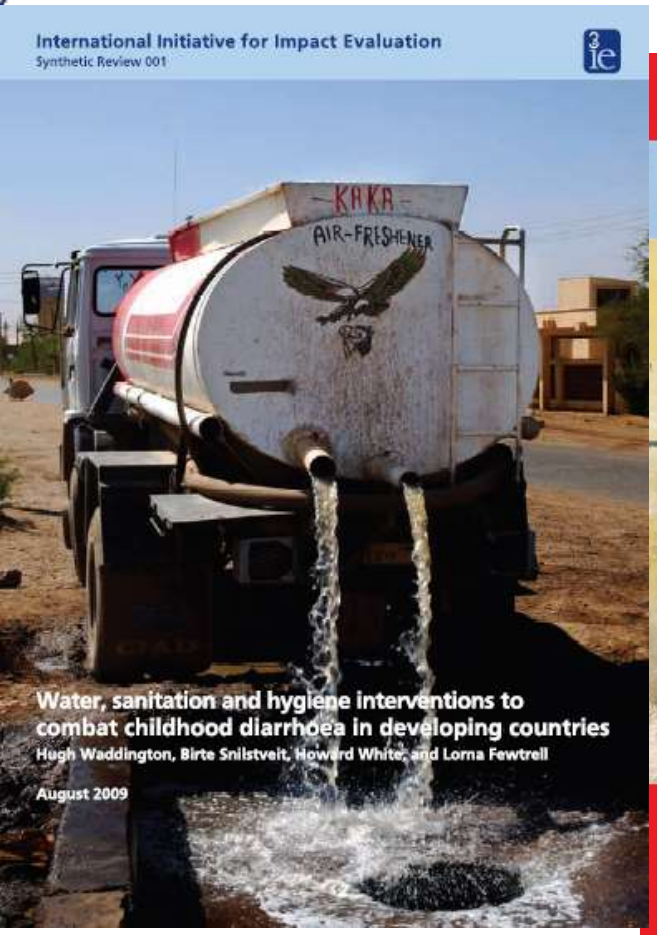
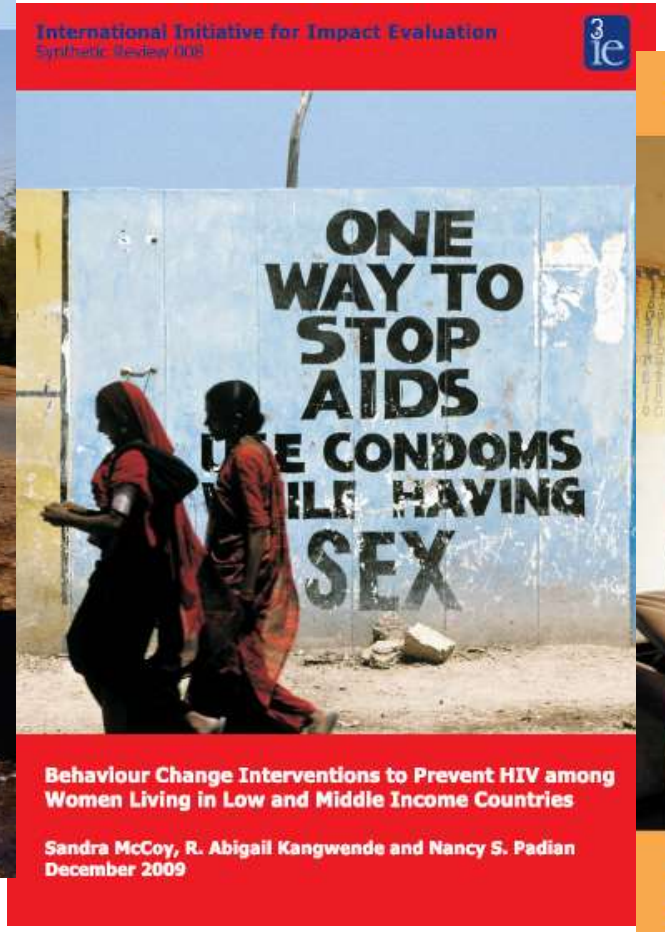
How to design and conduct a systematic review of education interventions?

Birte Snilstveit and Martina Vojtkova, 3ie

International Initiative for Impact Evaluation

Systematic Reviews





Why do a Systematic review?

"Small loans enable people in poverty to earn an income and provide for their families… Each successful business feeds a family, employs more people and eventually helps empower a whole community"

"Impact studies done on the Grameen Bank by independent researchers find that 5% of borrowers come out of poverty every year…the status of women has been enhanced"

"Women participants in microcredit programs often experience important self-empowerment… there is a strong indication from borrowers that microcredit improves their lives"

So what do reviews synthesising all the existing high quality evidence say?

"All impact evaluations of microfinance suffer from weak methodologies and inadequate data [which] can lead to misconceptions about the actual effects of a microfinance programme…It remains unclear under what circumstances, and for whom, microfinance has been and could be of real, rather than imagined, benefit to poor people"

Duvendack et al. (2011)

So what do reviews synthesising all the existing high quality evidence say?

"From those studies deemed comparable and of minimum acceptable quality, we can conclude that overall the effect of microcredit on women's control over household spending is weak'"

Vaessen et al. (forthcoming)

Why do we need systematic reviews?

- Sheer amount and flow of information/ research
- Variable quality of research outputs – evidence often anecdotal, unsystematic,
- Need to ‘separate the wheat from the chaff’
- Problems of publication bias – file drawer problem
- Need for the balance of evidence
- Limitations of single studies: time, context, sample specific

Aim of systematic reviews

"attempt to discover the consistencies and account for the variability in similar-appearing studies"

"seeking generalisations also involves seeking the limits and modifiers of generalisations"

"identify the contextual-specificity of available research and evidence"

(Cooper and Hedges, 1994:4).

