



Department of Performance Monitoring and Evaluation

Report on the Assessment of Government Evaluations

Evaluation Title:	Formative Evaluation of Textbooks and Workbooks
Evaluation Number:	421
Evaluation Completion Date:	28 February 2013
Period of Evaluation:	October 2011 - February 2013
Submitted:	21 August 2014 by Thandeka Mhlantla
Approved:	21 August 2014 by Mike Leslie

Evaluation Details

Evaluation Title:	Formative Evaluation of Textbooks and Workbooks
Evaluation Number:	421
Evaluation Completion Date:	28 February 2013
Created:	20 January 2014 by Mike Leslie
Submitted:	21 August 2014 by Thandeka Mhlantla
Approved:	21 August 2014 by Mike Leslie
Period of Evaluation:	October 2011 - February 2013
Known Cost:	R 4,119,707.00
Known Cost Type:	Audited
Initiated By:	The Department of Basic Education
Initiated By Internal:	Yes
Undertaken By:	The Australian Council for Educational Research (ACER)
Undertaken By Internal:	Yes

Assessors

Thandeka Mhlantla	thandeka@pdg.co.za
-------------------	--------------------

Assessment Documents

Document Name:	Document Type:	Added By:	Added On:
Consultative Summarised Workbook & Textbook Review Report.pdf	Any other relevant documentation pertaining to the evaluation process	Mike Leslie	20 January 2014
Consultative UNICEFDBEEvaluation of workbooks and textbooks200813.pdf	Evaluation report	Thandeka Mhlantla	26 February 2014
Summary Report on the Formative Evaluation of Workbooks and Textbooks South Africa.pdf	Any other relevant documentation pertaining to the evaluation process	Thandeka Mhlantla	26 February 2014
Summary assessment of the technical proposals.docx	Any other relevant documentation pertaining to the evaluation process	Thandeka Mhlantla	26 February 2014
ACER BAFO Tech Proposal.pdf	Any other relevant documentation pertaining to the evaluation process	Thandeka Mhlantla	26 February 2014
Critical Reader ToR_3 April 2012.doc	Terms of Reference (ToR) for the evaluation	Thandeka Mhlantla	26 February 2014
Final ReportExtension2013.pdf	A letter of engagement or contractual agreement with the service provider	Thandeka Mhlantla	26 February 2014
INCEPTION REPORT submitted 4.6.12.docx	An inception report	Thandeka Mhlantla	26 February 2014

Statement of Work & TOR.pdf	Terms of Reference (ToR) for the evaluation	Thandeka Mhlantla	26 February 2014
Supplement 6 - Contracts Review Committee (Final Draft).docx	Any other relevant documentation pertaining to the evaluation process	Thandeka Mhlantla	26 February 2014
Assessment Report at Moderation.pdf	Assessment Report at Moderation	Mike Leslie	27 February 2014

Quality Assessment Summary

The overall score of this evaluation has been rated at 3.58 out of 5 on the likert scale applied to assess the quality of government evaluations. This means the evaluation scored above average according to the standards mentioned on the quality assessments.

The highest score in the overarching considerations was in partnership approach which scored 3.94 out of 5. This is evident in that protocol measures of creating a steering committee to oversee different project milestones and to ensure the direction of the evaluation were well managed. Throughout the evaluation the relevant stakeholders made the process of interaction strong and easier to manage. One of the other contributors here could have been because of the good communication engagements that were occurring throughout the process of the evaluation between all the relevant stakeholders. Good feedback and interrogation of different milestones outcomes allowed for good direction and shape for the evaluation to take place.

The reporting phase of the evaluation scored 3.89 out of 5 which could be attributed to the good executive summary and the clearly reported structure of the evaluation. Through the collection of qualitative and quantitative data through methodologies such as surveys, focus groups, classroom observations, individual interviews, and case studies. These assisted with providing strong recommendations for each of the relevant stakeholders made for a good report to contribute to the policy context of proposals for the Action Plan 2014 for education.

The lowest score of the evaluation resulted from the capacity development overarching consideration. This was noted at 2.67 out of 5. Although it was planned to occur in the inception report there was little evidence that was substantiated as to whether this actually took place during the evaluation process.

In terms of the follow up, use and learning, it is clear that this is an important evaluation for the DBE and also for ACER. It not only focused on one of the most pressing issues that contributes to the improvement of the quality of education in the country, it also allowed for recommendations to be easily adaptable by the DBE.

