planning, monitoring and evaluation

Department:
Planning, Monitoring and Evaluation
REPUBLIC OF SOUTH AFRICA

Africa Centre for Evidence

Working together to make evidence-informed decision-making a reality

UNIVERSITY OF JOHANNESBURG

EARLY GRADE MATHEMATICS STUDY EVIDENCE MAP

A rapid internal analysis report

October 2018

ANALYSIS HIGHLIGHTS

- Having searched almost 6,000 citations, the evidence map includes 161 studies local and international studies investigating EGM teaching and learning. This includes 127 studies from South Africa.
- The large majority of the evidence-base is comprised of diagnostic studies (n=98) with impact evaluations contributing 41 studies and evidence of implementation and design presenting the minority of the evidence (n=12). To underline this trend that the majority of the evidence is descriptive rather than evaluative, the most frequent study design identified refers to descriptive assessments (n=38).
- In terms of the quality of the identified evidence-base, the majority of the studies are of poor or very poor quality (52%). Only, 24% of the included studies are rated as of high quality.
- The most researched EMG interventions identified in the evidence-base refer to: pedagogy, cognitive inputs, teacher content knowledge, learner & teacher support material, accountability and motivation.
- Lack of parents / community centred studies as well as subsequent interventions targeted at parents / communalities
- Overall, the investigated EGM interventions had short implementation periods with two-thirds of the interventions being applied for less than one year.

1. INTRODUCTION TO THE REPORT

This report presents a rapid internal analysis of the Early-grade Mathematic Study (EGMS) Evidence Map. The evidence map and this rapid internal report were co-produced by researchers at the Department for Planning, Monitoring and Evaluation (DPME) and the Africa Centre for Evidence (ACE). This report aims to provide a description of the evidence-base contained in the produced evidence map highlighting key patterns and features in the size and nature of the available evidence-base. The report serves as a compliment to the actual evidence map which can be accessed through DPME's research unit. The evidence map provides a more holistic and interactive way to engage with the available evidence-base and can be tailored to users' own preferences. The main aim of this report is to communicate key descriptive findings on the available evidence-base at a glance. It presents an aggregate view of the evidence-base and more granular analyses can be requested from the research team or be conducted by using the evidence map itself. This report is intended for internal use by DPME staff.

From January until August 2018 DPME conducted an evidence map of research on early-grade mathematics in South Africa. This evidence map was conducted in partnership with ACE and the Department of Basic Education (DBE). The project entitled 'EVIDENCE MAPPING: Educational interventions to support the teaching and learning of early grade mathematics in SA' was part of a wider effort to evaluate and build evidence on effective interventions to improve the teaching and learning of early grade mathematics in South Africa. Under this larger body of work, DBE and DPME have conducted and aim to continue to conduct a range of impact evaluations to understand what works in improving early-grade reading and mathematics outcomes in the country. Following the encouraging results of an early-grade *reading* impact study in 2017, an early-grade *mathematics* impact study is scheduled for 2019. The aim of the evidence mapping project was to better understand what types of early-grade *mathematics* programmes to pilot and to assess in the impact study. That is, the evidence mapping project aimed to inform the future primary impact evaluation study. The evidence map thus had three core objectives:

(1) To undertake a scoping study to identify and recommend policy interventions which impact on the teaching and learning of early grade numeracy with strong theories of change, as well as cost structures that would be sustainable on a large scale.
(2) To develop a responsive evidence-base to support long-term planning and policy development in the area of early-grade mathematics teaching and learning in South Africa.
(3) To promote the use of evidence synthesis methodologies as part of decision-makers' toolkit for evidence-based policy-making.

2. RESEARCH METHOD AND PROCESS

Evidence mapping is a research methodology part of the family of methods for evidence synthesis, e.g. systematic reviews, meta-analysis, rapid evidence assessments. Since 2015, DPME has developed its own tailored methodology for evidence mapping and applied this methodology in four broad policy areas to date. The Department has applied the evidence maps as an interactive knowledge management tool to facilitate the development, engagement and use of an evidence-base for decision-making. DPME's evidence maps are based on gold standard research processes as applied in systematic reviews to systematically and transparently search and collate an evidence-base, organise

and appraise this evidence, and to visualise and support its use through an interactive online interface and knowledge management platform. More information on DPME's evidence mapping methodology can be found in the Departmental Guidance Note on evidence mapping (DPME 2016) as well as in Appendix A of this report.

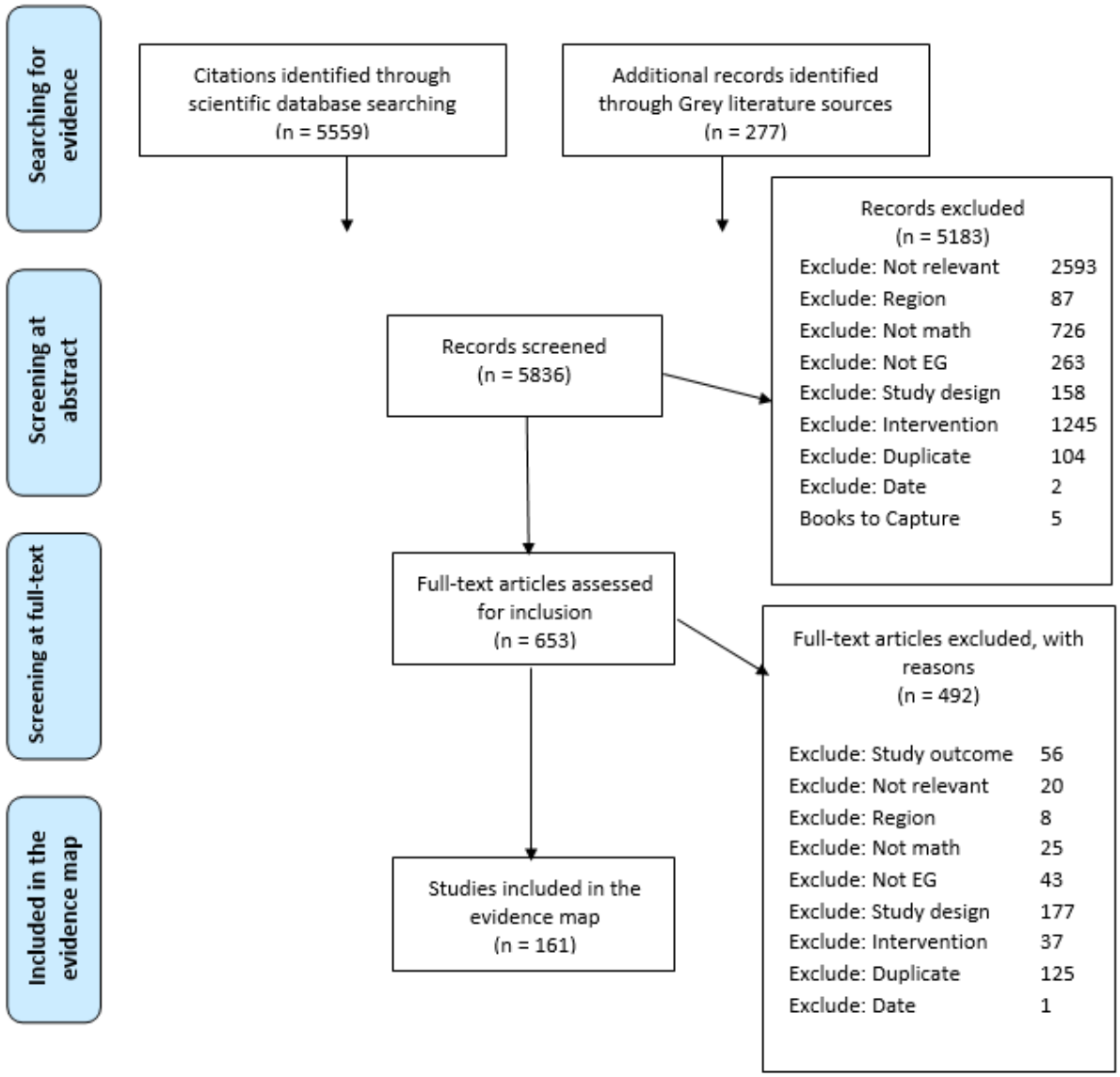
The evidence mapping research project was co-produced between DPME and ACE. The same applies to this report. As part of the co-production process researchers within DPME and ACE are match-made and engage on research- and policy engagement- steps jointly. As the evidence mapping process requires both research and policy engagement activities, the co-production approach is an essential component of the methodology.

