, M., Brignardello-Petersen, R., Mustafa, R. A., Heus, P., Lasserson, T., Opiyo, N., Kunnamo, I., Sinclair, D., Garner, P., Treweek, S., Tovey, D., Akl, E. A., Tugwell, P., Brozek, J. L., Guyatt, G., & Schunemann, H. J. (2016). Improving GRADE evidence tables part 3: Detailed guidance for explanatory footnotes supports creating and understanding GRADE certainty in the evidence judgments. *Journal of Clinical Epidemiology*, 74, 28–39. <https://doi.org/10.1016/j.jclinepi.2015.12.006>

Santesso, N., Glenton, C., Dahm, P., Garner, P., Akl, E. A., Alper, B., Brignardello-Petersen, R., Carrasco-Labra, A., De Beer, H., Hultcrantz, M., Kuijpers, T., Meerpohl, J., Morgan, R., Mustafa, R., Skoetz, N., Sultan, S., Wiysonge, C., Guyatt, G., & Schünemann, H. J. (2020). GRADE guidelines 26: Informative statements to communicate the findings of systematic reviews of interventions. *Journal of Clinical Epidemiology*, 119, 126–135. <https://doi.org/10.1016/j.jclinepi.2019.10.014>

Senarathna, T. (2015). Women in peace processes—Real inclusivity or “just add women”? https://www.academia.edu/32694680/Women_in_peace_processes_real_inclusivity_or_just_add_women_

Shemilt, I., Mcdaid, D., Marsh, K., Henderson, C., Bertranou, E., Mallender, J., Drummond, M., Mugford, M., & Vale, L. (2013). Issues in the incorporation of economic perspectives and evidence into Cochrane reviews. *Systematic Reviews*, 2, 83. <https://doi.org/10.1186/2046-4053-2-83>

SIDA. (2015). Gender Toolbox Brief: Women, Peace and Security (Gender Toolbox, p. 4) [Brief]. SIDA. <https://www.sida.se/contentassets/3a820dbd152f4fca98bacde8a8101e15/women-peace-and-security.pdf>

Signorelli, M. C., Hillel, S., de Oliveira, D. C., Ayala Quintanilla, B. P., Hegarty, K., & Taft, A. (2018). Voices from low-income and middle income countries: A systematic review protocol of primary healthcare interventions within public health systems addressing intimate partner violence against women. <https://doi.org/10.1136/bmjopen-2017-019266>

Sigsworth, R., & Kumalo, L. (2016). Women, peace and security: Implementing the Maputo Protocol in Africa. 24.

Snilstveit, B. (2012). Systematic reviews: From 'bare bones' reviews to policy relevance. *Journal of Development Effectiveness*, 4(3), 388–408. <https://doi.org/10.1080/19439342.2012.709875>

Snilstveit, B., & al. (2015). Interventions for improving learning outcomes and access to education in low- and middleincome countries (No. 24; Systematic Review 24, p. 891). 3ie. https://www.3ieimpact.org/sites/default/files/2019-01/SR24-education-review_2.pdf

Sonnenfeld, A., Chirgwin, H., International Initiative for Impact Evaluation (3ie), Berretta, M., International Initiative for Impact Evaluation (3ie), Longman, K., Independent consultant, Krämer, M., German Institute for Development Evaluation (DEval), Snilstveit, B., & International Initiative for Impact Evaluation (3ie). (2020). Building peaceful societies: An evidence gap map (2020th ed.). International Initiative for Impact Evaluation (3ie). <https://doi.org/10.23846/EGM015>

Spangaro, J., Adogu, C., Ranmuthugala, G., Davies, G. P., Steinacker, L., & Zwi, A. (2013). What Evidence Exists for Initiatives to Reduce Risk and Incidence of Sexual Violence in Armed Conflict and Other Humanitarian Crises? A Systematic Review. *PLOS ONE*, 8(5), e62600. <https://doi.org/10.1371/journal.pone.0062600>

Spangaro, J., Zwi, A. B., Adogu, C., Ranmuthugala, G., Davies, G. P., Steinacker, L., University of London, Social Science Research Unit, & Evidence for Policy and Practice Information and Co-ordinating Centre. (2013). What is the evidence of the impact of initiatives to reduce risk and incidence of sexual violence in conflict and post-conflict zones and other humanitarian crises in lower- and middle-income countries?: A systematic review.

Speake, B. (2013, February 11). A Gendered Approach to Peacebuilding and Conflict Resolution. E-International Relations. <https://www.e-ir.info/2013/02/11/a-gendered-approach-to-peacebuilding-and-conflict-resolution/>

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A.-W., Churchill, R., Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., ... Higgins, J. P. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, 355, i4919. <https://doi.org/10.1136/bmj.i4919>

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, 343, d4002. <https://doi.org/10.1136/bmj.d4002>

Strickland, R. (2004). To Have and To Hold: Women's Property and Inheritance Rights in the Context of HIV/AIDS in Sub-Saharan Africa. 84.

Sweetman, C., & Rowlands, J. (2016). Introduction: Working on gender equality in fragile contexts. *Gender & Development*, 24(3), 337–351.
<https://doi.org/10.1080/13552074.2016.1248033>

Tabachnick, B., & Fidell, Linda. S. (2007). *Using Multivariate Statistics*. In Boston: Allyn & Bacon (Vol. 3).