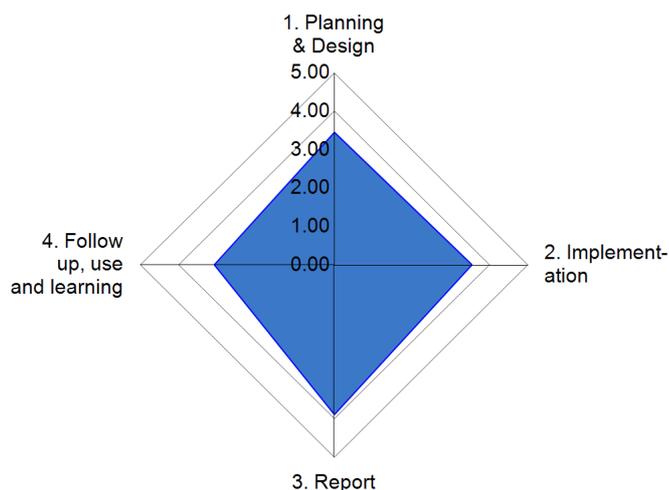
The evaluation has already become a proud document of the DBE which is seen as a good reflection on the work they are trying to achieve in basic education in South Africa. It is available on the DBE website and has been widely engaged with by the department.

Quality Assessment Scores

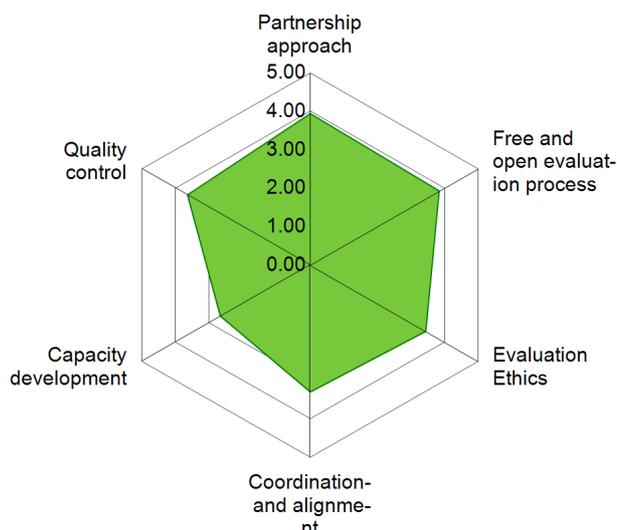
Phase of Evaluation	Score
1. Planning & Design	3.46
2. Implementation	3.55
3. Report	3.89
4. Follow-up, use and learning	3.08
Total	3.58

Overarching Consideration	Score
Partnership approach	3.94
Free and open evaluation process	3.85
Evaluation Ethics	3.44
Coordination and alignment	3.30
Capacity development	2.67
Quality control	3.64
Total	3.58

Scores: Phases of Evaluation



Scores: Overarching Considerations



Phase of Evaluation	Area of Evaluation	Score
1. Planning & Design	1.1. Quality of the TOR	3.58
1. Planning & Design	1.2. Adequacy of resourcing	3.36
1. Planning & Design	1.3. Alignment to policy context and background literature	2.00
1. Planning & Design	1.4. Appropriateness of the evaluation design and methodology	3.50
1. Planning & Design	1.5. Project management (Planning phase)	5.00
2. Implementation	2.1. Evaluation ethics and independence	3.64
2. Implementation	2.2. Participation and M&E skills development	3.50
2. Implementation	2.3. Methodological integrity	3.60
2. Implementation	2.4. Project management (Implementation phase)	2.00
3. Report	3.1. Completeness of reporting structure	3.83
3. Report	3.2. Accessibility of content	4.00
3. Report	3.3. Robustness of findings	3.77
3. Report	3.4. Strength of conclusions	4.29
3. Report	3.5. Suitability of recommendations	4.23
3. Report	3.6. Consideration of reporting risks and ethical implications	3.31
3. Report	3.7. Project management (Reporting phase)	4.00
4. Follow-up, use and learning	4.1. Resource utilisation	2.60
4. Follow-up, use and learning	4.2. Evaluation use	3.21
Total	Total	3.58

1. Planning & Design

1.1. Quality of the TOR

Standard: 1.1.1. The evaluation was guided by a well-structured and complete TOR or a well-structured and complete internal evaluation proposal

