Instructions for Impact Evaluation Repository Screening

Please refer to the following step-by-step document to screen potential studies for inclusion in the Impact Evaluation Repository (IER). The document was prepared by Anjini Mishra (amishra@3ieimpact.org) and Drew Cameron (dcameron@3ieimpact.org) with guidance from 3ie technical staff.

This document represents the third stage of screening in 3ie's Impact Evaluation Repository Search and Screening Protocol. Typically, studies will have passed through two initial stages (both a title, and abstract review) before being screened here. This third review stage is a two-step blinded review by two screeners. Each screener completes the checklist in step 1, followed by the checklist in step 2 if applicable. Scores by the two screeners are then compared by a third party. Those studies receiving unanimous scores of "Yes" or "No" between the two screeners are either accepted or rejected respectively. Studies receiving conflicting scores or those with any "Unclear" scores are reviewed by the third party for rejection or inclusion in the repository.

Studies with potential to be included in the IER may also come to 3ie's attention through other channels. In this case, this screening tool may be used by a single reviewer to include or reject a study.

Screening step 1

Step 1 requires a thorough reading of the study abstract, and the full text of the article when necessary (some abstracts and online records contain sufficient information to screen studies in step 1). Screeners must select "Yes", "No", or "Uncertain" for each item in the Step 1 checklist.

- If all categories in Step 1 are scored "Yes", the final score should be coded **"Yes"** and the screener should proceed to step 2 to screen the study for methodological rigor.
- If any of the categories in step 1 are scored "Unclear" (while the rest are "Yes"), a detailed explanation for each "Unclear" score should be provided, the final score in step 1 should be coded **"Unclear"**, and the screener should proceed to step 2 to screen the study for methodological rigor.
- If any categories in step 1 are scored "No", the screener should provide a detailed explanation as to why the study was rejected in at least one category and provide a final score of **"No"**. Any "No" score means that the final score for Step 1 is a "No."

Each study receives a final score of "Yes", "No", or "Unclear" after all categories (1-6 or 6a) are scored. Screeners should only select "Unclear" if they have exhausted all attempts to provide a final score of "Yes" or "No". Note that the *Effectiveness* category (item 6a in the step 1 screening checklist) should only be completed if the study is a randomized controlled trial in the biomedical sciences. Any category with a score of "Unclear" will be reviewed by IER managers.

Step 1: Screening checklist

Note: This screening checklist corresponds to a separate spreadsheet where scores are entered and comments should be recorded as necessary.

- *For studies receiving a score of "No," provide detailed comments in the space provided for **AT LEAST ONE** category for which you have selected "No". Multiple explanations are unnecessary.*
- *For studies scoring "Unclear," please provide a detailed explanation for **EVERY** category for which you have selected "Unclear".*
- *Explanations are unnecessary for studies scoring "Yes" (for which all categories are "Yes").*

1. Is the study in English?	**Yes**	**Unclear**	**No**

Some studies may be written in another language, yet feature title and/or abstract reference information in English. Note the publication information and country of origin or locate the full text if you are unsure of the full text language.

- If the study was written in English, **select "Yes"**.
- If you have searched, yet it is unclear whether the full text is available in English, **select "Unclear"**.
- If the full text of the study is only available in a language other than English, **select "No"**.

2. Is the study new (not already in the IER)?	**Yes**	**Unclear**	**No**

Verify that study does not already exist in the IER (here: http://www.3ieimpact.org/en/evidence/impact-evaluations/). Search in the IER for **both** the title and author names. Some studies may have different titles than those that currently appear in the IER (such as working papers that are later published as journal articles), or may have different primary authors. If the study appears in the IER under a different name, confirm the most recent version of the paper and note which version should be included in the IER.

- If the study is not present in the IER, **select "Yes"**.
- If it is unclear whether the study (or another version) already exists in the IER, **select "Unclear"**.
- If the study already exists in the IER, **select "No"**.