Key features of a Systematic Review

1. Unbiased: SR covers ALL literature, published and unpublished, no language restrictions

2. Rigorous: SRs use evidence appropriately, and critically appraise all literature included and excluded

3. Transparent: SR methods clear so reproducible

4. Global public good: updated as new evidence emerges

A systematic review is NOT a literature review

Characteristic	Literature reviews	Systematic reviews and meta-analysis
Perspective	Espousal of position	Neutral representation
Coverage	Selective, usually published literature only	Exhaustive (or representative), published and unpublished
Audience	Scholars	Policy makers (and practitioners)
Process	Unclear decision making process	Clear protocol for inclusion or exclusion, data extraction, analysis and reporting
Critical appraisal	At best, study design mentioned only	In-depth assessment of internal & external (and construct) validity
Synthesis method (impacts)	Null-hypothesis significance testing (‘vote counting’)	Synthesis of effect sizes, studies weighted by precision
Study team	Carried out individually	Conducted in teams, double-coding

What makes a systematic review 'systematic'?

1) Scoping: defining answerable question, methods set out in study protocol

2) Comprehensive search to identify published and unpublished studies, in any language

3) Application of study inclusion criteria (PICOS) to determine what gets in*

4) Critical appraisal of study quality, to assess how reliable is the evidence*

5) Data extraction and organisation

6) Synthesis of evidence (outcomes along causal chain)

7) Interpreting results (policy and practice, research recommendations)

8) Improving and updating reviews as new evidence emerges

* Should be conducted by 2 team members working independently

Policy question	Type of review	Example question
Does an intervention work?	Review of effectiveness, drawing on evidence from impact evaluations	What is the effect of community monitoring on education outcomes
What are the important implementation considerations for selected policy options?	Review of implementation, drawing on wider range of evidence, including process evaluations and qualitative studies	What are the barriers and facilitators to effective implementation of SBM?
What are stakeholders' experiences or preferences?	Review of participant experiences and 'views', drawing on qualitative studies and descriptive quantitative studies	Are children, parents and teachers satisfied with the education services provided?
At what cost?	Review of cost-effectiveness and cost-benefit	Cost-effectiveness of de-worming for improving learning outcomes
How do interventions x, y and z compare?	Comparative effectiveness reviews	Relative effectiveness of conditional vs unconditional cash transfers

Theory based systematic reviews

- Draw on methods from traditional reviews of effects
  - appropriate use of evidence is key

- Theory-based systematic reviews:
  1. Traditional effectiveness reviews +
     '+' = analysis of process & implementation
  2. Reviews draw on a program theory (logic model)
  3. Optional: Include additional studies to populate program theory, both quantitative and qualitative

1. Scoping: defining answerable question

- Policy questions often address a ‘big’ issue – how to reduce poverty?
- Need to unpack policy issue into more specific researchable questions that are evaluable

- What is the problem/ issue? For whom? What are the causes/ risk factors?

- What works? How?

- What are the barriers and facilitators of intervention effectiveness?

- Is the intervention acceptable?

The Importance of a Well Formulated Question

The question(s) will guide many aspects of the review process, including:

- determining eligibility criteria
- searching for studies
- critical appraisal of primary studies
- data to be extracted from included studies
- analysis to be undertaken
- presentation of findings

Defining scope means:

- A precise statement of the primary objective, ideally in a single sentence, possibly one or more secondary objectives
- Answerable question: (cost) effectiveness, drivers of change, stakeholder experiences/views, or a combination?
- PICOs
  - What's the relevant population?
  - What's the intervention?
  - If relevant, what's the comparison?
  - What outcomes of interest?
  - What study designs are appropriate to answer the review question?
- Theory of change: how is the intervention supposed to work?

Example: developing scope for EER

The primary objective of this review is to identify, assess and synthesise evidence on the effects of education interventions on children's access to education and learning in low- and middle-income countries.

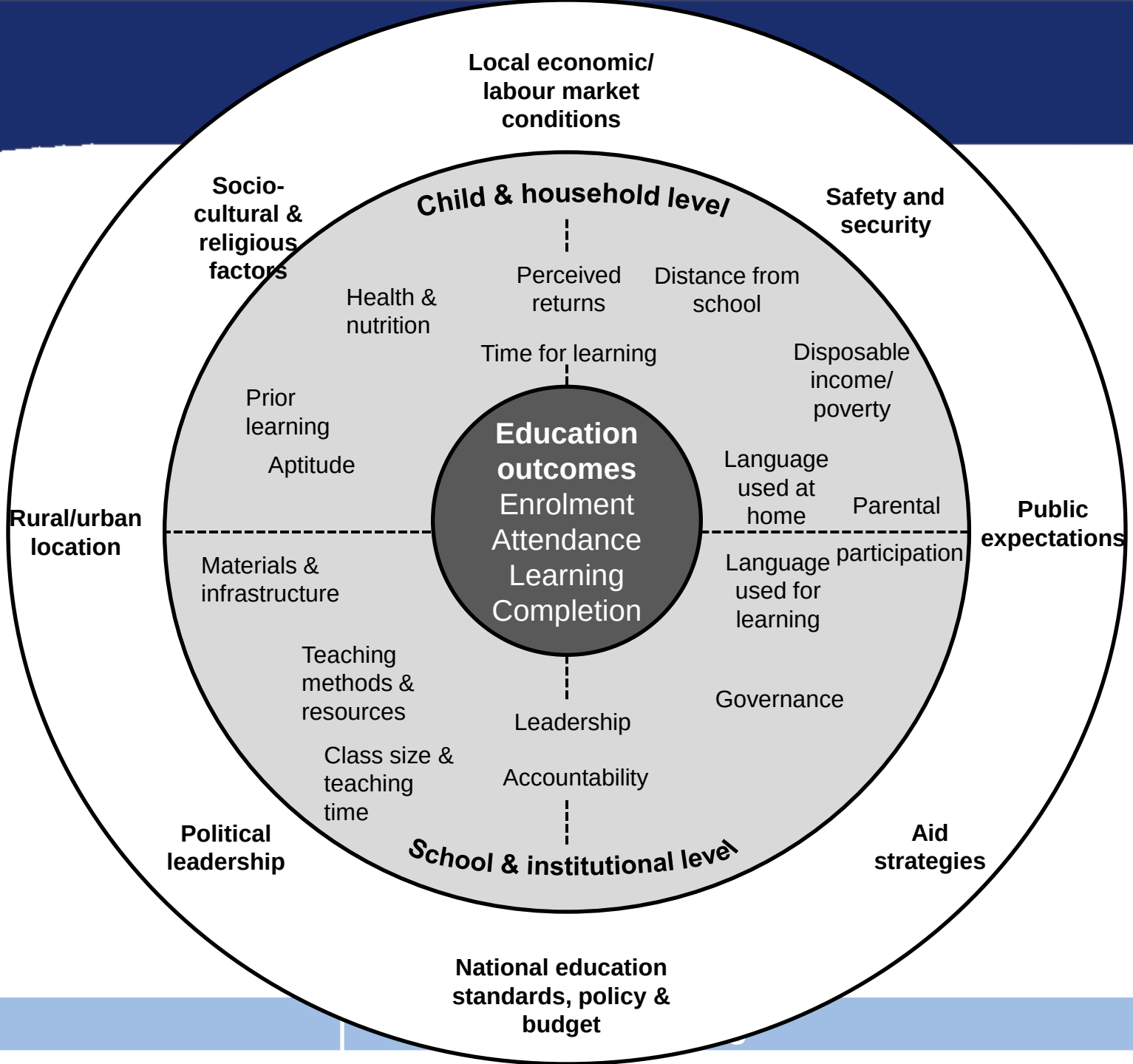
We will also aim to assess how education interventions affect different sub-groups of participants by incorporating sub-group analyses, and will also include a broader range of evidence to address questions relating to process, implementation and costs.