This report presents a rapid internal analysis of the data set contained in the full evidence map. It borrows from rapid response approaches to informing policy decision-making (Mijumbi 2017). ACE and DPME researchers analysed the data within the evidence map over the course of 3 working day including research activities and joint discussion. The data was accessed through DPME's back-end capture platform of the evidence map and supplemented with ACE's systematic review software. Frequencies and cross-tabulation of the *aggregate* data contained in the evidence-base were conducted to inform this report. As a result, this rapid internal report only provides a high-level summary and the size and nature of the collated early-grade mathematics evidence-base.

3. DESCRIPTION OF THE EVIDENCE-BASE

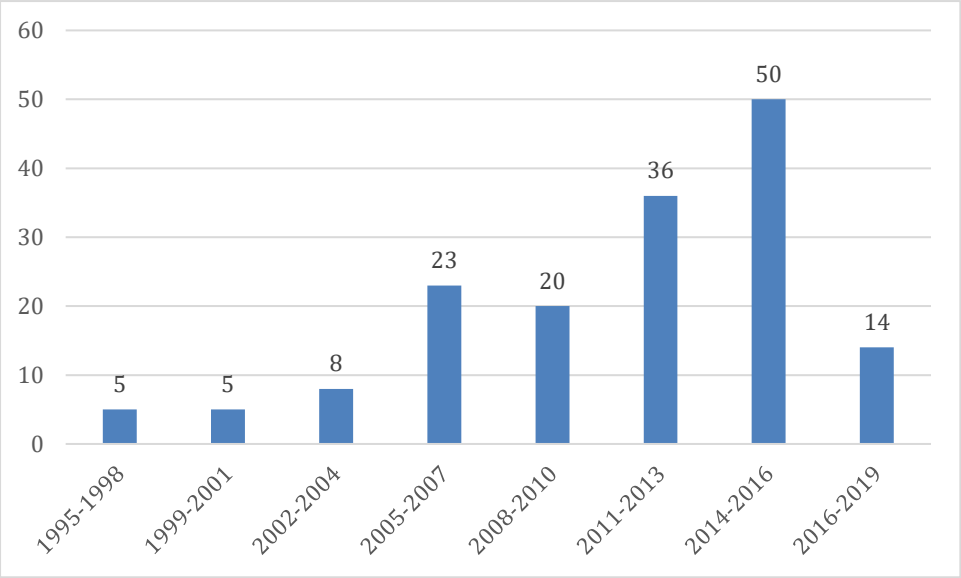
The PRISMA flow chart presented in Figure 1 indicates that a total of 5,836 records were screened at an abstract level with 5,559 of results attributed to formal/academic searches and 277 grey literature searches. Of the 5,836 only 653 of the studies made it to full-text screening. Reasons of exclude at an abstract level were mostly: Not relevant citation (n=2,593), intervention (n=2,145), and not mathematics outcomes (n=726). Of the 653 studies screened at full-text, **161 of the studies were included in the evidence map**. The predominant reasons for exclude at full-text were: not relevant study design (n=177) and duplicate (n=125). The reason for high numbers of Exclude: Study design is attributed to the inclusion criteria being strictly counter-factual based impact evaluations for international studies. The reason for high numbers of Exclude: Duplicates is that various overlapping databases were consulted within the formalised search strategies employed, hence, a large number of duplicates were found.

Figure 1 Prisma Flow Chart



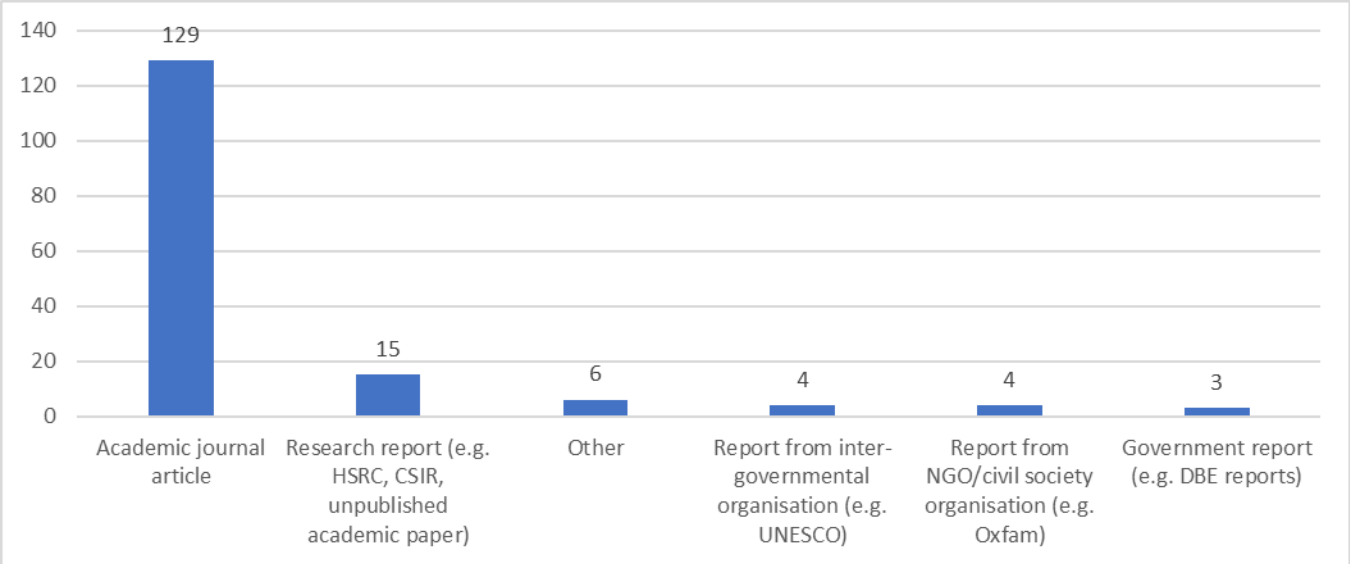
Of the 161 included studies, 44 were grey literature studies with the remaining 117 studies coming from formal/academic searches. We included studies from the year 1990 onwards. The earliest included study was published in 1995, thus there were no studies included prior to 1995. Figure 2 below illustrates the distribution of research over three year periods beginning at 1995. The figure illustrates an increasing trend from the year 2001, peaking between the period 2012-2015, where 45 studies included in the map were published. This is a stark contrast to the 5 studies included from 1995-1998. The small number of studies in the 2017-2019 category is explained by the search period for the evidence map. Our searches were conducted in February 2018 leaving studies published in the remainder of 2018 and 2019 outside the scope of the evidence map.