Comment and Analysis: The TOR was well-structured. It was concise and provided the rationale for looking at textbooks and workbooks. It gave a good understanding of the background of the work and the need for the evaluation to be done. The breakdown of the contents that were highlighted included looking at the purpose and objectives of the evaluation which address efficiency (quality) and effectiveness (utilisation). It clearly states the evaluation questions that need to be considered also taking into account the different levels of the results chain. Methodology is mentioned but not in great detail. Reporting and deliverables are stipulated and expectations that the DBE had about the chosen service provider. It was clearly outlined within the TOR that the evaluation report needed to include the following • Executive summary • Background • Evaluation approach and methodology • Findings (by evaluation criteria and questions) • Conclusion • Recommendations • Reference • Annexure. The primary users of the evaluation are Department of Basic Education and counterparts at provincial levels. The ToR indicated that the consulting firm/institution would have to put together a gender-balanced multidisciplinary team which has demographic sensitivity that is suitable to address the with different categories of textbooks/workbooks and evaluation methods.

Rating: 4

Standard: 1.1.2. The purpose of the evaluation stated in the TOR (or an internal evaluation proposal) was clear and explicit

Comment and Analysis: The purpose of the evaluation was clearly stated in the TOR. The evaluation seeks to address two globally accepted criteria: the efficiency (quality) and effectiveness (utilisation) of textbooks and workbooks used in schools. The TOR (2010: 3) states, "the purpose of the formative evaluation is to provide feedback on performance of language and Mathematics text books to guide adjustments to future editions. The primary users of the evaluation are Department of Basic Education and counterparts at provincial levels. The evaluation will assess the management, coordination, support and resources in relation to curriculum objectives, outcomes and coverage as described above. An explicit outcome of the evaluation will be a tool which may be used in future to do the same exercise, in addition to an evaluation report".

Rating: 5

Standard: 1.1.3. The evaluation questions in the TOR (or an internal evaluation proposal) were clearly stated and appropriate for addressing the evaluation purpose

Comment and Analysis: The TOR did not explicitly map out questions however, the purpose of the evaluation shows that there is a need to include all types of schools in the evaluation and to also ensure the best outcomes are achieved. Since this particular standard addresses the TOR highlighting the questions one will understand that this would be a shortcoming towards the document and therefore a low score is given for this standard.

However, within consultative process, three main questions were listed in the Inception Report that ACER delivered which were identified by UNICEF/DBE,:

1. How are the textbooks and workbooks being utilised?
2. What is happening in schools where the workbooks and textbooks are not being utilised?
3. Do the workbooks and textbooks have the characteristics of quality text/workbooks

Rating: 2

Standard: 1.1.4. The approach and type of evaluation was suited to the purpose and scope of the evaluation TOR (or an internal evaluation proposal)

Comment and Analysis: The evaluation is conceptualised as a formative assessment of an on-going intervention which has been classified in terms of the implementation evaluation typology. A Realist Evaluation using participatory action research was highlighted in the inception report and was suited to meet the needs of the requirements of the TOR in order to get the best possible outcomes out of the main issues needing to be discussed and found. This type of evaluation is crucial to bringing together various sources of information to best find outcomes that answer questions around the utilisation of workbooks and their quality. The evaluation will assess the management, coordination, support and resources in relation to curriculum objectives, outcomes and coverage.

Rating: 4

Standard: 1.1.5. The TOR (or an internal evaluation proposal) identified the intended users of the evaluation and their information needs

Comment and Analysis: The DBE and UNICEF were identified as users of the evaluation but there is no explicit indication as to what their information needs are in relation to the project. Passing reference is made to making recommendations in line with Action Plan 2014 but the intended users and their specific information needs could have been better expressed in the ToR.

Rating: 2

Standard: 1.1.6. Key stakeholders were involved in the scoping of the TOR and choosing the purpose of the evaluation

Comment and Analysis: Key stakeholders within the DBE were involved in the scoping of the TOR and choosing the purpose of the evaluation. Ms Carol Deliwe at the DBE was seen to be an integral part of the DBE for ensuring that the consultative process between the client and service delivery was beneficial too both parties.