Thomas, J., & Brunton, J. (2010). EPPI-Reviewer 4: Software for Research Synthesis.

Thomas, J., & Harden, A. (2008a). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8(1), 45.
<https://doi.org/10.1186/1471-2288-8-45>

Thomas, J., & Harden, A. (2008b). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8(1), 45.
<https://doi.org/10.1186/1471-2288-8-45>

Thomas, J., McNaught, J., & Ananiadou, S. (2011). Applications of text mining within systematic reviews. *Research Synthesis Methods*, 2(1), 1–14.
<https://doi.org/10.1002/jrsm.27>

Thomas, J., O'Mara-Eves, A., & Brunton, G. (2014). Using qualitative comparative analysis (QCA) in systematic reviews of complex interventions: A worked example. *Systematic Reviews*, 3(1), 67. <https://doi.org/10.1186/2046-4053-3-67>

Tian, N., Fleurant, A., Kuimova, A., Wezeman, P., & Wezeman, S. (2019). Trends in world military expenditure, 2018. 12.

Tol, W. A., Stavrou, V., Greene, M. C., Mergenthaler, C., van Ommeren, M., & García Moreno, C. (2013a). Sexual and gender-based violence in areas of armed conflict: A systematic review of mental health and psychosocial support interventions. *Conflict and Health*, 7(1), 16. <https://doi.org/10.1186/1752-1505-7-16>

Tol, W. A., Stavrou, V., Greene, M. C., Mergenthaler, C., van Ommeren, M., & García Moreno, C. (2013b). Sexual and gender-based violence in areas of armed conflict: A systematic review of mental health and psychosocial support interventions. *Conflict and Health*, 7(1), 16. <https://doi.org/10.1186/1752-1505-7-16>

Tyrer, R. A., & Fazel, M. (2014). School and Community-Based Interventions for Refugee and Asylum Seeking Children: A Systematic Review. *PLOS ONE*, 9(2), e89359.
<https://doi.org/10.1371/journal.pone.0089359>

UN Development Programme. (2019). Parliament as partners supporting women peace and security agenda. UNDP.
<https://www.undp.org/content/undp/en/home/librarypage/crisis-prevention-and-recovery/parliament-as-partners-supporting-women-peace-and-security-agend.html>

UN Women. (n.d.-a). Thematic Area Dashboard | UN Women Data Hub. Retrieved November 27, 2020, from <https://data.unwomen.org/data-portal/vaw-wps>

UN Women. (n.d.-b). Thematic Area Dashboard | UN Women Data Hub. Retrieved November 27, 2020, from <https://data.unwomen.org/data-portal/vaw-wps>

UN Women. (2015). Preventing Conflict, transforming justice, securing the peace.

UN Women. (2020). Facts and figures: Women, peace, and security | What we do. UN Women. <https://www.unwomen.org/en/what-we-do/peace-and-security/facts-and-figures>

UNICEF (Ed.). (2006). Women and children: The double dividend of gender equality. UNICEF.

United Nations (Ed.). (2010). Achieving gender equality, women's empowerment and strengthening development cooperation: Dialogues at the Economic and Social Council. United Nations.

United Nations Human Rights Office of the High Commissioner. (2014). Women's rights are human rights. United Nations. <https://www.ohchr.org/documents/events/whrd/womenrightsarehr.pdf>

United Nations Peacebuilding Support Office. (2010). UN Peacebuilding: An Orientation. United Nations. https://www.un.org/peacebuilding/sites/www.un.org.peacebuilding/files/documents/peacebuilding_orientation.pdf

United Nations Security Council. (2000). Resolution 1325 (2000) [Resolution]. [https://undocs.org/pdf?symbol=en/S/RES/1325\(2000\)](https://undocs.org/pdf?symbol=en/S/RES/1325(2000))

United Nations Security Council. (2016). Resolution 2282 (2016) [Resolution]. <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N16/118/51/PDF/N1611851.pdf?OpenElement>

United Nations Security Council. (2018). Resolution 2419 (2018) [Resolution]. [https://undocs.org/pdf?symbol=en/S/RES/2419\(2018\)](https://undocs.org/pdf?symbol=en/S/RES/2419(2018))

United States Institute for Peace. (2012). Gender, War, and Peacebuilding. United States Institute of Peace. <https://www.usip.org/publications/2012/09/gender-war-and-peacebuilding>

UNOCHA. (2009). Human Security in Theory and Practice: Application of the Human Security Concept and the United Nations Trust Fund for Human Security. <https://www.unocha.org/sites/dms/HSU/Publications%20and%20Products/Human%20Security%20Tools/Human%20Security%20in%20Theory%20and%20Practice%20English.pdf>

USAID. (2007). Women and Conflict: An Introductory guide for programming (p. 29). USAID. https://www.usaid.gov/sites/default/files/documents/1865/toolkit_women_and_conflict_an_introductory_guide_for_programming.pdf

Valentine, J. C., Aloe, A. M., & Lau, T. S. (2015). Life After NHST: How to Describe Your Data Without “p-ing” Everywhere. *Basic and Applied Social Psychology*, 37(5), 260–273. <https://doi.org/10.1080/01973533.2015.1060240>