Figure 2 Publication Distribution



Amongst the type of research, academic journal articles clearly dominate as illustrated in Figure 3 with 129 studies. This was followed by research reports with 15 studies. The six studies captured as 'other' were mainly PhD theses. The remainder of the studies were different types of reports clustered between 3-4 studies each.

Figure 3 Types of Research

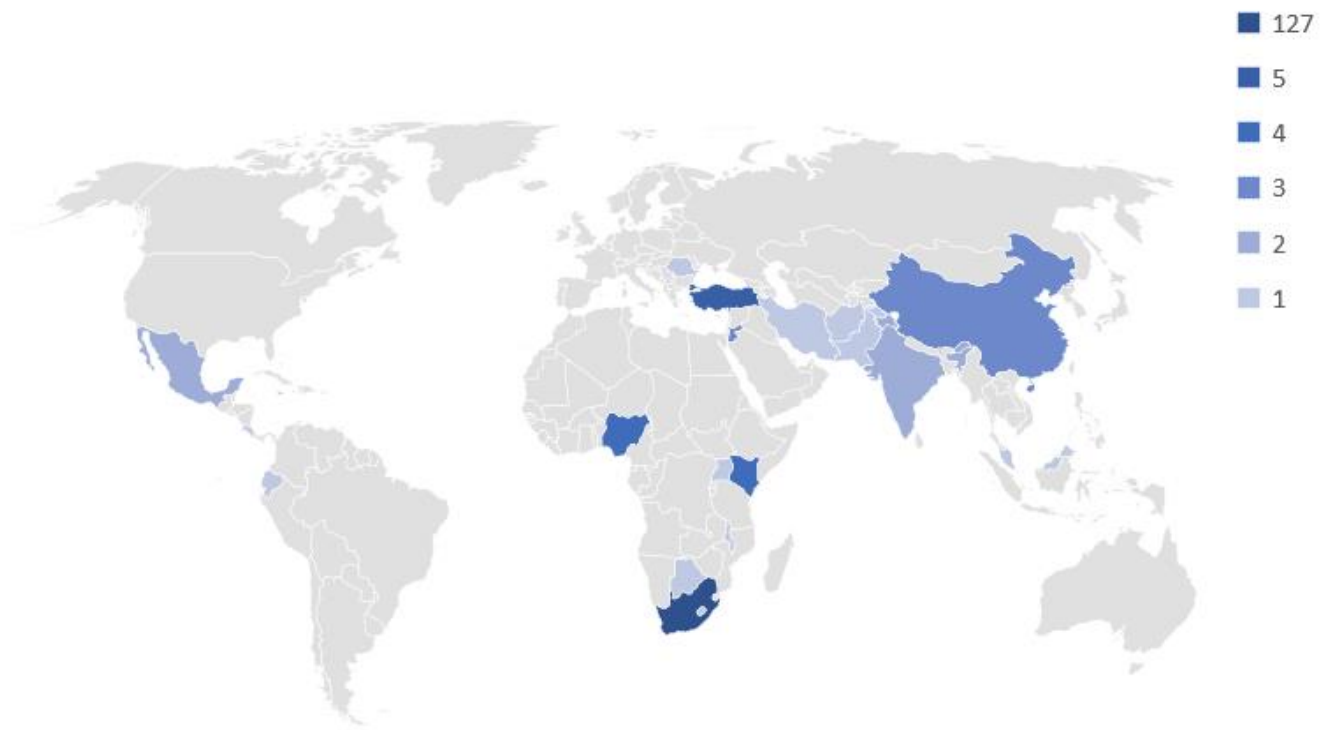


4. NATURE OF EVIDENCE

Country and Region

Figure 4 below provides the geographical distribution of the evidence included in the EGMS map. Given the more inclusive inclusion criteria for South African studies, the evidence on EGM is concentrated in the southern part of the African continent, South Africa (127/161) in particular, and very sparsely distributed across the African region. Except for a few studies drawn from Nigeria (4/161) , Kenya (4/161) and Uganda(1/161), Malawi(1/161), Lesotho(1/161), and Botswana (1/161), there exist scant evidence on EGM across the whole African continent. Though also limited, the other evidence emanates from the Asian and Pacific region including three studies each from China and Jordan, two studies from India, and one study from Malaysia, Jamaica, Iran, Pakistan and Afghanistan each. A small number of studies are drawn from two European countries, namely Turkey (5/161) and Romania (1/161), with evidence from Turkey surpassing any of the evidence from any single country in Asian and Pacific region. Lastly, there is evidence of EGMS that emerges from the Latin American region including two studies from Mexico, and one study from Costa Rica and Ecuador each.

Figure 4 Geo-map of the included evidence-base



Type of the evidence

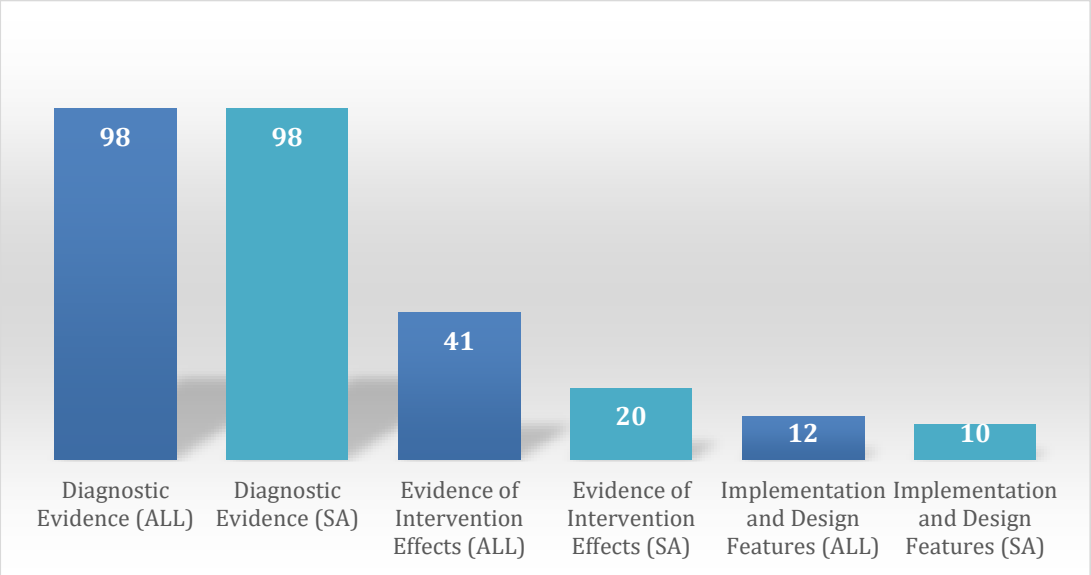
Figure 5 illustrates the types of evidence for all studies and for South African studies. The type of evidence included in the map were categorised into 3 groups: evidence of interventions effects, implementation and design evidence, and diagnostic evidence. Figure 5 is broken down by all studies and studies conducted in South Africa. The reason for sorting studies in this way is due to utilising

different inclusion criteria for South African studies and international studies. Implementation and design evidence and diagnostic evidence type of studies were included only for South African studies. The overall overview is that majority of the studies are diagnostic evidence (n-98), followed by Evidence of intervention effect type studies (41) and lastly, Implementation and design feature types of evidence with 12 studies.