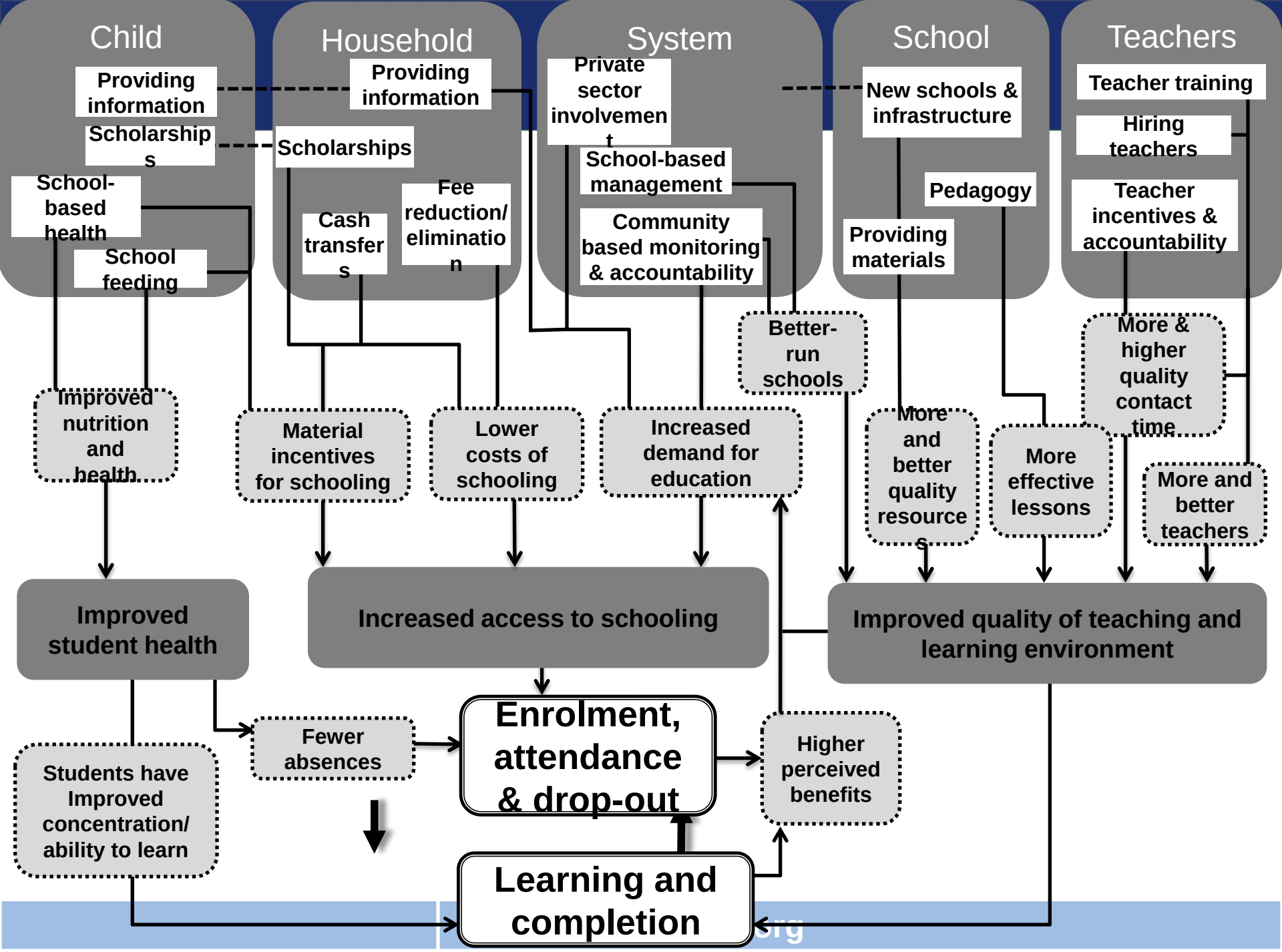
From objectives to research questions

- To achieve these objectives we aim to answer the following questions:
- (1a): What are the effects of different education interventions on enrolment, attendance, dropout rates, completion and learning outcomes for primary and secondary school age children in low-and middle-income countries?
- (1b): How do education interventions affect different sub-groups of participants (according to gender, age, sibling and gender order, urban or rural location, or socio-economic status)?
- (2a): What intervention and implementation features are associated with relative success and failure in improving educational outcomes?
- (2b): What are the contextual barriers to, and facilitators of, the effectiveness of educational interventions?