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, 36(1), 1–48. <https://doi.org/10.18637/jss.v036.i03>

Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1(2), 112–125. <https://doi.org/10.1002/jrsm.11>

Waddington, H., Aloe, A. M., Becker, B. J., Djimeu, E. W., Hombrados, J. G., Tugwell, P., Wells, G., & Reeves, B. (2017). Quasi-experimental study designs series—paper 6: Risk of bias assessment. *Journal of Clinical Epidemiology*, 89, 43–52. <https://doi.org/10.1016/j.jclinepi.2017.02.015>

Waddington, H., Stevenson, J., Sonnenfeld, A., & Gaarder, M. (2018). PROTOCOL: Participation, inclusion, transparency and accountability (PITA) to improve public services in low- and middle-income countries: a systematic review. *Campbell Systematic Reviews*, 14(1), 1–69. <https://doi.org/10.1002/CL2.205>

Waddington, H., White, H., Snilstveit, B., Hombrados, J. G., Vojtkova, M., Davies, P., Bhavsar, A., Eysers, J., Koehlmoos, T. P., Petticrew, M., Valentine, J. C., & Tugwell, P. (2012a). How to do a good systematic review of effects in international development: A tool kit. *Journal of Development Effectiveness*, 4(3), 359–387. <https://doi.org/10.1080/19439342.2012.711765>

Waddington, H., White, H., Snilstveit, B., Hombrados, J. G., Vojtkova, M., Davies, P., Bhavsar, A., Eysers, J., Koehlmoos, T. P., Petticrew, M., Valentine, J. C., & Tugwell, P. (2012b). How to do a good systematic review of effects in international development: A tool kit. *Journal of Development Effectiveness*, 4(3), 359–387. <https://doi.org/10.1080/19439342.2012.711765>

Waddington, H., White, H., Snilstveit, B., Hombrados, J. G., Vojtkova, M., Davies, P., Bhavsar, A., Eysers, J., Koehlmoos, T. P., Petticrew, M., Valentine, J. C., & Tugwell, P. (2012c). How to do a good systematic review of effects in international development: A tool kit. *Journal of Development Effectiveness*, 4(3), 359–387. <https://doi.org/10.1080/19439342.2012.711765>

Warren, E., Post, N., Hossain, M., Blanchet, K., & Roberts, B. (2015). Systematic review of the evidence on the effectiveness of sexual and reproductive health interventions in humanitarian crises. *BMJ Open*, 5, e008226. <https://doi.org/10.1136/bmjopen-2015-008226>

WebCite query result. (n.d.). Retrieved June 21, 2021, from <https://www.webcitation.org/5tTRTc9yJ>

White, H. (2009). Theory-based impact evaluation: Principles and practice. *Journal of Development Effectiveness*, 1(3), 271–284. <https://doi.org/10.1080/19439340903114628>

Wilson, D. B., Weisburd, D., & McClure, D. (2011). Use of DNA testing in police investigative work for increasing offender identification, arrest, conviction and case clearance. *Campbell Systematic Reviews*, 7(1), 1–53. <https://doi.org/10.4073/csr.2011.7>

World Bank. (2020). Classification of Fragile and Conflict-Affected Situations. World Bank. <https://www.worldbank.org/en/topic/fragilityconflictviolence/brief/harmonized-list-of-fragile-situations>

World Health Organisation. (n.d.). WHO Gender Responsive Assessment Scale: Criteria for assessing programmes and policies. World Health Organisation. Retrieved November 26, 2020, from https://www.who.int/gender/mainstreaming/GMH_Participant_GenderAssessmentScale.pdf

Analysis and reporting

Amaral, S., Bhalotra, S., and Prakash, N. (2019). *Gender, Crime and Punishment: Evidence From Women Police Stations In India*. 63.

BMZ, F. M. for E. C. and D. (n.d.-a). *Further Information On Peace And Security*. Federal Ministry for Economic Cooperation and Development. Retrieved November 6, 2020, from http://www.bmz.de/en/issues/state_fragility/dokumente/index.html

BMZ, F. M. for E. C. and D. (n.d.-b). *German And International Activities In Fragile States*. Federal Ministry for Economic Cooperation and Development. Retrieved November 6, 2020, from http://www.bmz.de/en/issues/state_fragility/deutschesengagement/index.html

BMZ, F. M. for E. C. and D. (n.d.-c). *State Fragility – A Challenge For Development Policy*. Federal Ministry for Economic Cooperation and Development. Retrieved November 6, 2020, from http://www.bmz.de/en/issues/state_fragility/index.html

Bouta, T., Frerks, G., and Bannon, I. (2004). *Gender, Conflict, And Development*. The World Bank. <https://doi.org/10.1596/0-8213-5968-1>

Brief_33_Gender_in_conflict.pdf. (n.d.). Retrieved November 5, 2020, from https://www.iss.europa.eu/sites/default/files/EUISSFiles/Brief_33_Gender_in_conflict.pdf

CDA Collaborative Learning. (2019). *Sub-sector Review Of Evidence From Reconciliation Programs* (p. 34). Peacebuilding Evaluation Consortium. <https://www.cdacollaborative.org/wp-content/uploads/2019/08/PEC-Reconciliation-Sector-Review-Final-March-2019-2.pdf>