Upon comparing South African studies to all studies, there is only one major discrepancy, which expectedly lies with evidence of intervention effects. Less than half of the evidence of intervention effects were conducted in South Africa. This is in stark contrast to the 98 diagnostic evidence studies. It provides the perception that studies in South Africa are focussed more on identifying underlying barriers or contextual factors to EGM teaching and learning as opposed to evaluating the impact of EGM interventions.

Implementation and design feature evidence provides interesting insights into the international studies. Only 2 of the 21 evidence of intervention effect studies for international countries incorporated empirical evidence on implementation and design features.

Figure 5 Type of the evidence

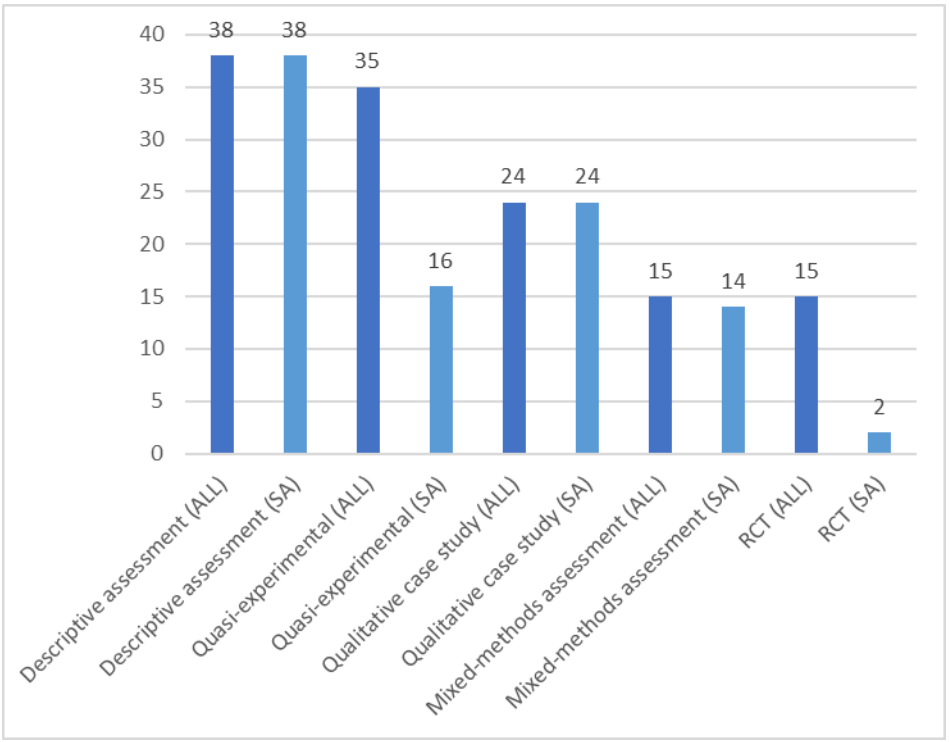


Study methodology

An analysis of the type of methods used in the mapped studies indicates that majority of the studies utilised descriptive assessments followed by quasi-experimental design (Figure 6). The other common methods are qualitative case studies, mixed-methods assessments and RCTs. The full list of methods used can be found in Table 1 in the Appendices. It is worth noting that amongst the least utilised methods are natural experiments, regression discontinuity design, qualitative evaluations and comparative case studies.

Figure 6 also provides the types of methods in South Africa. Given the scale of studies from South Africa, there is very little differentiation between the most popular methods overall and the most popular methods for South African studies. However, RCTs and quasi-experimental studies indicate large discrepancies: only two RCTs were identified in South Africa out of a total of 15 for all countries and 16 Quasi-experimental studies out of a total of 35 studies for all countries.

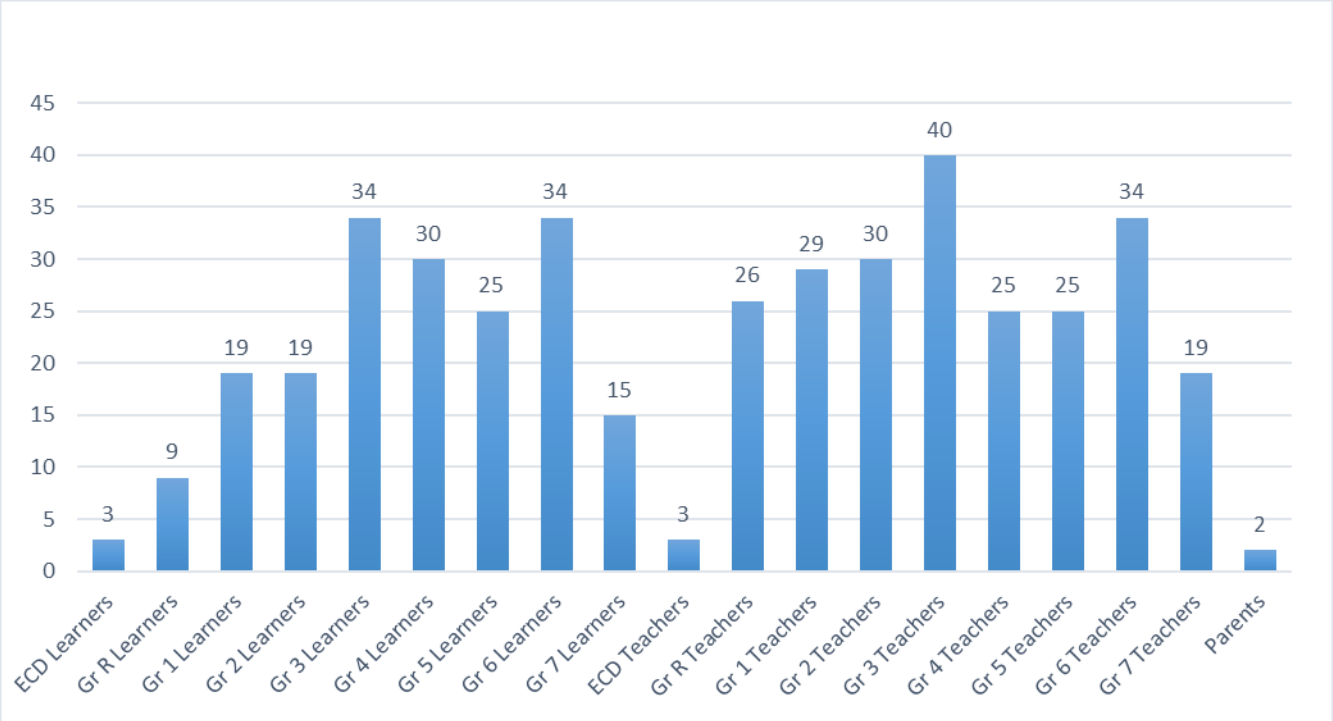
Figure 2 Study methods



Target group

The main targeted group in EGM studies is Grade 3 and 6, for both learners and teachers as shown in Figure 8. This might be related to the previous annual national assessments (ANA) taking place in these grades and authors making use of this dataset. There is minimum focus on ECD mathematics in EGM studies. Interestingly, in Grade R,1 and 2 most of the evidence shows skewness towards teachers rather than learners themselves.