3. Is the study published?	**Yes**	**Unclear**	**No**

To be included in the IER, studies must be "published" either in a journal, in a book (possibly as a book chapter), as a report from an organization, or in a working paper series. Studies available as non-attributed reports or working papers released only by the original author, or on a university website (not as part of a paper series), or with no publication information should be rejected.

- If the study is published, **select "Yes"**.
- If it is unclear whether the study was published, **select "Unclear"**.
- If the study is not published, **select "No"**.

4. Did the program or intervention take place in a developing country?	**Yes**	**Unclear**	**No**

Please refer to the attached lists of countries in Tables I, II, and III in the Appendix.
Note: Studies collecting data (even in part) in a developing country should be included (this includes countries categorized as low- or middle-income according to the World Bank). Please refer to Table III in the Appendix for a list of developing countries. For studies taking place in "transitional countries" refer to Table II in the Appendix to determine if the study data were collected during a historical period in which the country was classified as "under-developed". Studies taking place before 1987 should only be excluded if they are listed on the "Developed Countries" list (Table I in Appendix A).

- If the study took place in a developing country, **select "Yes".**
- If it is not clear whether the study took place in a developing country, **select "Unclear".**
- If the study only took place in a developed country, **select "No".**

5. Is the study of at least one specific policy, program, or intervention?	**Yes**	**Unclear**	**No**

Studies in the IER examine the effects of specific programs, policies, or interventions. Studies which only investigate natural or market-based occurrences, or that report on the findings of controlled laboratory experiments with no discernable development intervention should be excluded.

- If the study examines a specific policy, program, or intervention, **select "Yes".**
- If it is unclear whether the study examines a specific policy, program, or intervention, **select "Unclear".**
- If the study does not examine a specific policy, program, or intervention, **select "No".**

6. Does the study use at least one of the following methods?	**Yes**	**Unclear**	**No**

Studies in the IER must compare a treatment condition to a counterfactual (what would happen in the absence of the treatment). Note that studies using a *pipeline* or *phased-in* approach are only valid if they also utilize one of the following methods. The following are considered rigorous approaches to counterfactual analysis in impact evaluation:

i. Randomized Controlled Trial (RCT).
ii. Regression Discontinuity Design (RDD).
iii. Propensity Score Matching (PSM) for non-randomized studies based on participant self-selection, or other Matching Methods.
iv. Instrumental Variable (IV) estimation (or other methods using an instrumental variable such as the Heckman Two Step approach).
v. Difference-in-Differences (DD), or a fixed or random effects model with an interaction term between time and intervention for baseline and follow-up observations.

- If the study uses at least one of the methods (i-v) above, **select "Yes".**
- If it is unclear whether the study uses one of the methods above, **select "Unclear".**
- If it is totally clear that the study does not use any of the methods above, **select "No".**

6a. If the study is an RCT, does it answer an effectiveness question?	**Yes**	**Unclear**	**No**

Studies may exist anywhere on the continuum of effectiveness vs. efficacy. Typically, efficacy studies examine treatment outcomes under highly controlled conditions. Effectiveness studies go beyond laboratory trials and examine interventions in real world settings. Note that RCTs that only address the biomedical efficacy of a drug or treatment should be excluded. The following are screening guidelines to help make this judgment:

If any of these conditions are met in addition to methodological criteria in #6 above**, select "Yes":**

a. The intervention under study promotes a social, economic, or behavioral change either as one of the final measured outcomes or as a mechanism within the theory of change (beyond the self-administration of a drug). *For example, the study may include health / behavioral messaging, training, provision of information, or screening / surveillance for specific disease conditions.*
b. The study measures any other outcomes in addition to or beyond purely biomedical indicators (such as returns to education, economic productivity, quality of life, disability adjusted life years, or spillover effects)
c. The study measures the cost-effectiveness or cost-benefit of the treatment(s).
d. The study records any additional formative information that could guide the design or execution of future studies. For example, an RCT that also measures acceptability of a

particular treatment (measuring respondent satisfaction with treatment not merely a rate of compliance or uptake) would be included.
e. The treatment is both prepared and delivered by a community health worker, or trained layperson (such as a parent, teacher, or community member and not merely one of the program or study enumeration team).
f. The program or outcomes measured answer, or attempt to answer, a question relevant to the roll-out of international development policies or interventions.