Classification Of Fragile And Conflict-affected Situations. (n.d.). World Bank. Retrieved November 6, 2020, from <https://www.worldbank.org/en/topic/fragilityconflictviolence/brief/harmonized-list-of-fragile-situations>

Cockburn, C., & Žarkov, D. (2002). *The postwar moment: Militaries, masculinities and international peacekeeping*. London: Lawrence & Wishart. <https://repository.uhn.ru.nl/handle/2066/62653>

Cohn, C. (2017). Beyond the “Women, Peace and Security” Agenda: Why We Need a Feminist Roadmap for Sustainable Peace (p. 18). Consortium on Gender, Security and Human Rights. https://genderandsecurity.org/sites/default/files/Cohn_-_Beyond_the_Women_Peace_and_Security_Agenda_Why_We_Need_a_Feminist_Roadmap_for_Sustainable_Peace.pdf

Das, N., Yasmin, R., Ara, J., Kamruzzaman, M., Davis, P., Behrman, J., Roy, S., and Quisumbing, A. R. (2013). *How Do Intrahousehold Dynamics Change When Assets Are Transferred To Women? Evidence From Brac’s Challenging The Frontiers Of Poverty Reduction — Targeting The Ultra Poor Program In Bangladesh* (SSRN Scholarly Paper ID 2405712). Social Science Research Network. <https://doi.org/10.2139/ssrn.2405712>

Di Pierro, J. (2021). CAPACITY BUILDING: Is It Only A Matter Of Training? *Lear*.
<https://www.learlab.com/insights/capacity-building-is-it-only-a-matter-of-training/>

Dudwick, N., and Kuehnast, K. (n.d.). *Gender and Fragility: Ensuring a Golden Hour*. 8.

FCSList-FY06toFY20.pdf. (n.d.). Retrieved November 6, 2020, from
<http://pubdocs.worldbank.org/en/176001594407411053/FCSList-FY06toFY20.pdf>

FINAL States of Fragility Highlights document.pdf. (n.d.). Retrieved November 6, 2020, from
<http://www.oecd.org/dac/conflict-fragility-resilience/docs/FINAL%20States%20of%20Fragility%20Highlights%20document.pdf>

Fragility Index for a Differentiated Approach. (n.d.). 16.

Gender Development Index (GDI) | Human Development Reports. (n.d.). Retrieved November 6, 2020, from
<http://hdr.undp.org/en/content/gender-development-index-gdi>

Gender In Fragile And Conflict-affected Environments. (n.d.). GSDRC. Retrieved November 5, 2020, from
<https://gsdrc.org/topic-guides/gender/gender-in-fragile-and-conflict-affected-environments/>

Gholami, M., Pakdaman, A., Montazeri, A., and Virtanen, J. I. (2017). Evaluation Of The Impact Of A Mass Media Campaign On Periodontal Knowledge Among Iranian Adults: A Three-month Follow-up. *PloS One*, 12(1), e0169668.
<https://doi.org/10.1371/journal.pone.0169668>

Global indexes. (n.d.). Vision Of Humanity. Retrieved November 6, 2020, from
<https://www.visionofhumanity.org/maps/>

Goetz, A. M. (2018, March 13). What Does The New Women, Peace, And Security Index Measure? *IPI Global Observatory*. <https://theglobalobservatory.org/2018/03/what-does-wps-index-measure/>

GPI_2020_web.pdf. (n.d.). Retrieved November 6, 2020, from
https://visionofhumanity.org/wp-content/uploads/2020/10/GPI_2020_web.pdf

Hdr2019_technical_notes.pdf. (n.d.). Retrieved November 6, 2020, from
http://hdr.undp.org/sites/default/files/hdr2019_technical_notes.pdf

Hemsteede, R. (2018). *Conditional Or Unconditional Cash Transfers? From Ideology To Policy Dialogue*. Socialprotection.Org.
<https://socialprotection.org/discover/blog/conditional-or-unconditional-cash-transfers-ideology-policy-dialogue>

IDEA. (2021). *Gender Quotas | International IDEA*. <https://www.idea.int/data-tools/data/gender-quotas/quotas#what>

Khasnabis, C., Motsch, K. H., Achu, K., Jubah, K. A., Brodtkorb, S., Chervin, P., Coleridge, P., Davies, M., Deepak, S., Eklindh, K., Goerdt, A., Greer, C., Heinicke-Motsch, K., Hooper, D., Ilagan, V. B., Jessup, N., Khasnabis, C., Mulligan, D., Murray, B., ... Lander, T. (2010). Self-help groups. In *Community-Based Rehabilitation: CBR Guidelines*. World Health Organization. <https://www.ncbi.nlm.nih.gov/books/NBK310972/>