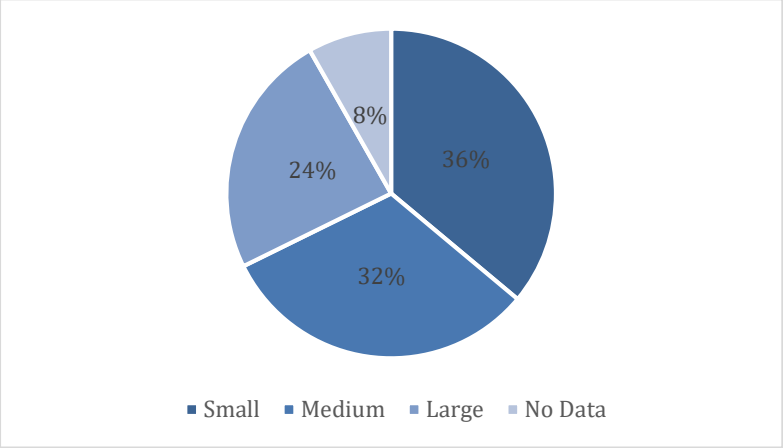
Figure 3 Target group



Scale of the research

A third of the studies included in the map were of small scale and medium scale each (Figure 9) . Only a quarter of the included studies could be described as large scale. A small percentage of studies (9%) did not report the scale of investigation.

Figure 4 Scale of the research

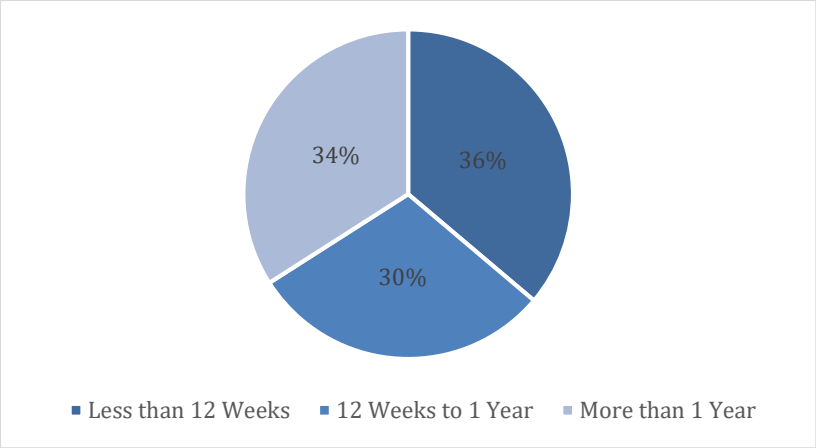


Intervention implementation length

Although majority of the studies have not reported on implementation lengths, of those that have, Figure 10 below illustrates that implementation lengths are evenly distributed between implementation less than 12 weeks (36%), implementation between 12 weeks and 1 year (30%) and

implementation of more than 1 year (34%). However, it is worth mentioning that Figure 10 represents only a subsample of the total includes as the majority of studies have not reported on implementation lengths. A possible explanation of the limited studies reporting on implementation lengths could be in relation to a substantial amount of the studies part of the evidence map were diagnostic in nature.