Rating: 4

1.2. Adequacy of resourcing

Standard:	1.2.1. The evaluation was adequately resourced in terms of time allocated
Comment and Analysis:	The evaluation was envisioned to be completed within a year of the inception report although delayed procurement processes meant that initially a shorter time frame of 6 months was requested, later adjusted to a year. The original provision in the TOR was on balance a fair allocation, although this was challenged by delays in beginning the evaluation thereby providing insufficient time allocations at the start of implementation, as discussed later.
Rating:	2
Standard:	1.2.2. The evaluation was adequately resourced in terms of original budget
Comment and Analysis:	The initial evaluation was budgeted for R3 762 759.00 which was a more than reasonable original budget for the evaluation over the period of a year. Later, a request was made for an additional desk-based review at a cost of R356 948.00. Both of these budgets appear generous for the type and scope of the work requested.
Rating:	4
Standard:	1.2.3. The evaluation was adequately resourced in terms of staffing and skills sets
Comment and Analysis:	ACER provided experts in education from within their organisation at different levels to assist with conducting the evaluation. All varied in skills set. The team also encompassed key national researchers from UNISA to assist ACER key staff members who had educational experience but not necessarily contextual knowledge of the basic education system in South Africa. By involving national experts in education and research assistants at post graduate level from UNISA, this created a more sound and credible unit of technical representation.
Rating:	4
Standard:	1.2.4. Where appropriate, the evaluation planned to incorporate an element of capacity building of partners/staff responsible for the evaluand
Comment and Analysis:	The evaluation planned to incorporate an element of capacity building for research assistants doing their post-graduate studies in universities in South Africa. There is mention about capacity building for researchers that was provided by ACER before data collection began but this was not explicitly targeted at capacity building to be passed on to DBE staff members.
Rating:	3

1.3. Alignment to policy context and background literature

Standard: 1.3.1. There was evidence that a review of the relevant policy and programme environments had been conducted and used in planning the research

Comment and Analysis: There is passing reference to the policy environment in the ToR. The inception report did however mention the need for the evaluation to be contextualised in light of departmental policies in 2011 and 2012. In short, there was not clear evidence that relevant policy and programme environments were reviewed at the planning stage.

Rating: 2

Standard: 1.3.2. There was evidence of a review of appropriate literature having been conducted and used in planning the research

Comment and Analysis: There is limited evidence that a literature review informed the ToR but the inception report does frame the research in terms of a brief review of literature relevant to development interventions. There is not evidence of a review of literature relevant to the educational sector or textbooks and workbooks at the planning stage.

Rating: 2

1.4. Appropriateness of the evaluation design and methodology

Standard: 1.4.1. There was explicit reference to the intervention logic or the theory of change of the evaluand in the planning of the evaluation

Comment and Analysis: From inception reporting stages the logic of the intervention was acknowledged and this shaped the methodology of the evaluation. The intervention logic could have been better defined and provided a breakdown of the different stakeholders' involvement required in order to create change and impact.

Rating: 3

Standard: 1.4.2. Key stakeholders were consulted on the design and methodology of the evaluation

Comment and Analysis: Members of the DBE were consulted with ACER consultants during the design stages of the evaluation. Within the designing and finalisation of the inception phase there was good consultation processes that facilitated contribution from the client (DBE) in order to enable buy-in from stakeholders. Key stakeholders were consulted within this process through a steering committee forum which involved a mix of DBE people.

Rating: 4

Standard: 1.4.3. The planned methodology was appropriate to the questions being asked

Comment and Analysis: The TOR highlighted its specifications on the methodology which stipulated that "the sample must include poorly performing schools, as well as those with multigrade teaching, and foundation phase grades, with some specialist schools included (Dinaledi). The sample will reflect the variety of contexts that characterise the South African schooling system, including the eleven languages of learning and teaching in schools as well as the eleven language subjects". This was the only form of methodology mentioned in the ToR. ACER highlighted that a mixed method approach would be the best to adequately answer questions posed by the evaluation and the following methods to be used for the evaluative tools:

1. Questionnaire for students
2. Questionnaire for teachers
3. Questionnaire for school managers
4. Interview structures and protocols for PED representatives
5. Interview structure and protocols for parents (focus groups)
6. Classroom observation rubric (pedagogical practice)
7. Classroom observation rubric (learner behavior/performance)
8. Rubric for assessing book quality in the desktop review

They highlighted that these would best represent the purpose of the evaluation, including answering the questions, and meet the needs of the DBE.

There are three categories of research method: an overview of sample schools across the country based on questionnaires and interviews; three school case studies; and the use of a rubric, a checklist, to evaluate the quality of the Workbooks.

Rating: 4

Standard: 1.4.4. Sampling was appropriate and adequate given focus and purpose of evaluation

Comment and Analysis: Out of the 48 000 schools in SA, 231 schools ended up constituting the sample size for the analysis. Of those, it was decided that 11 schools would be the case studies that included provincial office representative interviews, parent focus groups, and classroom observations (of teachers and students). The focus of the evaluation demanded rich content delivery from questionnaires, interviews, observations and desktop analysis and so a larger sample could have been richer and added more depth to the study, but what was planned as a good sample. The sampling was statistically representative for what the evaluation sought to achieve and gave rich representation for the data collection.