- Klugman, J., Nagel, R., & Viollaz, M. (2021). *Women's empowerment as a path to peace*. 23.
- Miller, C., Tsoka, M., Boyd-Boffa, M., Msiska, K., Brooks, M., Mtonga, D., Chambers, E., Cunningham, E., Mwanza, M., Henninger, N., Chimwaza, D., and Khonje, D. (2008). *Impact Evaluation Report External Evaluation of the Mchinji Social Cash Transfer Pilot*.
- OECD Highlights documents_web.pdf*. (n.d.). Retrieved November 6, 2020, from https://www.oecd.org/dac/conflict-fragility-resilience/docs/OECD%20Highlights%20documents_web.pdf
- Perova, E., and Reynolds, S. A. (2017). Women's Police Stations And Intimate Partner Violence: Evidence From Brazil. *Social Science and Medicine* (1982), 174, 188–196. <https://doi.org/10.1016/j.socscimed.2016.12.008>
- Pompa, C. (2014). *TVET and skills training in fragile and conflict affected countries*. 28.
- Search for Common Ground. (2016). *Community Dialogue Design Manual*. Common Ground Institute. <https://www.sfcg.org/wp-content/uploads/2017/07/CGI-Anglais-interactive.pdf>
- States of Fragility 2015: Meeting Post-2015 Ambitions | en | OECD*. (n.d.). Retrieved November 6, 2020, from <http://www.oecd.org/dac/states-of-fragility-2015-9789264227699-en.htm>
- Strickland, R. (n.d.). *Gender Equity and Peacebuilding: From Rhetoric to Reality*. 48.
- Sweetman, C., and Rowlands, J. (2016). Introduction: Working On Gender Equality In Fragile Contexts. *Gender and Development*, 24(3), 337–351. <https://doi.org/10.1080/13552074.2016.1248033>
- UNFPA. (2015). *UNFPA Women and Girls Safe Spaces Guidance*. <https://www.unfpa.org/sites/default/files/resource-pdf/UNFPA%20UNFPA%20Women%20and%20Girls%20Safe%20Spaces%20Guidance%20%5B1%5D.pdf>
- VSL Associates. (2021). *The VSLA Methodology*. VSL Associates. <https://www.vsla.net/the-vsla-methodology/>
- White, H., Menon, R., and Waddington, H. (2018). *Community-driven development: Does it build social cohesion or infrastructure? A mixed-method evidence synthesis*. 1.
- Women, peace and security index 2019/20 inclusion, justice and security*. (2019). Georgetown Institute for Women, Peace and Security (GIWPS).
- Women, Peace, and Security Index. (n.d.). *Georgetown Institute of Women Peace and Security*. Retrieved November 6, 2020, from <https://giwps.georgetown.edu/the-index/>
- World Bank. (2021a). *Community-Driven Development* [Text/HTML]. World Bank. <https://www.worldbank.org/en/topic/communitydrivendevelopment>
- World Bank. (2021b). *Microfinance And Economic Development* [Text/HTML]. World Bank. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/107171511360386561/Microfinance-and-economic-development>

Zuckerman, E., & Greenberg, M. (2004). The gender dimensions of post-conflict reconstruction: An analytical framework for policymakers. *Gender & Development*, 12(3), 70–82. <https://doi.org/10.1080/13552070412331332330>

Other systematic reviews

A Review of the Evidence Developed for a Technical Consultation on Expanding Access to Injectable Contraception. (n.d.). 51.

A Review Of The Evidence Developed For A Technical.pdf. (n.d.). Retrieved June 21, 2021, from

<https://www.fhi360.org/sites/default/files/media/documents/A%20Review%20of%20the%20Evidence%20Developed%20for%20a%20Technical%20Consultation%20on%20Expanding%20Access%20to%20Injectable%20Contraception.pdf>

Arango, D. J., Morton, M., Gennari, F., Kiplesund, S., and Ellsberg, M. (n.d.). *Interventions To Prevent Or Reduce Violence Against Women And Girls: A Systematic Review Of Reviews*. 61.

Arango et al. - Interventions to Prevent or Reduce Violence Again.pdf. (n.d.). Retrieved June 21, 2021, from

<https://openknowledge.worldbank.org/bitstream/handle/10986/21035/927130NWP0Wome00Box385382B00PUBLIC0.pdf?sequence=1>

Baird, S., Ferreira, F. H. G., Özler, B., and Woolcock, M. (2013). Relative Effectiveness Of Conditional And Unconditional Cash Transfers For Schooling Outcomes In Developing Countries: A Systematic Review. *Campbell Systematic Reviews*, 9(1), 1–124.

<https://doi.org/10.4073/csr.2013.8>

Brody, C., Hoop, T. de, Vojtkova, M., Warnock, R., Dunbar, M., Murthy, P., and Dworkin, S. L. (2017). Can Self-help Group Programs Improve Women's Empowerment? A Systematic Review. *Journal of Development Effectiveness*, 9(1), 15–40.

<https://doi.org/10.1080/19439342.2016.1206607>

De Koker, P., Mathews, C., Zuch, M., Bastien, S., and Mason-Jones, A. J. (2014). A Systematic Review Of Interventions For Preventing Adolescent Intimate Partner Violence. *Journal of Adolescent Health*, 54(1), 3–13.

<https://doi.org/10.1016/j.jadohealth.2013.08.008>

Eggers del Campo, I., and Steinert, J. I. (2020). The Effect Of Female Economic Empowerment Interventions On The Risk Of Intimate Partner Violence: A Systematic Review And Meta-analysis. *Trauma, Violence, and Abuse*, 1524838020976088.