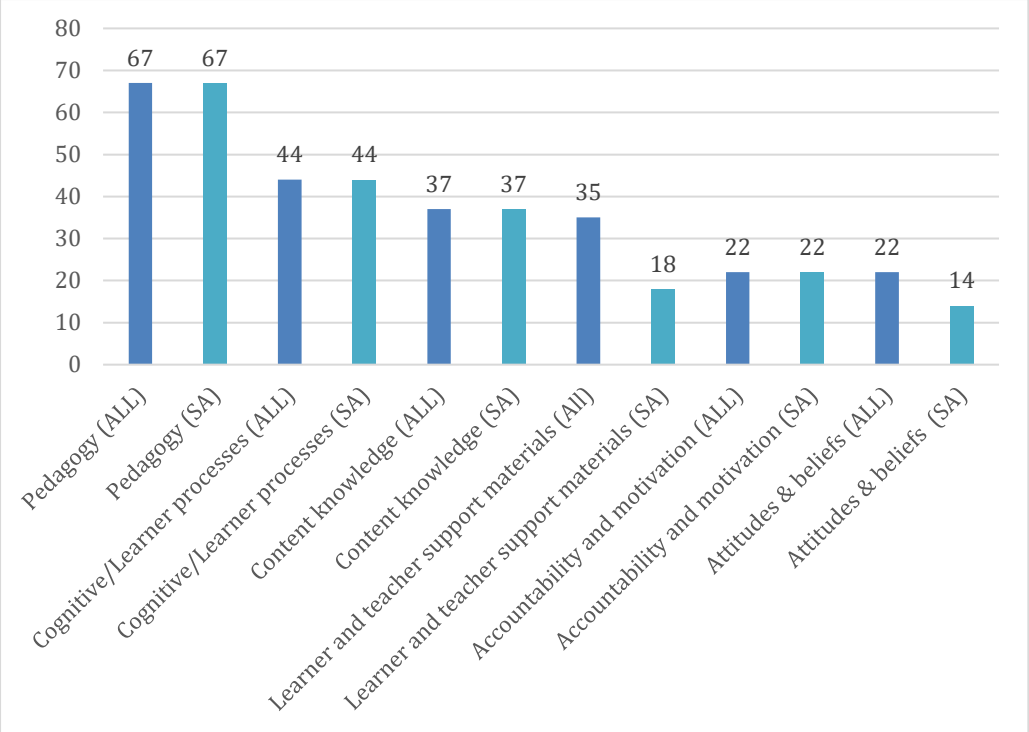
Figure 10 Implementation length



5. PATTERNS IN THE EVIDENCE-BASE

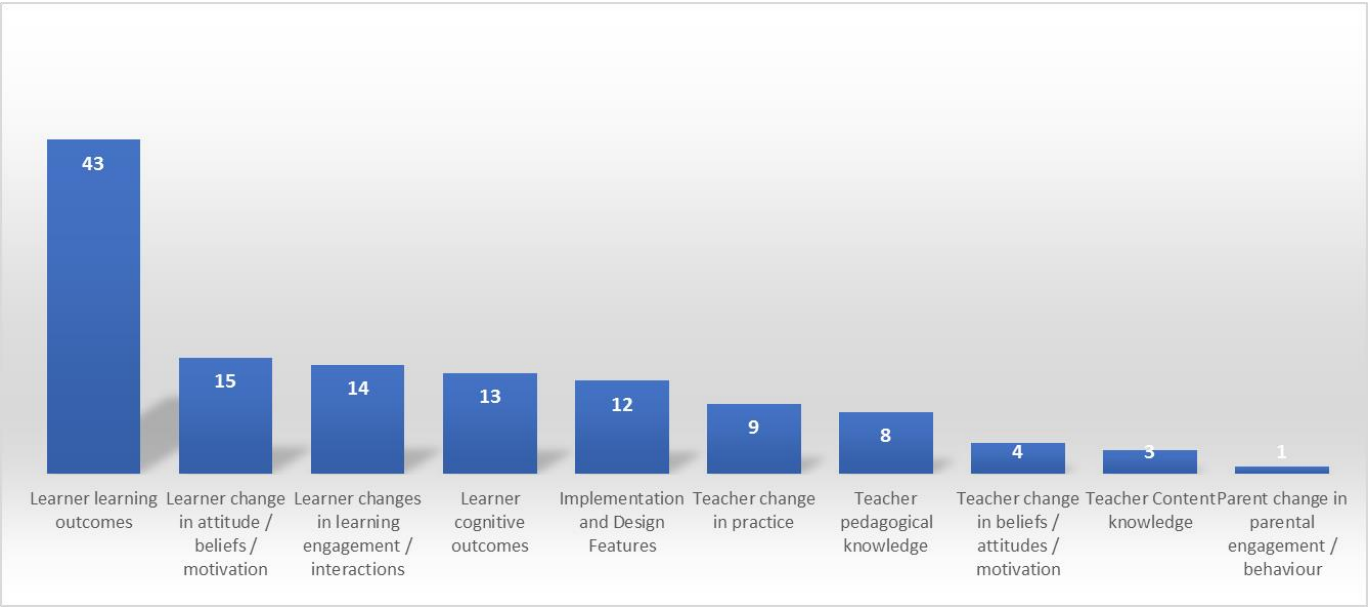
In total, our evidence map included 416 interventions that are categorised into 27 groups. Of the 27 intervention groups, Figure 11 tabulates the Top 6 most popular interventions of the studies included in the map. The most frequently assessed interventions are pedagogical interventions (n=67) and there is quite a big gap until the second most frequently reported intervention, which is comprised of a cluster of intervention groups that range from 44 to 35 studies. The rest of the intervention grouping on the top 6 list fall within the 37 to 22 studies range. The remaining interventions can be found in Table 2 in the Appendices. Additionally, we illustrate the interventions in South Africa relative to the overall number of intervention groups. Considering that a vast number of studies are in South Africa, it is unsurprising that there is very little variance between the two depictions in the graph.

Figure 5 Top 6 interventions



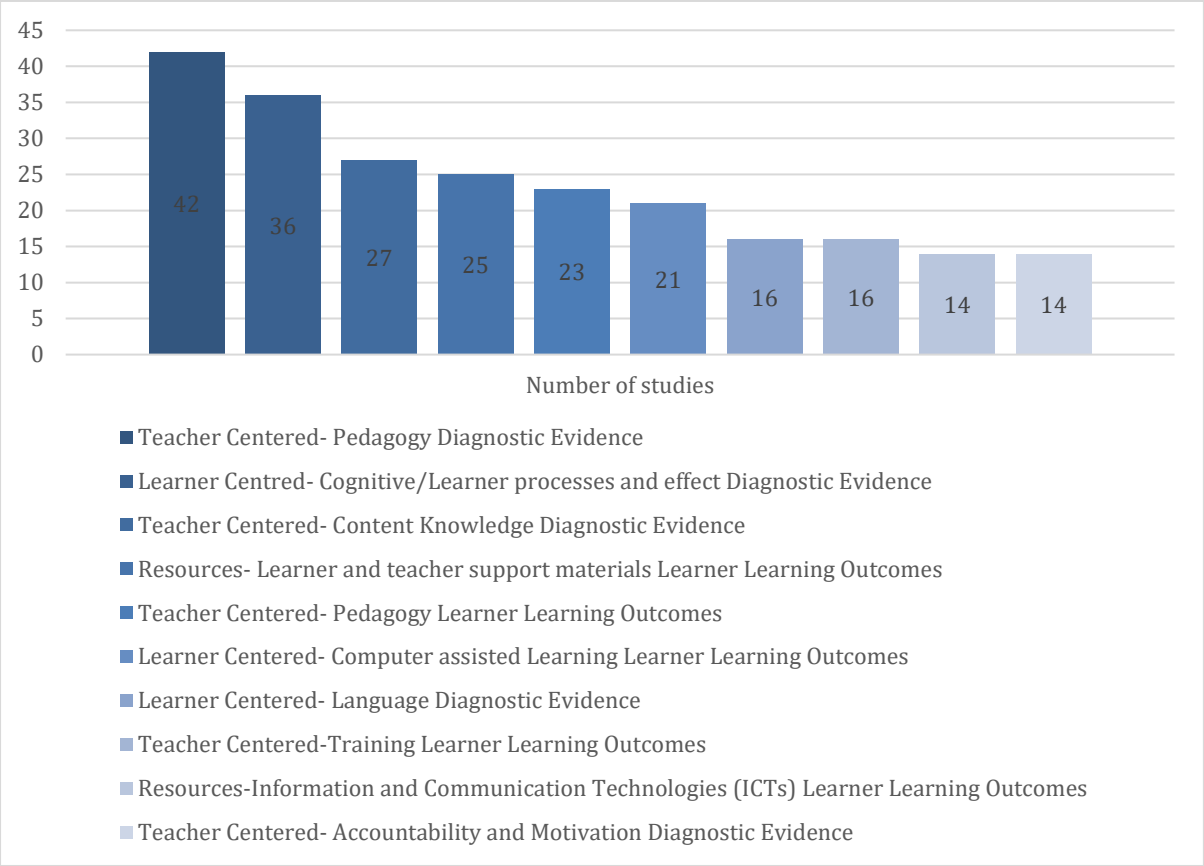
The most popular outcomes are illustrated in Figure 12 below. Overall, most studies had focussed on learner outcomes specifically. Of the learner centred outcomes, the most common was learning outcomes, reported in 43 studies. The remaining of learner centred outcomes lies between 15 to 13 studies. Teacher-related outcomes range between 9 and 3 studies in terms of frequency of reporting and this indicates that learner centred outcomes were reported more than teacher centred outcomes.

Figure 6All outcomes



In addition to assessing the frequency of interventions and outcomes, we further investigated particular configurations of interventions and outcomes. These are reported in Figure 13 which represents the top 10 configurations of interventions and outcomes reported in the included studies. The most prominent configurations were Teacher-Centered: Pedagogy and Learner-Centred: Cognitive/Learner processes and effect, both of which are assessed as Diagnostic Evidence. This was then followed by clusters of four configurations ranging between 27-21 studies and 16-14 studies. Interestingly, half of interventions within the illustrated configurations were teacher-centered, with learner-centered reported in three configurations and resources in two. In terms of outcomes, Learner learning outcomes (n=5) and Diagnostic evidence (n=5) were the only two reported amongst the 10 most frequent intervention/outcome configurations.