Rating: 4

Standard: 1.4.5. There was a planned process for using the findings of the evaluation

Comment and Analysis: Little detail is provided as to how the planned findings of the evaluation will be infused to inform the department's activities. Although one could argue that for the mere fact that the department is commissioning this evaluation, they necessarily need to use the findings, there is no information mentioning how the use of the findings will either contribute to the delivery or use of these books in the future.

Rating: 2

1.5. Project management (Planning phase)

Standard:	1.5.1. The inception phase was used to develop a common agreement on how the evaluation would be implemented
Comment and Analysis:	ACER made a variety of changes to adapt to the needs and requirements of the DBE and other key stakeholders. There was good consultative processes that allowed for the scope to not only reflect on the concerns and needs of the client, but also key stakeholders. During the deliberation phase of the inception report ACER and the DBE had consecutive meetings to ensure awareness and satisfaction of the intended product.
Rating:	5

2. Implementation

2.1. Evaluation ethics and independence

Standard: 2.1.1. Where data was gathered in contexts where ethical sensitivity is high, appropriate clearance was achieved through an ethics review board; e.g. in evaluation involving minors, institutions where access usually requires ethical or bureaucratic clearance, and situations where assurances of confidentiality was offered to participants

Comment and Analysis: Because this was an evaluation on the delivery of school materials, no parental consent needed to be given, although many schools did notify parents nonetheless. The teachers union was notified in time about the evaluation which also assisted in ensuring that there would be good engagement in classrooms and teachers were aware of the mandate of the evaluation. An accompanying letter from the DG went out with all evaluators to schools and schools were contacted in advance about the evaluation and observation processes. Where necessary the DBE supported ACER with the correct letters to enable them access the relevant sampled schools and provided permission in line with departmental protocols.

Rating: 4

Standard: 2.1.2. Where external, the evaluation team was able to work freely without significant interference

Comment and Analysis: There was no evidence of interference/intimidation from the DBE or UNICEF in terms of how they saw the evaluation take on a direction. The methodology was free conducted by ACER to then inform the DBE on its recommendations.

Rating: 4

Standard: 2.1.3. The evaluation team was impartial and there was no evidence of conflict of interest

Comment and Analysis: There were some initial issues that were experienced in the inception phase of the project. There was change in the scope of the evaluation which resulted in there being a conflict with the main project manager from ACER. However, this was quickly resolved with a new programme manager stepping in. This resulted in the approval of the inception report with ACER and the DBE within a two week turnaround.

Rating: 3

2.2. Participation and M&E skills development

Standard: 2.2.1. Key stakeholders were consulted through a formalised mechanism or institutional arrangement

Comment and Analysis: The steering committee was formed as a formal mechanism to allow for stakeholder input and engagement to occur during the evaluation process. Using this committee also allowed for buy-in and also more assistance in ensuring that the process for the evaluation could go smoothly. This process also allowed DBE to involve some of their staff members from different units to engage in the evaluation.

Rating: 4

Standard:	2.2.2. Where appropriate, an element of capacity building of partners responsible for the evaluand was incorporated into the evaluation process
Comment and Analysis:	By allowing for a stakeholder engagement to occur, the evaluation allowed for an interactive learning process to be achieved. A mechanism to test the evaluative tools and to better understand the data, the DBE also used some of its internal statisticians and required them to run the data again as a form of verification but also to learn how to use different models designed externally that can fit in their educational programmes. This also largely drew interest on ensuring the credibility of the data and information that the evaluation provided.
Rating:	3
Standard:	2.2.3. Where appropriate, the evaluation team incorporated an element of skills development amongst the evaluators (e.g. students, interns, recent graduates, etc)
Comment and Analysis:	ACER has a large pool of employees, almost 300 researchers internally. The evaluation used a large amount of people from the policy analysis and programme design division. These people were either specialists in the field of education or specialists in evaluation. Researchers who were from South Africa were selected from UNISA and used on the ground to conduct some of the interviews/observations. It is stated in the Inception Report that "prior to the commencement of data collection in September, ACER staff and KNRs will train junior local researchers in application of the developed evaluation rubrics, interviewing skills and classroom observation techniques (2012: 33). It is "envisaged that this process will provide significant capacity building opportunities for the postgraduate students involved, but equally will allow research to be appropriately informed by the local context. One training will be held in either Pretoria or Johannesburg depending on which location is easiest for all participants. Additional trainings may also be held in other cities as necessary" (ACER, 2012:33). It was not clear to what extent this was fully implemented, but there was evidence that the evaluation team did grow some internal evaluation capacities.
Rating:	4
Standard:	2.2.4. Peer review of the agreed evaluation design and methodology occurred prior to undertaking data collection
Comment and Analysis:	No explicit peer review was mentioned to have taken place prior to undertaking the evaluation although a broader Advisory Group reviewed all submissions and was understood to fulfill this function. ACER welcomed engagement and guidance from the Advisory Group throughout the life cycle of the evaluation.
Rating:	3