If it is unclear whether the study meets any of these conditions (a-f), **select "Unclear"**.
Note that in erring on the side of inclusion, studies which are "Unclear" should likely be included.)

If the study meets none of these conditions (a-f), **select "No".**

Final Score:	**Yes**	**Unclear**	**No**

Screening step 2

All studies identified as "Yes" or "Unclear" during step 1 should be screened in step 2. Step 2 requires a full-text reading of the study to determine whether it uses one of the required impact evaluation methods. First, locate the full text of each record (as possible). Next, indicate whether the final results are reported in either the abstract or full text of the study (Yes, Unclear, or No). For those studies scoring "Yes" or "Unclear", evaluate the study based on the screening criteria below in item 2.

Screeners should provide detailed explanations for all "No" and "Unclear" scores. Screeners should *not*assess the quality of the study, merely whether at least one of the listed identification strategies was used appropriately.

Step 2: Full text study methodology review			
1. Are impact evaluation results reported?	Yes	Unclear	No

Study results (outcomes or impacts of the intervention) must be reported for the study to be included in the IER. Some examples of studies that do not report results include a study protocol (randomized controlled trials often publish study protocols, though usually contain the word "protocol" in the study title), a process evaluation (reporting on the success or failure of an intervention to deliver certain program components, but not on the actual effect of these components on beneficiary populations), or studies reporting only on baseline findings (baseline measurements taken before the intervention has taken place).

- If the study results are reported either in the abstract or full text of the article, **select "Yes".**
- If it is unclear from the abstract and full text whether the study reports evaluation results, **select "Unclear".**
- If study results are not reported, **select "No".**

2. If scoring "Yes" or "Unclear" in #1 above, identify the primary identification strategy utilized in the study from the options (i-v) below and assess that method on the scoring criteria. For studies that have multiple identification strategies, only one identification strategy needs to be scored "Yes" for the study to be accepted. Please review each strategy until either scoring "Yes" for any one, or scoring "No" or "Unclear" for all

methods utilized in the study (i-v).

i. Randomized Controlled Trial (RCT)	Yes	Unclear	No

Randomized Controlled Trials use random assignment to allocate the intervention amongst members of the eligible population. Participants must have an equal chance of allocation to either the treatment or control group.

- If some form of centralized allocation mechanism or a random component in the sequence generation process is described (whether the allocation is by group/cluster or individual) so as to develop a randomly assigned treatment and control group, **select "Yes".**
- If the paper does not provide details on the randomization process, or if the paper may use a quasi-randomization process which is not equivalent to true randomization, **select "Unclear".**
- If there is any failure in the allocation mechanism that could affect the randomization process, **select "No".**

ii. Regression Discontinuity Design (RDD)	Yes	Unclear	No

Regression Discontinuity Designs identify treatment and comparison groups as being those just either side of some observable threshold value of a variable. This variable may be a score or an observed characteristic (e.g. age or land holding) used by program staff to determine the eligible population, or it may be a variable found to distinguish participants from non-participants through data analysis.

- If allocation is made based on a pre-determined discontinuity on a continuous variable(s) with a clearly defined cutoff point and individuals cannot affect the assignment variable(s) in response to knowledge of the participation decision rule, **AND** a clearly defined treatment and control group is established at both sides of the cut-off point, **select "Yes".**
- If the assignment variable is non-blinded, or it is unclear whether participants can affect it in response to knowledge of the allocation mechanism, or the assignment variable(s) or cutoff point are unclear, **select "Unclear".**
- If there is evidence that participants altered the assignment variable prior to assignment, or there are/is no assignment variable(s) or no cutoff point, **select "No".**

iii. Propensity Score Matching (PSM) or Other Matching Method	Yes	Unclear	No

Propensity Score Matching compares outcomes in a treatment group to those in a comparison group that is constructed through matching based on propensity scores. Propensity scores are the probability of participating in the intervention based on a metric of observed characteristics (by combining these characteristics in a single "score"). Selected characteristics must not be affected by the intervention.
Other matching methods, such as "covariate matching," must a) match treatment to control observations using explicit criteria, b) use variables for matching that explain the measured outcomes, and c) check for balance.