Summary of inclusion criteria

Study Characteristic	Inclusion criteria
Population	Primary and secondary school age children in mainstream education in LMICs
Intervention	Interventions with primary focus on educational outcomes
Comparison	No intervention, different education intervention
Outcomes	Primary outcomes: enrolment, attendance, drop-out, completion, learning (numeracy, measures of cognitive and problem solving skills, and composite assessment scores); Secondary outcomes: teacher attendance, teacher performance
Study Type	1a and 1b: Experimental studies and quasi-experimental studies 2a and 2b: Studies included to address 1a and 1b + qualitative studies, descriptive quantitative studies, process evaluations, project documents linked to interventions studied in included experimental and quasi-experimental studies
Timeframe	Studies published from 1990 onwards


Local economic/ labour market conditions
Socio-cultural & religious factors
Safety and security
Child & household level
Perceived returns
Distance from school
Health & nutrition
Time for learning
Disposable income/ poverty
Prior learning
Aptitude
Education outcomes
Enrolment
Attendance
Learning
Completion
Language used at home
Parental participation
Rural/urban location
Public expectations
Materials & infrastructure
Language used for learning
Teaching methods & resources
Governance
Leadership
Class size & teaching time
Accountability
School & institutional level
Political leadership
Aid strategies
National education standards, policy & budget


Child
Household
System
School
Teachers
Providing information
Scholarships
School-based health
School feeding
Providing information
Scholarships
Cash transfers
Fee reduction/ elimination
Private sector involvement
School-based management
Community based monitoring & accountability
New schools & infrastructure
Pedagogy
Providing materials
Teacher training
Hiring teachers
Teacher incentives & accountability
Better-run schools
More & higher quality contact time
Improved nutrition and health
Material incentives for schooling
Lower costs of schooling
Increased demand for education
More and better quality resources
More effective lessons
More and better teachers
Improved student health
Increased access to schooling
Improved quality of teaching and learning environment
Fewer absences
Enrolment, attendance & drop-out
Higher perceived benefits
Students have Improved concentration/ ability to learn
Learning and completion

The devil is in the detail…

Intervention level	Intervention type
Child level	School feeding programmes
	School-based health programs
	Providing Information to children
	Merit based scholarships
Household level	Interventions reducing costs: Cash transfers
	Interventions reducing costs: Scholarships and allowances
	Interventions reducing costs: Reducing or eliminating school user fees
	Providing information to parents
School level	Pedagogy interventions
	New schools & infrastructure
	Interventions providing materials
Teacher level	Teacher incentives and accountability
	Teacher training
	Hiring additional teachers
System level	Decentralisation and local community participation: School-based management (SBM)
	Decentralisation and local community participation: Community based monitoring and accountability interventions
	Public private partnerships and private provision of schooling

Defining study designs

Types of study designs (Qs 1a and 1b)

Experimental and quasi-experimental study designs that allow for causal inference. Specifically, we will include:

- Studies where participants are randomly assigned to treatment and comparison group (experimental study designs);
- Studies where assignment to treatment and comparison group is based on other known allocation rules, including a threshold on a continuous variable (regression discontinuity designs) or exogenous geographical variation in the treatment allocation (natural experiments);
- Studies with non-random assignment to treatment and comparison group, provided they include pre-and post-test measures of the outcome variables of interest to ensure equity between groups on the baseline measure, as well as use appropriate methods to control for selection bias and confounding, such as statistical matching or regression adjustment

Exercise 1

As 2015 approaches, the international community is developing the post-2015 education agenda. With enrolment in primary education in developing regions reaching 90 per cent, focusing on getting children into school is no longer enough. The global education consultation process highlighted widespread consensus that good quality education, with a strong focus on learning, is the most important priority for the post-2015 education agenda.

The UN has commissioned a range of different teams to conduct systematic reviews to inform the post-2015 action plan. Each team has to choose a different intervention area that they think has potential to achieve this broader objective and review the evidence to help guide decision-making.

Exercise:

- Define your research objective:
- Propose a systematic review research question(s):
- Define the PICOS (populations, interventions, comparisons, outcomes, study designs):

2. Search:

Range of strategies

- Electronic search engines including general and subject specific academic databases
- Search sources of grey (unpublished literature)
- Search sources of non-English literature
- ‘Hand-search’ organisation web-sites, journals, library shelves
- Literature ‘snowballing’: bibliographies of included studies, forward citation tracking
- Contacting key researchers and intervention agencies

Developing the search strategy

- Identify key concepts, their synonyms (free text) and thesauri and MeSH terms
- Where should truncation characters be used? Eg: school* will retrieve school, schools, schooling
- What limiting features are available to narrow results?(Time period)
- Decide how to combine terms using boolean operators:

- AND: both terms must be present in order for a record to be retrieved, e.g: cash transfer AND Philippines