<https://doi.org/10.1177/1524838020976088>

Emezi, Y., & Images, G. (2021). *Governance, Crime and Conflict Initiative Evidence Wrap-Up*. 109.

LaCroix, J. M., Snyder, L. B., Huedo-Medina, T. B., and Johnson, B. T. (2014). Effectiveness Of Mass Media Interventions For HIV Prevention, 1986–2013: A Meta-analysis. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 66, S329.

<https://doi.org/10.1097/QAI.0000000000000230>

Langer, L., Erasmus, Y., Tannous, N., Obuku, E., Ravat, Z., Chisoro, C., Opondo, M., Nduku, P., Tripney, J., van Rooyen, C., and Stewart, R. (2018). Women In Wage Labour: A Systematic Review Of The Effectiveness And Design Features Of Interventions Supporting Women's Participation In Wage Labour In Higher-growth And/Or Male-dominated Sectors In Low- And Middle-income Countries. In *EPPI-Centre, Social Science Research Unit, UCL Institute of Education: London, UK*. [Report]. EPPI-Centre, Social Science Research Unit, UCL Institute of Education.
<https://eppi.ioe.ac.uk/cms/Coursesseminars/Seminarsevents/Womeninwagelabour/tabid/3749/Default.aspx>

Lassi et al. - 2015—*Impact Of Service Provision Platforms On Maternal .pdf*. (n.d.). Retrieved June 21, 2021, from
<https://conflictandhealth.biomedcentral.com/track/pdf/10.1186/s13031-015-0054-5.pdf>

Lassi, Z. S., Aftab, W., Ariff, S., Kumar, R., Hussain, I., Musavi, N. B., Memon, Z., Soofi, S. B., and Bhutta, Z. A. (2015). Impact Of Service Provision Platforms On Maternal And Newborn Health In Conflict Areas And Their Acceptability In Pakistan: A Systematic Review. *Conflict and Health*, 9(1), 25. <https://doi.org/10.1186/s13031-015-0054-5>

Manley, J., Gitter, S., and Slavchevska, V. (2013). How Effective Are Cash Transfers At Improving Nutritional Status? *World Development*, 48, 133–155.
<https://doi.org/10.1016/j.worlddev.2013.03.010>

McQueston, K., Silverman, R., and Glassman, A. (2013). The Efficacy Of Interventions To Reduce Adolescent Childbearing In Low- And Middle-income Countries: A Systematic Review. *Studies in Family Planning*, 44(4), 369–388. <https://doi.org/10.1111/j.1728-4465.2013.00365.x>

Providing access to economic assets for girls and young women in low-and-lower middle-income countries: A systematic review of the evidence. (2012). Social Science Research Unit.

Stark, L., Robinson, M. V., Seff, I., Gillespie, A., Colarelli, J., and Landis, D. (2021). The Effectiveness Of Women And Girls Safe Spaces: A Systematic Review Of Evidence To Address Violence Against Women And Girls In Humanitarian Contexts. *Trauma, Violence, and Abuse*, 1524838021991306. <https://doi.org/10.1177/1524838021991306>

Vaessen, J., Rivas, A., Duvendack, M., Jones, R. P., Leeuw, F., Gils, G. van, Lukach, R., Holvoet, N., Bastiaensen, J., Hombrados, J. G., and Waddington, H. (2014). The Effects Of Microcredit On Women's Control Over Household Spending In Developing Countries: A Systematic Review And Meta-analysis. *Campbell Systematic Reviews*, 10(1), 1–205.
<https://doi.org/10.4073/csr.2014.8>

White, H., Menon, R., & Waddington, H. (2018). *Community-driven development: Does it build social cohesion or infrastructure?* (p. 52). 3ie.
https://www.3ieimpact.org/sites/default/files/2019-01/wp30-cdd_0.pdf

Other publications in the 3ie Systematic Review Series

The following reviews are available at

<http://www.3ieimpact.org/evidence-hub/publications/systematic-reviews>

Strengthening intergroup social cohesion in fragile situations. 3ie Systematic Review 46. Sonnenfeld, A, Doherty, J, Berretta, M, Shisler, S, Snilstveit, B, Eyers, J, Castaman, K, Gupta, R, Anda Leon, MD, Franich, A, Yavuz, C, Baafi, A and Obaid, R, 2021.

Effects of electricity access interventions on socio-economic outcomes in low- and middle-income countries. 3ie Systematic Review 45. Moore, N, Glandon, D, Tripney, J, Kozakiewicz, T, Shisler, S, Eyers, J, Zalfou, R, Leon, MDA, Kurkjian, V, Snilstveit, B and Perdana, A. 2020.

Incentives for climate mitigation in the land use sector – the effects of payment for environmental services (PES) on environmental and socio-economic outcomes in low- and middle-income countries: a mixed-method systematic review. 3ie Systematic Review 44. Snilstveit, B, Stevenson, J, Langer, L, da Silva, N, Rabat, Z, Nduku, P, Polanin, J, Shemilt, I, Eyers, J, Ferraro, PJ, 2019.

Does incorporating participation and accountability improve development outcomes? Meta-analysis and framework synthesis. 3ie Systematic Review 43. Waddington, H, Sonnenfeld, A, Finetti, J, Gaarder, M and Stevenson, J, 2019.