2.3. Methodological integrity

Standard: 2.3.1. The methods employed in the process of the evaluation were consistent with those planned

Comment and Analysis: The study envisioned to do a case study of 11 schools, which was the agreed size in the inception phase, but this needed to be revised to 3. This was not due to ACER or DBE fault but largely due to issues of communication and contact of relevant people to assist with allowing for evaluations to occur. This was the only technical glitch that changed the methods of evaluation in the study. One could argue that by the use of only 3 case studies, the representivity was skewed and not fully representative of the desired outcomes as planned. Apart from minor changes and revisions from the inception report, the methodology was welcomed by the DBE.

Rating: 3

Standard: 2.3.2. A pilot of data collection instrumentation occurred prior to undertaking data collection

Comment and Analysis: Evidence could not be found that a pilot was done on the data collection instruments. No evidence was given in the report that suggests this was done or not. However, the questionnaires were designed by ACER and consideration was taken by the committee on the relevance of the questions and whether they served the purpose of the evaluation.

Rating: 2

Standard: 2.3.3. Data collection was not compromised by fieldwork-level problems or unplanned diversions from original intentions

Comment and Analysis: The main issues that could have been seen as compromising the data collection were issues that surrounded choosing a sampling population. Conversations were held between the DBE and ACER around choosing the correct sample. No fieldwork was compromised, but due to high level technical issues the reduction of the case study schools were reduced. As a result, this may have limited investigations into how schooling is experienced in South Africa in order to identify factors that have an impact on the efficiency and effectiveness of textbooks/workbooks.

Rating: 3

Standard: 2.3.4. Forms of data gathering were appropriate given the scope of evaluation

Comment and Analysis: Data were collected through: structured observations of classrooms; questionnaires for Heads of Department, students, teachers, school governing body representatives; semi-structured interviews were conducted with school leaders; semi-structured focus groups were conducted with students, parents and teachers. Furthermore, a quality rubric for assessment of the textbooks was developed for the purpose of the review. In total, these represented more than appropriate forms of data gathering given the scope of the evaluation.

Rating: 4

Standard:	2.3.5. The data analysis approach and methods were appropriate and sufficient given the purposes of the evaluation
Comment and Analysis:	The data analysis approach involved reviewing the secondary data and empirical findings of the study to assess the utilisation of workbooks in schools. Various correctional studies were done to do cluster analysis that made understanding the nature of the study more rich and in-depth. Triangulation of data was done through qualitative and quantitative data collected during interviews, focus groups, classroom observations and questionnaires.
Rating:	4
Standard:	2.3.6. Key stakeholders were significantly engaged as part of the methodology
Comment and Analysis:	Key stakeholders in education were involved as part of the methodology. Parents, teachers, and school governing boards were all included as a source of data. The 50 sampled schools were representative of low-resource schools, as this was the intended audience for the workbooks that had been developed. Case study schools were drawn from the sample of survey schools and were chosen in consultation with DBE. Surveys were prepared for four teachers in each of the sampled schools. Two teachers were required to complete the survey from the mathematics perspective and two from the language perspective.
Rating:	4
Standard:	2.3.7. The methodology included engaging beneficiaries appropriately as a key source of data and information
Comment and Analysis:	Learners as the ultimate beneficiaries of the intervention were involved as a key source of data through the structured observation of classrooms, semi-structured interviews, focus groups as well as questionnaires related to the utilisation of workbooks. The different types of methods employed provided a comprehensive triangulation of data collection methods for intervention beneficiaries and exemplifies best practice.
Rating:	5