Figure 7 Configuration of interventions and outcomes



Next, we investigated evidence gaps in the available evidence-base for both interventions and outcomes. In order to identify gaps within interventions a criteria was set of a minimum of 5 studies per intervention. The following gaps were identified:

- Learner-centered-- Health and Nutrition (n=2)
- Parents / Care-givers / Community Centred-- Involvement (n=4)
- Parents / Care-givers / Community Centred-- Context (n=3)
- Parents / Care-givers / Community Centred-- Language (n=0)
- Mathematics curriculum-- ECD (3)

In order to identify gaps within outcomes/ study design a criteria was set of a minimum of 5 studies per intervention. The following gaps were identified:

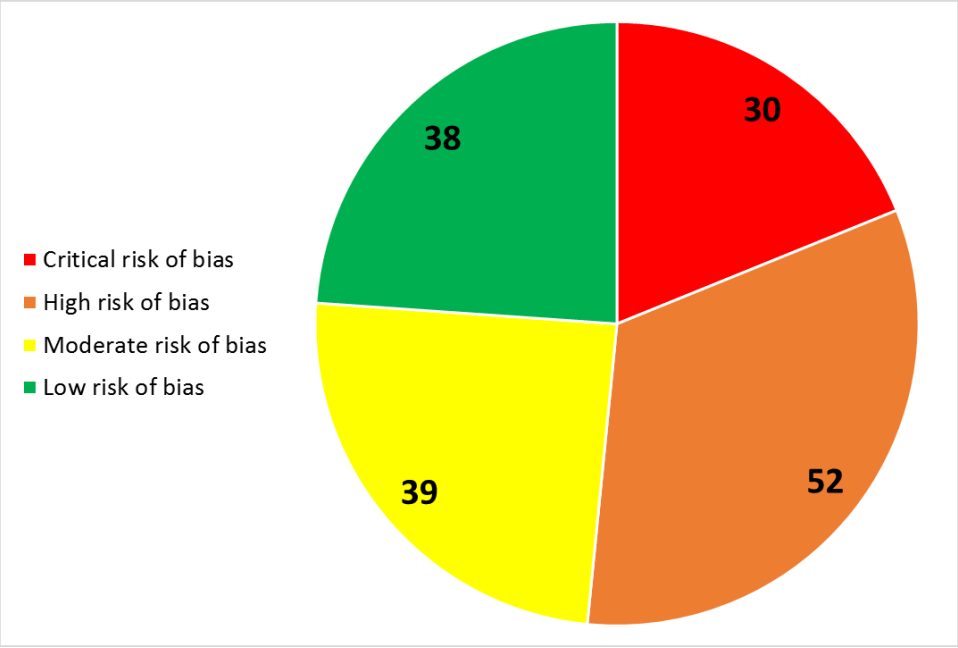
- Parent change in parental engagement / behaviour (n=1)
- Parent change in beliefs / attitudes (n=0)
- Parent: Other (n=0)

Considering the information above, it is clear that there are major gaps within the literature in terms both interventions and outcomes with respect to parent based studies.

Quality of the Evidence

Figure 14 below indicates the quality of the included evidence-base. We used a rigorous critical appraisal tool to assess the quality of each included study consistently (Appendix H). Overall, the quality of the included evidence was low. A total of 52% of the studies was assessed as either of critical risk of bias (n=30) or high risk of bias (n=52). This leaves studies rated of a high risk of bias comprising one third of the entire evidence-base alone. Only a quarter (24%) of studies was rated of a low risk of bias and can therefore be regarded as fully trustworthy. A similar proportion of studies (24%) was rated of moderate risk of bias indicating some concerns about the trustworthiness of the evidence. In sum, overall, this leaves the evidence-base of poor quality and only a small subset of studies can be recommended without reservation in assessing intervention effects and to inform policy. A similar picture emerges when only looking at the South Africa studies: 51% of all studies are of either critical (14%) or of high risk of bias (37%). This is in contrast to 49% of South African studies that are either of moderate (29%) or low (20%) risk of bias.

Figure 8 Quality of the included evidence-base

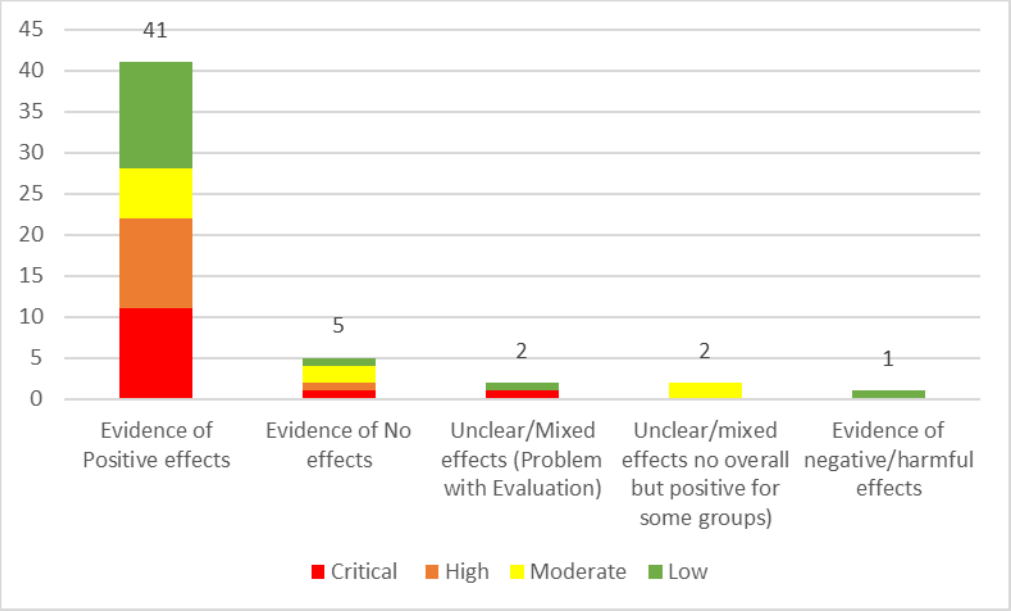


6. EFFECTIVENESS OF THE INTERVENTIONS

In terms of intervention effectiveness, a majority of the mapped studies reported evidence of positive effects (n=41), with five reporting evidence of no effects. Two studies indicated unclear /mixed effects based on a problem in the evaluation. Only one study revealed evidence of negative/harmful effects. Given the stark contrast between published positive and published negative effects, there is an indication of publication bias within the available evidence-base.