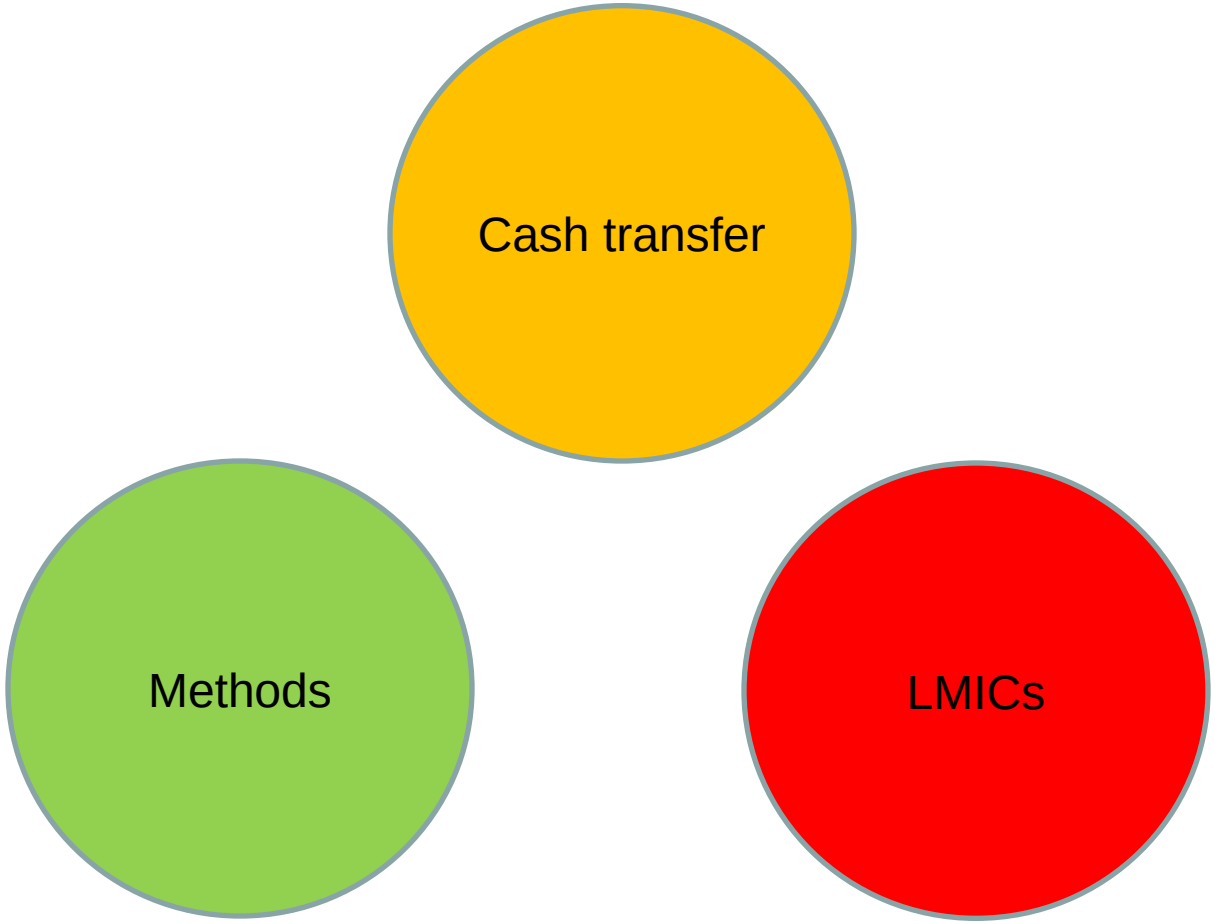
- OR: Either term may be present in order for a record to be retrieved (used to search for related terms or synonyms), e.g: India OR China OR Philippines

- NOT: Used between two terms to ensure that the second term will not appear in any of the results. (NB! Be careful!), e.g. education NOT adult

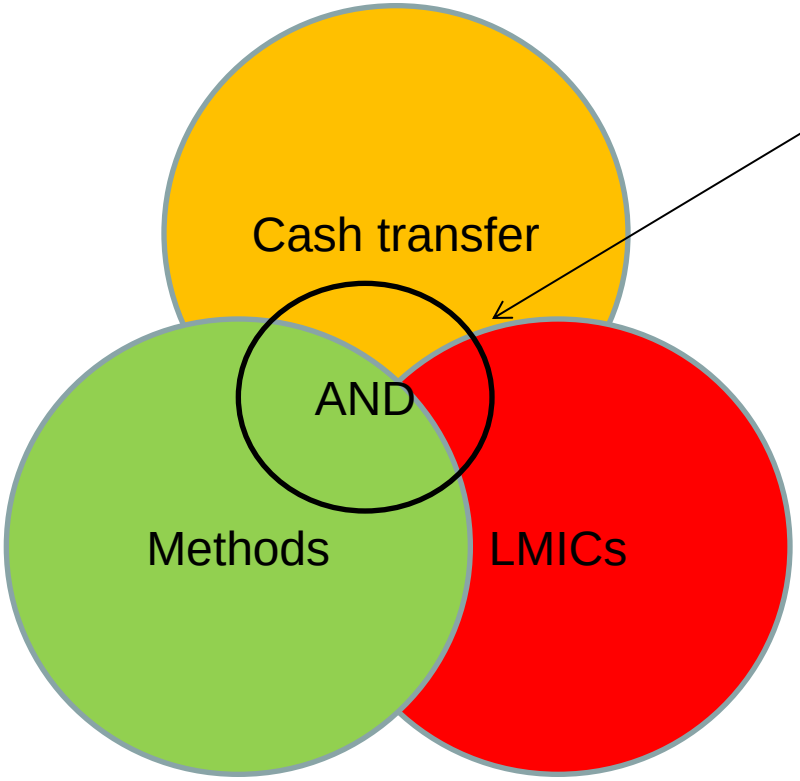
- NEAR: Used to retrieve records where terms appear in proximity to each other, eg: Cash NEAR/3 transfer will retrieve studies where cash appear within 3 words of transfer

- ADJ: Used to retrieve records where terms appear adjacent to each other, eg: cash ADJ transfer will only retrieve studies where transfer is placed immediately next to cash

Key search concepts



Key search concepts



Studies with key terms in title/ abstract

Example strategy

(scholarship* OR subsid* OR stipend* OR grant* OR donation OR bursary OR bursaries OR “tuition relief” OR “user payment*” OR “merit aid” OR “merit based aid” OR “merit-based aid” OR “merit award”)

Use AND to combine with LMICs terms to get any study of interventions with any of the LMIC terms in title or abstract (set to search these fields)

Academic databases EER

Africa Wide: http://www.ebscohost.com/academic/africa-wide-information
Academic Search Premier: http://www.ebscohost.com/academic/academic-search-premier
Applied Social Science Index and Abstracts (ASSIA): www.csa.com/factsheets/assia-set-c.php
CAB Abstracts
Econlit
Education Resources Information Center (ERIC)
International Bibliography of the Social Sciences (IBSS): http://search.proquest.com/ibss?accountid=149134
PAIS International (Public Administration Information Systems)
PsycInfo
Sociofile/SocIndex
Sociological Abstracts: http://search.proquest.com/socabs
Web of Science: Social Science Citation Index (SSCI) and Arts & Humanities Citation Index (AHCI): http://ip-science.thomsonreuters.com/cgi-bin/jrnlst/jloptions.cgi?PC=SS
Worldwide Political Science Abstracts
Health databases to search only using health terms:
Global Health (CABI) (only school feeding and health terms)
Embase (only school feeding and health terms)
Medline (only school feeding and health terms)