2.4. Project management (Implementation phase)

Standard:	2.4.1. The evaluation was conducted without significant shifts to scheduled project milestones and timeframes
Comment and Analysis:	The proposed milestones of the evaluation were adhered to by ACER by ACER insofar as possible. This was due to changes in the scope of the evaluation and delays in appointment. The initial evaluation was conducted between the submission of the inception report in June 2012 and the finalisation of the evaluation report in February 2013. Although requested to be completed within 6 months it was then extended to include an additional request from the DBE to add further desktop review to acquire a deeper understanding in the programme.
Rating:	2

3. Report

3.1. Completeness of reporting structure

Standard: 3.1.1. Executive summary captures key components of the report appropriately

Comment and Analysis: The executive summary clearly articulates the background/motivation for the evaluation and the process of the methodology. It does a good job of explicitly highlighting the key areas the evaluation addressed and goes mentions the key findings of each of the evaluation questions that support the purpose of the evaluation. It also goes into summarising some of the recommendation that came out of the evaluation.

Rating: 4

Standard: 3.1.2. The context of the development intervention is explicit and presented as relevant to the evaluation

Comment and Analysis: The development of the intervention is mentioned throughout the evaluation. Within the section of the literature review, one is able to identify a link towards the consideration and concern around the challenge within the South African education system with regards to the quality of effective textbooks and workbooks. The initiative is specific to the evaluation needs to interrogate and address the standard of these books and ways in which to develop improvement to these books and their effects on learners and teachers.

Rating: 4

Standard: 3.1.3. There is a clear rationale for the evaluation questions

Comment and Analysis: The rationale for the evaluation question steps from the key data processes that the evaluation addressed in order to meet with the concerns DBE expressed about the quality and effectiveness of teaching and learning; in particular, to drive improvement in the Basic Education sector. There is a clear use of the questions to fully investigate whether the intervention will succeed and how it contributes to the use of the books.

Rating: 4

Standard: 3.1.4. The scope or focus of the evaluation is apparent in the report

Comment and Analysis: The scope of the evaluation is clearly articulated by stating that "the study assess the quality of utilisation of workbooks and the quality of textbooks in South African school" (ACER, 2013: 24). It highlights how teachers and pupils utilise these books, but explicitly states that it does not plan to understand what impact these books have on South African teachers. This is proven throughout the reports data gather methodologies and data analysis all talk to the research questions which are formulated around the scope of the evaluation.

Rating: 4

Standard:	3.1.5. A detailed methodology is outlined in the relevant section of the report to the point that a reader can understand the data collection, analysis and interpretation approaches used
Comment and Analysis:	The document clearly outlines the direction the methodology took. It was broken into three segments which looked at (a) a large scale study drawing on a representative sample of schools, (b) three school case studies and (c) a desk review of workbooks and textbooks. Within each of the components is was clearly outlined what each of the segments of the methodology would entail and how this would be done. It was also brought forward to the steering committee to understand and either agree or disagree on the methodology that ACER presented. ACER was able to determine all the relevant instruments needed to conduct the evaluation to provide quality monitoring and good data quality.
Rating:	3
Standard:	3.1.6. Acknowledgement of limitations of all aspects of the methodology and findings are clearly and succinctly articulated
Comment and Analysis:	The evaluation highlighted that there could be limitations of the method (self-reporting on questionnaires, messages lost in translation, small qualitative component, causality cannot be established). There were some limitations that arose that could not be controlled and as a result were mentioned by interviewees.
Rating:	4
Standard:	3.1.7. Key findings are presented in a clear way; they are made distinct from uncertain or speculative findings; and unused data is not presented in the body of the report
Comment and Analysis:	Findings were clearly articulated and was included in the report. The findings were also divided into the evaluation questions for the executive summary. The report provided substantial findings that were articulated by graphs and narrative explanation of the relevant data collection tools. No statements were given to state whether any speculative findings were gathered or if there was uncertain findings. All the information gathered pertained to the use and need of the evaluation.
Rating:	4
Standard:	3.1.8. Conclusions and recommendations are clear and succinctly articulated
Comment and Analysis:	Evidence provided nine recommendations that could assist the DBE with the ability to contribute towards the improvement of both the effectiveness and utilisation of workbooks and textbooks in South Africa and provide feedback to guide adjustments to future editions. Each recommendation was clearly combed out with logical steps to be taken and followed. The conclusion provided insightful to note that there is a need to recognise the role teachers play in steering basic education and the use of these textbooks and workbooks.
Rating:	4

3.2. Accessibility of content

Standard: 3.2.1. The final evaluation report is user-friendly, written in accessible language and its content follows a clear logic

Comment and Analysis: The final evaluation is available on the DBE website and is a proud piece of writing for ACER too. It has been written in English and is accessible on the DBE website for all those who wish to access it. It has also been summarized into an easier read for the general public, with graphical representation and a full colour copy that makes for a more fun read. The only issue would be to know whether parents, pupils and teachers who were interviewed managed to access these findings after the evaluation either on the website or ACER/DBE distributed the results to them. This would be a good way to understand how the process of information would have translated to the direct users of the product.