- If the study matches participant and non-participant observations using explicit criteria (relevant baseline or time-invariant characteristics) in a way that addresses potential issues of self-selection bias and the variables in the selection equation are unaffected by the intervention, **select "Yes".**
- If it is not clear whether relevant time-invariant characteristics (or relevant time varying characteristics in the case of panel data) are accounted for in the matching, **select "Unclear".**
- If the study did not use appropriate propensity score matching or other matching techniques, **select "No".**

iv. Instrumental Variable Estimation (IV)	Yes	Unclear	No

An instrumental variable helps to identify causal impacts of an intervention when participation is partially determined by beneficiaries. Instrumental variables must be correlated with participation or enrollment in the intervention, and exogenous to the outcome variable(s) (except through participation). IV estimation may also include methods such as the Heckman two-step procedure, 2SLS, LIML, etc. where an appropriate exogenous instrumental variable is used to control for possible self-selection bias.

- If an appropriate instrumental variable is used which is exogenously generated (for example, due to a "natural" experiment or random allocation), **select "Yes".**
- If the exogeneity of the instrument is unclear or unconvincing, **select "Unclear".**
- If the study did not use an appropriate instrumental variable estimation strategy, **select "No".**

v. Difference-in-Differences or Double Difference (DD)	Yes	Unclear	No

Double difference estimation measures the change in the outcome variable(s) observed in the treatment group compared to the change observed in the control group over time. Difference-in-differences may also include triple-difference estimation to examine impact heterogeneities.

- If the study uses a difference-in-differences multivariate estimation method, **select "Yes".**
- If the study uses a fixed or random effects model with an interaction term between time and intervention to establish a trend between baseline and follow-up panel data, but do not explicitly mention "difference-in-differences", **select "Unclear".**
- If the study does not use a difference-in-differences estimation method, **select "No".**

Final Score	Yes	Unclear	No

Appendix - Country lists

The following (tables I, II, and III) are taken from the World Bank.[1]

Table I. Developed countries	
Andorra	Australia
Austria	Bahamas
Belgium	Bermuda
Brunei Darussalam	Canada
Cayman Islands	Channel Islands
Curacao	Denmark
Faeroe Islands	Finland
France	French Polynesia
Germany	Greenland
Hong Kong (SAR)	Iceland
Ireland	Israel
Italy	Japan
Kuwait	Liechtenstein
Luxembourg	Monaco
Netherlands	New Zealand
Norway	Qatar
San Marino	Singapore
Sint Maarten (Dutch Part)	Spain
St. Martin (French Part)	Sweden
Switzerland	Taiwan
Turks and Caicos Islands	United Arab Emirates
United Kingdom	United States
Virgin Islands (U.S.)	

Note: Table I includes all nations that have continuously exceeded the threshold for "high income country" status since either fiscal year 1987 (when the delineation was first established in 1989), or since the establishment of each nation.