Impact of financial inclusion in low- and middle-income countries: a systematic review of reviews. 3ie Systematic Review 42. Duvendack, M and Mader, P, 2019.

Agricultural input subsidies for improving productivity, farm income, consumer welfare and wider growth in low- and middle-income countries: a systematic review. 3ie Systematic Review 41. Hemming, DJ, Chirwa, EW, Ruffhead, HJ, Hill, R, Osborn, J, Langer, L, Harman, L, Coffey, C, Dorward, A and Phillips, D, 2018.

Vocational and business training to improve women's labour market outcomes in low- and middle-income countries: a systematic review. 3ie Systematic Review 40. Chinen, M, De Hoop, T, Balarin, M, Alcázar, L, Sennett, J, and Mezarina, J, 2018.

Interventions to improve the labour market for adults living with physical and/or sensory disabilities in low- and middle-income countries: a systematic review. 3ie Systematic Review 39. Tripney, J, Roulstone, A, Vigurs, C, Hoglebe, N, Schmidt, E and Stewart, R, 2017.

The effectiveness of contract farming in improving smallholder income and food security in low- and middle-income countries: a mixed-method systematic review. 3ie Systematic Review 38. Ton, G, Desiere, S, Vellema, W, Weituschat, S and D'Haese, M (2017)

Interventions to improve the labour market outcomes of youth: a systematic review of training, entrepreneurship promotion, employment services and subsidized employment interventions. 3ie Systematic Review 37. Kluve J, Puerto S, Robalino D, Romero JM, Rother F, Stöterau J, Weidenkaff F and Witte M (2017)

Promoting handwashing and sanitation behaviour change in low- and middle-income countries: a mixed-method systematic review. 3ie Systematic Review 36. Buck, ED, Remoortel, HV, Hannes, K, Govender, T, Naidoo, S, Avau, B, Veegaete, AV, Musekiwa, A, Lutje, V, Cargo, M, Mosler, HJ, Vandekerckhove, P and Young T (2017)

Incorporating the life cycle approach into WASH policies and programmes: A systematic review. 3ie Systematic Review 35. Annamalai, TR, Narayanan, S, Devkar, G, Kumar, VS, Devaraj, R, Ayyangar, A and Mahalingam, A (2017)

Effects of certification schemes for agricultural production on socio-economic outcomes in low- and middle-income countries: a systematic review 34. Oya, C, Schaefer, F, Skolidou, D, McCosker, C and Langer, L (2017)

Short-term WASH interventions in emergency response: a systematic review. 3ie Systematic Review 33. Yates, T, Allen, J, Joseph, ML and Lantagne, D (2017)

Community monitoring interventions to curb corruption and increase access and quality of service delivery in low- and middle-income countries. 3ie Systematic Review 32. Molina E, Carella L, Pacheco A, Cruces, G and Gasparini, L (2016)

Effects and mechanisms of market-based reforms on access to electricity in developing countries: a systematic review. 3ie Systematic Review 31. Bensch, G, Sievert, M, Langbein, J, Kneppel, N (2016)

Youth gang violence and preventative measures in low- and middle-income countries: a systematic review (Part II), 3ie Systematic Review 30. Higginson, A, Benier, K, Shenderovich, Y, Bedford, L, Mazerolle, L, Murray, J (2016)

Youth gang membership and violence in low- and middle-income countries: a systematic review (Part I), 3ie Systematic Review 29. Higginson, A, Benier, K, Shenderovich, Y, Bedford, L, Mazerolle, L, Murray, J (2016)

Cash-based approaches in humanitarian emergencies: a systematic review, 3ie Systematic Review Report 28. Doocy, S and Tappis, H (2016)

Factors affecting uptake of voluntary and community-based health insurance schemes in low-and middle-income countries: a systematic review, 3ie Systematic Review 27. Panda, P, Dror, IH, Koehlmoos, TP, Hossain, SAS, John, D, Khan, JAM and Dror, DM (2016)

Parental, community and familial support interventions to improve children's literacy in developing countries: a systematic review, 3ie Systematic Review 26. Spier, E, Britto, P, Pigott, T, Roehlkapartain, E, McCarthy, M, Kidron, Y, Song, M, Scales, P, Wagner, D, Lane, J and Glover, J (2016)

Business support for small and medium enterprises in low- and middle-income countries: a systematic review, 3ie Systematic Review 25. Piza, C, Cravo, T, Taylor, L, Gonzalez, L, Musse, I, Furtado, I, Sierra, AC and Abdelnour, S (2016)

Interventions for improving learning outcomes and access to education in low- and middle-income countries: a systematic review, 3ie Systematic Review 24. Snilstveit, B, Stevenson, J, Phillips, D, Vojtkova, M, Gallagher, E, Schmidt, T, Jobse, H, Geelen, M, Pastorello, M, and Eysers, J (2015)

Economic self-help group programmes for improving women's empowerment: a systematic review, 3ie Systematic Review 23. Brody, C, De Hoop, T, Vojtkova, M, Warnock, R, Dunbar, M, Murthy, P and Dworkin, SL (2016)