Electronic libraries and registries of impact evaluations:

- AEA (American Economic Association) RCT Registry
- British Library of Development Studies (BLDS): http://blds.ids.ac.uk/
- JOLIS (Joint Libraries of the World Bank and IMF): http://external.worldbankimflib.org/external.htm
- 3ie Register of Impact Evaluation Published Studies: http://www.3ieimpact.org/en/evidence/impact-evaluations/
- 3ie RIDIE (Registry for International Development Impact Evaluations): http://ridie.3ieimpact.org/
- 3ie Systematic Reviews Database
- EPPI-Centre Evidence Library
- Campbell Library
- Cochrane Library (only health terms)

Reducing ‘file drawer effect’:

- **Dissertations:** Dissertations & Theses Database (Proquest); British Library Electronic Theses online Service (EtHOS):
- **Libraries of grey lit:** Open Grey, ELDIS
- **Research organisations and networks:** IPA, J-PAL, NBER (National Bureau of Economic Research), Bureau for Research and Economic Analysis of Development (BREAD)
- **Conference proceedings:** Proceedings for past American Economic Association (AEA), SSCI Conference proceedings
- **Relevant NGOs:**
- Government agencies, bilateral and international donors and implementing agencies
- Write to experts and organisations to identify relevant research
- Citation tracking: SSCI & Google Scholar
- Listservs

Search management

- Export the results to bibliographic management software (RefWorks, Reference Manager, EndNotes, Eppi Reviewer)
- Add Source code for each database searched (e.g ERIC1, PsycINFO1…)
- Remove duplicates
- Compile a Search History document listing the original search strategies

3. Application of study inclusion criteria (PICOS)

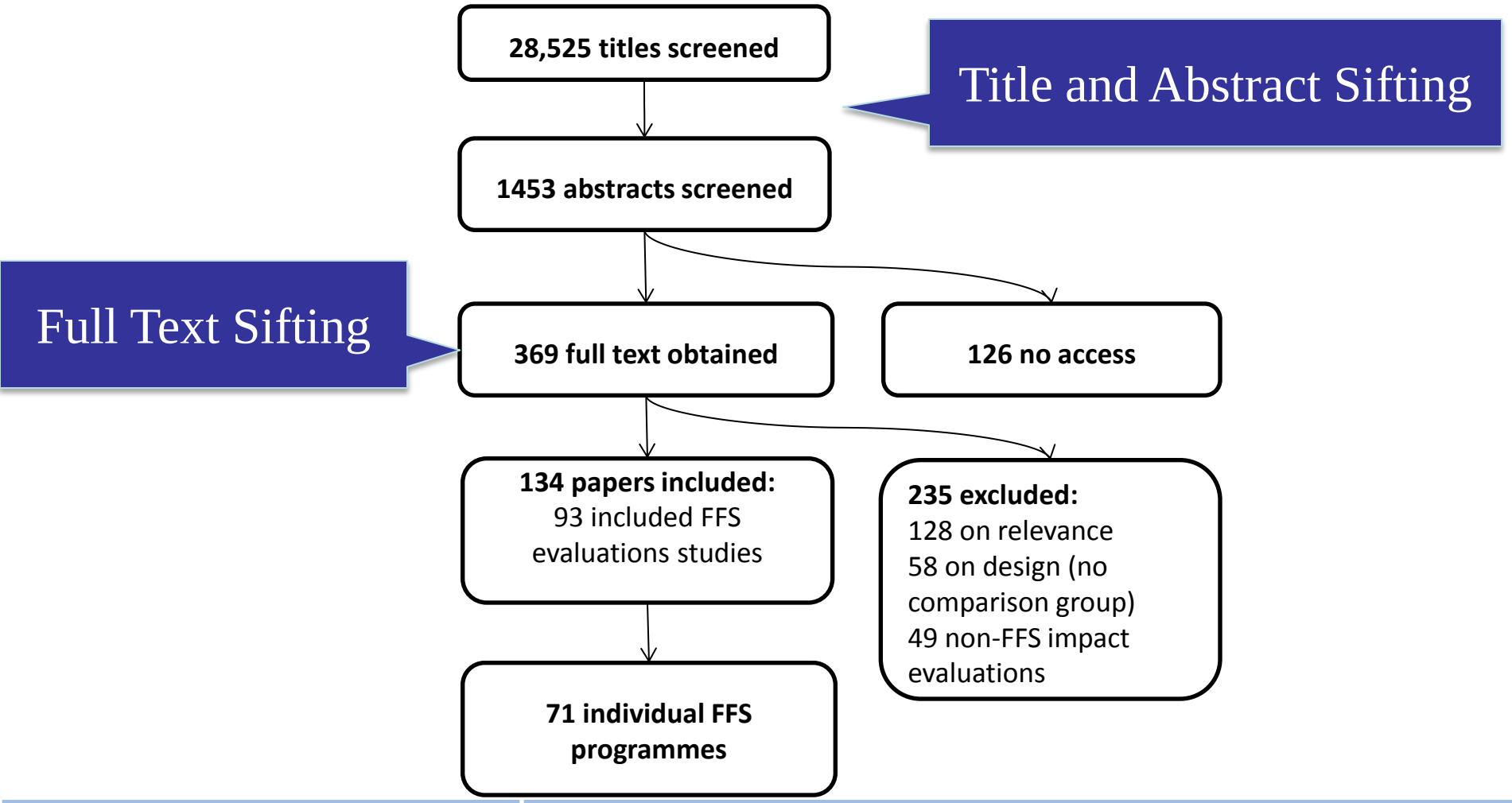
Filter evidence based on PICOS

- Over-inclusive at title and abstract
- Inclusion decisions by 2 people

3ie's experience:

- Consistently find 'treasures' of evidence we didn't know existed
- Rigorous evidence remains limited for many interventions (problem of 'empty reviews' if PICOS set too narrowly)
- Poor abstracts means high number of full-texts screened

Sifting and Developing a Flow Chart



4. Critical Appraisal

- Critical appraisal is an *essential part* of a systematic review
- Separates out high quality from low quality studies
- Uses *explicit and transparent criteria* for including and excluding primary studies for review
- These criteria *vary according to the methods of primary studies*
- Focus on conduct of study rather than design

Study validity: impact evaluations

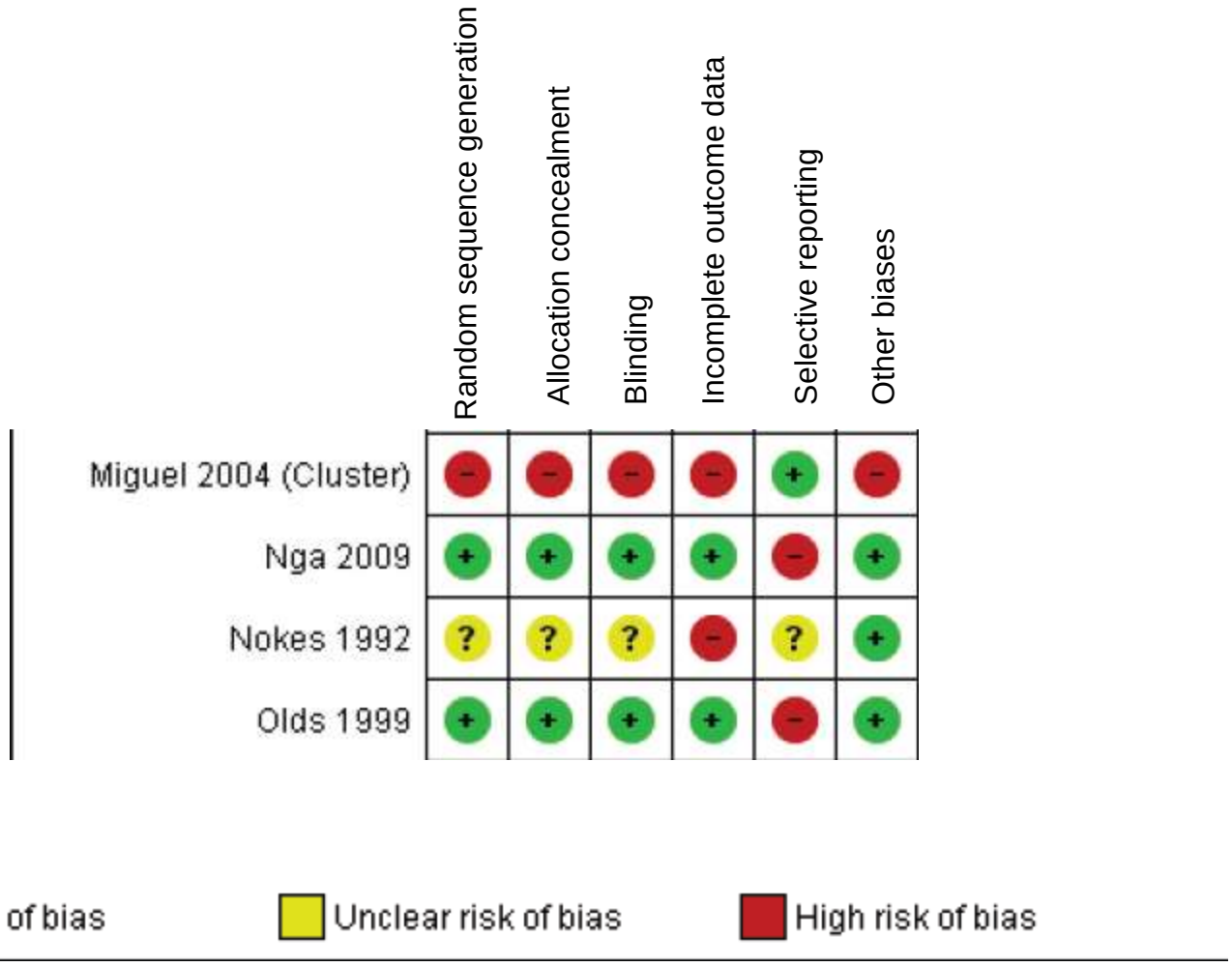
- Internal validity
  - Concerns the validity of the causal relationship estimated
- Statistical conclusion validity
  - Concerns the assumptions, estimation and calculation of effect estimates
- External validity
  - Concerns the generalizability of causal claims from study sample to broader population
- Construct validity
  - Concerns relevance of relationships measured to broader relationships of interest (treatment and outcome measures)