[1] http://siteresources.worldbank.org/DATASTATISTICS/Resources/OGHIST.xls

Table II. Transitional countries		
Country	**Developed period**	**Under-developed period**
American Samoa	1987-1989	1990-present
Aruba	1987-1990; 1994-present	1991-1993
Bahrain	1987-1989; 2001-present	1990-2000
Barbados	1989; 2000; 2002; 2006-present	1987-1988; 1990-1999; 2001; 2003-2005
Croatia	2008-present	1992-2007
Cyprus	1988-present	1987
Czech Republic	2006-present	1992-2005
Equatorial Guinea	2007-present	1987-2006
Estonia	2006-present	1991-2005
Guam	1987-1989; 1995-present	1990-1994
Greece	1996-present	1987-1995
Hungary	2006-present	1987-2006
Isle of Man	1987-1989; 2002-present	1990-2001
Latvia	2009	1991-2008; 2010-present
Macao (SAR)	1994-present	1987-1993
Malta	1989; 1998; 2000; 2002-present	1987-1988; 1990-1997; 1999; 2001
New Caledonia	1995-present	1987-1994
Northern Mariana Islands	1995-2001; 2007-present	1992-1994; 2002-2006
Oman	2007-present	1987-2006
Poland	2009-present	1987-2008
Portugal	1994-present	1987-1993
Puerto Rico	1989; 2002-present	1987-1988; 1990-2001
Republic of Korea	1995-1997; 2001-present	1987-1994; 1998-2000
Slovak Republic	2007-present	1992-2006
Slovenia	1997-present	1992-1996
Saudi Arabia	1987-1989; 2006-present	1990-2003
St. Kitts and Nevis	2011-present	1987-2010
Trinidad and Tobago	2006-present	1987-2005

Note: Table II includes nations that have (at some point since 1987) crossed the threshold of High Income Country status. In the event that relevant impact evaluations were conducted during a period of time in which the nation was *not* in the high income country category, that study should be included for further screening. Note that FY information is only available through 2011.

Table III. Developing countries (middle- or low-Income)			
Afghanistan	Egypt, Arab Rep.	Malaysia	South Africa
Albania	El Salvador	Maldives	South Sudan
Algeria	Eritrea	Mali	Sri Lanka
Angola	Ethiopia	Marshall Islands	St. Lucia
Antigua and Barbuda	Fiji	Mauritania	St. Vincent and the Grenadines
Argentina	Gabon	Mauritius	Sudan
Armenia	Gambia, The	Mexico	Suriname
Azerbaijan	Georgia	Micronesia, Fed. States	Swaziland
Bangladesh	Ghana	Moldova	Syrian Arab Republic
Belarus	Grenada	Mongolia	Tajikistan
Belize	Guam	Montenegro	Tanzania
Benin	Guatemala	Morocco	Thailand
Bhutan	Guinea	Mozambique	Timor-Leste
Bolivia	Guinea-Bissau	Myanmar	Togo
Bosnia and Herzegovina	Guyana	Namibia	Tonga
Botswana	Haiti	Nepal	Tunisia
Brazil	Honduras	Nicaragua	Turkey
Bulgaria	India	Niger	Turkmenistan
Burkina Faso	Indonesia	Nigeria	Tuvalu
Burundi	Iran, Islamic Rep.	Pakistan	Uganda
Cambodia	Iraq	Palau	Ukraine
Cameroon	Jamaica	Panama	Uruguay
Cape Verde	Jordan	Papua New Guinea	Uzbekistan
Central African Republic	Kazakhstan	Paraguay	Vanuatu
Chad	Kenya	Peru	Venezuela, RB
Chile	Kiribati	Philippines	Vietnam
China	Korea, Dem. Rep.	Romania	West Bank and Gaza
Colombia	Kosovo	Russian Federation	Yemen, Rep.
Comoros	Kyrgyz Republic	Rwanda	Zambia
Congo, Dem. Rep.	Lao PDR	Samoa	Zimbabwe
Congo, Rep.	Lebanon	São Tomé and Principe	**Former countries**
Costa Rica	Lesotho	Saudi Arabia	Czechoslovakia
Côte d'Ivoire (Ivory Coast)	Liberia	Senegal	Gibraltar
Cuba	Libya	Serbia	Mayotte
Djibouti	Lithuania	Seychelles	Netherlands Antilles
Dominica	Macedonia, FYR	Sierra Leone	Serbia and Montenegro
Dominican Republic	Madagascar	Solomon Islands	USSR
Ecuador	Malawi	Somalia	Yugoslavia

Note: All nations in table III have remained below the threshold for "developed" countries since 1989, or since their individual inception.