The identification and measurement of health-related spillovers in impact evaluations: a systematic review, 3ie Systematic Review 22. Benjamin-Chung, J, Abedin, J, Berger, D, Clark, A, Falcao, L, Jimenez, V, Konagaya, E, Tran, D, Arnold, B, Hubbard, A, Luby, S, Miguel, E and Colford, J (2015)

The effects of school-based decision-making on educational outcomes in low- and middle-income countries: a systematic review, 3ie Systematic Review Report 21. Carr-Hill, R, Rolleston, C, Pherali, T and Schendel, R, with Peart, E, and Jones, E (2015)

Policing interventions for targeting interpersonal violence in developing countries: a systematic review, 3ie Systematic Review 20. Higginson, A, Mazerolle, L, Sydes, M, Davis, J, and Mengersen, K (2015)

The effects of training, innovation and new technology on African smallholder farmers' wealth and food security: a systematic review, 3ie Systematic Review 19. Stewart, R, Langer, L, Rebelo Da Silva N, Muchiri, E, Zaranyika, H, Erasmus, Y, Randall, N, Rafferty, S, Korth, M, Madinga, N and de Wet, T (2015)

Community based rehabilitation for people with disabilities in low- and middle-income countries: a systematic review, 3ie Systematic Review 18. Lemmi, V, Gibson, L, Blanchet, K, Kumar, KS, Rath, S, Hartley, S, Murthy, GVS, Patel, V, Weber, J and Kuper H (2015)

Payment for environmental services for reducing deforestation and poverty in low- and middle-income countries: a systematic review, 3ie Systematic Review 17. Samii, C, Lisiecki, M, Kulkarni, P, Paler, L and Chavis, L (2015)

Decentralised forest management for reducing deforestation and poverty in low- and middle- income countries: a systematic review, 3ie Systematic Review 16. Samii, C, Lisiecki, M, Kulkarni, P, Paler, L and Chavis, L (2015)

Supplementary feeding for improving the health of disadvantaged infants and young children: a systematic and realist review, 3ie Systematic Review 15. Kristjansson, E, Francis, D, Liberato, S, Greenhalgh, T, Welch, V, Jandu, MB, Batal, M, Rader, T, Noonan, E, Janzen, L, Shea, B, Wells, GA and Petticrew, M (2015)

The impact of land property rights interventions on investment and agricultural productivity in developing countries: a systematic review, 3ie Systematic Review Report 14. Lawry, S, Samii, C, Hall, R, Leopold, A, Hornby, D and Mtero, F, 2014.

Slum upgrading strategies and their effects on health and socio-economic outcomes: a systematic review, 3ie Systematic Review 13. Turley, R, Saith, R., Bhan, N, Rehfuess, E, and Carter, B (2013)

Services for street-connected children and young people in low- and middle-income countries: a thematic synthesis, 3ie Systematic Review 12. Coren, E, Hossain, R, Ramsbotham, K, Martin, AJ and Pardo, JP (2014)

Why targeting matters: examining the relationship between selection, participation and outcomes in farmer field school programmes, 3ie Systematic Review 11. Phillips, D, Waddington, H and White, H (2015)

The impact of export processing zones on employment, wages and labour conditions in developing countries, 3ie Systematic Review 10. Cirera, X and Lakshman, R (2014)

Interventions to reduce the prevalence of female genital mutilation/cutting in African countries, 3ie Systematic Review 9. Berg, RC and Denison, E (2013)

Behaviour change interventions to prevent HIV among women living in low and middle income countries, 3ie Systematic Review 8. McCoy, S, Kangwende, RA and Padian, NS (2009)

The impact of daycare programs on child health, nutrition and development in developing countries, 3ie Systematic Review 7. Leroy, JL, Gadsden, P and Guijarro, M (2011)

Willingness to pay for cleaner water in less developed countries: Systematic review of experimental evidence, 3ie Systematic Review 6. Null, C, Hombrados, JG, Kremer, M, Meeks, R, Miguel, E and Zwane, AP (2012)

Community-based intervention packages for reducing maternal morbidity and mortality and improving neonatal outcomes, 3ie Systematic Review 5. Lassi, ZS, Haider, BA and Langou, GD (2011)

The effects of microcredit on women's control over household spending: a systematic review, 3ie Systematic Review 4. Vaessen, J, Rivas, A, Duvendack, M, Jones, RP, Leeuw, F, van Gils, G, Lukach, R, Holvoet, N, Bastiaensen, J, Hombrados, JG and Waddington, H, (2013).

Interventions in developing nations for improving primary and secondary school enrolment of children: a systematic review, 3ie Systematic Review 3. Petrosino, A, Morgan, C, Fronius, T, Tanner-Smith, E, and Boruch, R, 2016.

Interventions to promote social cohesion in Sub-Saharan Africa, 3ie Systematic Review 2. King, E, Samii, C and Snilstveit, B (2010)

Water, sanitation and hygiene interventions to combat childhood diarrhoea in developing countries, 3ie Systematic Review 1. Waddington, H, Snilstveit, B, White, H and Fewtrell, L (2009)

International Initiative for Impact Evaluation
London International Development Centre
20 Bloomsbury Square
London WC1A 2NS
United Kingdom

3ieuk@3ieimpact.org
Tel: +44 207 958 8351/8350

www.3ieimpact.org