Source: Shadish, Cook & Campbell 2001

Risk of Bias

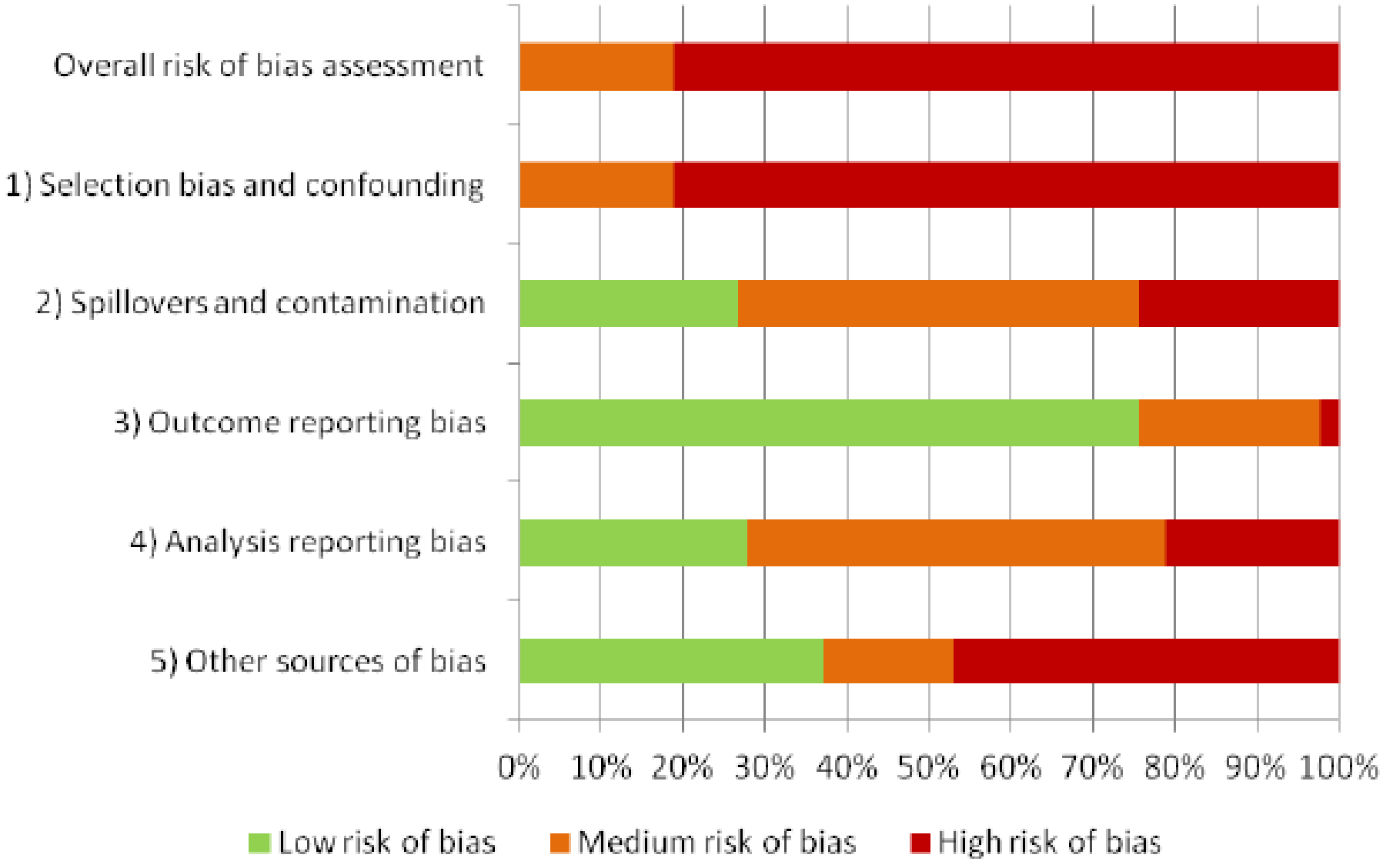
1. Baseline confounding and selection bias: was the allocation or identification mechanism able to control for baseline confounding and sample selection bias (censored data)?
2. Time-varying confounding: was the method of analysis executed adequately to ensure comparability of groups throughout the study?
3. Bias due to missing data: is the estimation method sensitive to non-random attrition?
4. Biases in outcome data collection: did the process of being observed cause motivation bias (Hawthorne and John Henry effects, courtesy bias, recall bias)?
5. Departures from intended interventions: was the study adequately protected against performance bias and survey effects?
6. Outcome & analysis reporting biases: was the study free from outcome reporting bias and analysis reporting bias?

Risk of bias: study by study results



Source: Garner et al. 2012

Risk of bias table: overall body of evidence



External Validity of studies

Questions to consider:

- Does the study have relevance to the 'real world' in which you are working?

- Is the sample in the study/report similar to the population you are dealing with?

- Are the outcomes in the study/report the same outcomes that are of interest to your work?

- Are there any contextual factors mentioned in the study/report that would limit its relevance to your work?

- Does the study examine efficacy (ideal conditions) or effectiveness (real world) of intervention?

5. Data extraction

- Coding protocol: essential feature of systematic review
- Goal: transparent and replicable, description of studies and extraction of findings
- Forms should be part of protocol and developed a-priori (although can be changed as long as recorded)
- Ideally independent extraction of data by 2 researchers, in particular for data to be used to calculate effect sizes and their variance
- Use excel, Eppi reviewer, Access, File maker or similar to organise data

Types of data:

Three main categories:

(1) descriptive data on study design, publication type, intervention and context for purposes of descriptive analysis of the body of research;

(2) data on the population, context, study design, intervention design and process and implementation for purposes of moderator analysis and qualitative synthesis (addressing question 2a and 2b);

(3) data on outcomes and sample for purposes of effect size calculation.

Typically multi-level: Study level, outcome level, effect size level

Exercise 2

Which key concepts would you combine for the search strategy?

Can you think of some associated terms that would be part of the search strategy?

What types of data would you need to collect?

Statistical procedures and synthesis

Additional features to detail in the protocol:

- How to calculate effect sizes
- How to deal with dependent effect sizes
- How to deal with missing data
- Methods of synthesis: decision rule, statistical model
- Analysis of heterogeneity
- Sensitivity analysis
- Publication bias

Useful resources

- MEC2IR: http://www.campbellcollaboration.org/artman2/uploads/1/Draft_revised_MEC2IR_reporting_standards_1.docx

http://www.campbellcollaboration.org/artman2/uploads/1/Draft_revised_MEC2IR_conduct_standards_1.docx

- Campbell protocol and review guidelines: http://www.campbellcollaboration.org/lib/project/328/
- Campbell training videos: http://www.campbellcollaboration.org/resources/training.php
- JDeff special issue on SRs (open access): http://www.tandfonline.com/toc/rjde20/4/3#.VAV11PldVc8
- Cochrane Handbook: http://www.cochrane.org/resources/handbook/
- David Wilson's effect size calculator: http://mason.gmu.edu/~dwilsonb/ma.html
- Michael Borenstein Introduction to Meta-Analysis: http://onlinelibrary.wiley.com/book/10.1002/9780470743386
- bsnilstveit@3ieimpact.org OR mvojtkova@3ieimpact.org

Thanks!

Please visit our website: www.3ieimpact.org