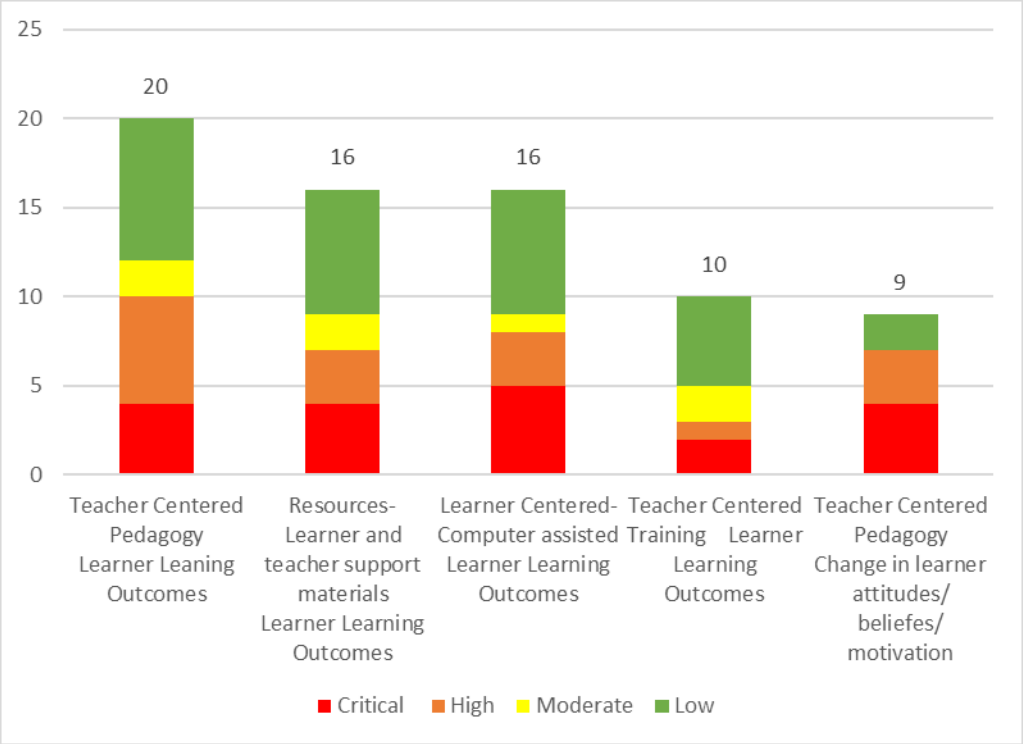
Of the 41 studies reporting on positive effects, 11 were assessed as critical risk of bias and high risk of bias respectively, 6 studies were assessed to have a moderate risk of bias and 13 a low risk of bias. Evidence of no effects reported in 5 studies had an overall proportionate distribution between critical, high, moderate and low risks of bias. Unclear studies/Mixed effects (Problem with Evaluation) and Unclear/mixed effects no overall but positive for some groups were found in 2 studies respectively. The former entailing one study critical risk of bias and one low risk of bias and the latter having both studies assessed as moderate risk pf bias. Evidence of negative/harmful effects was reported in just one study assessed to have a low risk of bias.

Figure 9 Intervention effects



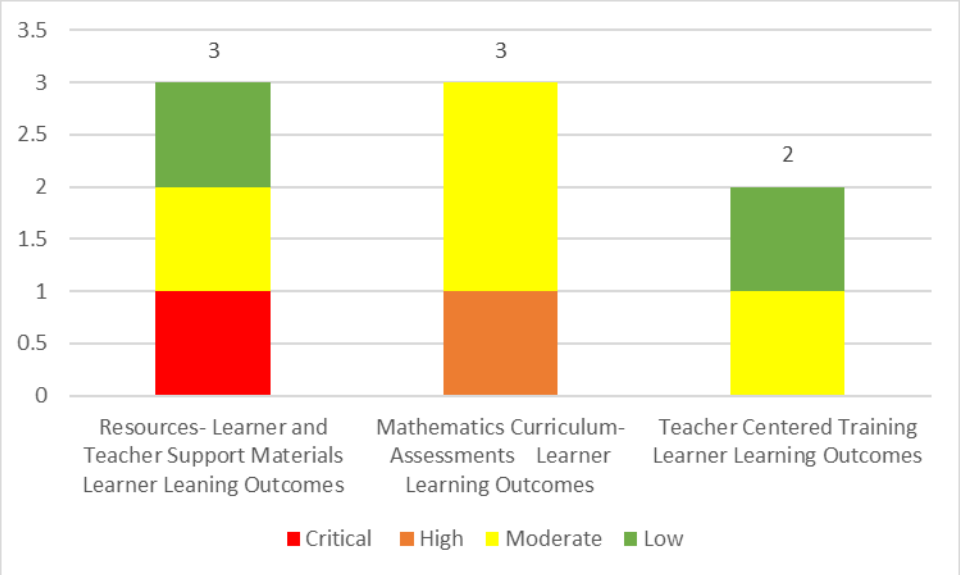
Breaking the evidence of positive effects down further, Figure 16 represents the top 5 configurations of interventions and outcomes that have reported positive effects in conjunction with the appraisals. The most prominent configuration leading to positive effects was Teacher-Centered: Pedagogy. Two configurations had 16 studies each with the lowest configuration having just 9 studies. With regard to the appraisals, the overall findings reflect the overall quality of the included evidence-base in Figure 16. Most of the configurations that have reported positive effects have more than half of the studies graded as either critical or high risk of bias.

Figure 10 Types of interventions with positive effects



Breaking the evidence of no effects down further Figure 17 represents the top 3 configurations of interventions and outcomes that have reported no effects together with appraisals for the studies. There were a few other studies that had only one configuration which are not represented in the figure due to space constraints. Overall, Evidence of no effects configurations were largely clustered between 1-3 studies with only three configurations representing more than a single study. Interestingly, critical risk of bias for studies reporting no effects contrast to the those of positive effects as well as the overall quality of the included evidence-base. Most of the studies configurations illustrate indicate that the majority of the studies are either moderate to low risk of bias.

Figure 11 Types of interventions with no effects



Results from the map of the remaining three categories are summarised below:

- Evidence of unclear/mixed effects with an evaluation problem had been reported in only two studies. Similarly, unclear/mixed effects no overall but positive for some groups had been reported in two studies only. Both, unclear/mixed effects with an evaluation problem and unclear/mixed effects (no overall but positive for some group) had studies with various intervention and outcome configurations.

- Evidence of negative/harmful effects was a rare finding in the literature being reported only in one study that had looked at various configurations of interventions and outcomes.

APPENDICES (ACE)

A. Description of the research process
B. Record of screened studies at title and abstract and at full-text
C. List of included studies
D. Folder of included studies and data extraction (https://drive.google.com/open?id=17Q2drBopjhW6l7QZ6fabqEk8J9MMlbwj)
E. Inclusion Criteria
F. Search Strategy and Results
G. Data Extraction Tool
H. Critical Appraisal Tool

Table 1: Full list of Methods

Quasi-experimental	35
Descriptive assessment	38
Qualitative case study	24
Mixed-methods assessment	15
RCT	15
Action research	12
Qualitative assessment	10
Ethnography	5
Mixed- Methods	2
Life histories	2
Natural experiment	1
Qualitative evaluation	1
Comparative case study	1
RDD	1

Table 2: Full list of Interventions

Pedagogy	67
Cognitive/Learner processes	44
Content knowledge	37
Learner and teacher support materials (LTSMs)	35
Accountability and motivation	22
Attitudes & beliefs	22
Computer-assisted learning	21
Language	20
Training	19
Curriculum	17
Information and Communication Technologies (ICTs)	16
Language	14
Assessments	14
Out-of-school offerings	9

Additional personnel	8
Oversight, supervision and monitoring	8
Instructional leadership	8
School readiness	7
Involvement	6
Context	4
Physical infrastructure	4
School functionality	4
Language	4
Health & nutrition	2
ECD	2
Language	1
Other	1