Rating: 4

Standard: 3.2.2. Quality of writing and presentation is adequate for publication including: adequate layout and consistent formatting; complete sentences and no widespread grammatical or typographical errors; consistency of style and writing conventions (e.g. tense, perspective (first person, third person); levels of formality; references complete and consistent with cited references in reference list and vice versa; etc.

Comment and Analysis: The quality of the writing is good and there was a good pool of resources that were analysed. The layout and language is appropriate and does a good job at interrogating all the findings. The graphs and statistical representation are all in context to the written explanation and provide for good analysis within the report. No errors of any sort were noted in the report.

Rating: 4

Standard: 3.2.3. Appropriate conventions are used in presentation of data (e.g. use of appropriate statistical language; reporting of p-values where appropriate; not reporting statistically insignificant findings as significant; clarifying disaggregation categories in constructing percentages; not using quantitative language in reporting qualitative data, etc.)

Comment and Analysis: The statistical language was used was sound and beneficial to help with the explanation and interpretation of the quantitative aspects of the evaluation. The evaluation engaged more with the use of the particular product (work books and text books with pupils/teachers and parents), certain technicalities needed to be encompassed and reflective to ensure reporting served to contribute to answering the evaluation questions. It is a reader-friendly document that transitions well from quantitative to qualitative analysis and description that brings in good comparison.

Rating: 4

Standard: 3.2.4. The use of figures and tables is such that it supports communication and comprehension of results; and data reported in figures and tables are readily discernible and useful to a reader familiar with data presentation conventions

Comment and Analysis: The consistency of data (figures and tables) in the report were well articulated and explored throughout the report. The use of technical tools/vehicles of analysis were also used by the DBE for sampling the usefulness of all the information provided in the full text report.

Rating: 4

3.3. Robustness of findings

Standard:	3.3.1. Data analysis appears to have been well executed
Comment and Analysis:	Questionnaires were summarised in the form of frequency graphs and histograms, which assisted in better explaining the status of what was happening. ACER also used path analysis (not very different from correlational analysis), which is based on some relational hypotheses about antecedents (about what might lead to what) will be carried out. Analysis of work/test book quality will be in the form of a comparative study. The use here would be to compare all the books with one another this would show how effective each book each. The sampled schools for case studies look at deepening the evaluation through depth in data collection and examination of each of the schools. The report also stated that "For the meta-analysis we will draw on data from sources such as the 2007 SACMEQ study, TIMSS and PIRLS to assess the relationship between text book availability and utilisation against learner performance" (ACER, 2012: 22).
Rating:	4
Standard:	3.3.2. Findings are supported by available evidence
Comment and Analysis:	All findings either had hard copy questionnaires which were analysed and assessed as evidentiary support. Some registers were collected to ensure that there was confirmation that interviews occurred with the relevant researchers. All evidence was handed over to the DBE. Some of the quotes given in the interview processes were included in the final report and informed some of the conclusions. The TOR also stipulated that part of the mandatory deliverables all "full transcripts of all in-depth interviews and focus group discussions in an electronic format, including a list of places visited and people met" (TOR: 2011: 7).
Rating:	5
Standard:	3.3.3. The evidence gathered is sufficiently and appropriately analysed to support the argument
Comment and Analysis:	Evidence gathered has not been interrogated by other writers or professionals in the education field who could bring in some critique to further solidify the arguments and evidence made by ACER. However, this does not take away from its arguments made which are supported well by documentation and evidence from most of the data collection tools that were used. One could argue that more case studies could have been used to strengthen and tighten the argument as there is room for more input.
Rating:	4
Standard:	3.3.4. There is appropriate recognition of the possibility of alternative interpretations
Comment and Analysis:	One could bring in counter-arguments about the findings of the evaluation and its interpretation of the findings. There was not much recognition that alternative interpretations could likely occur.
Rating:	2

Standard: 3.3.5. The report appears free of significant methodological and analytic flaws

Comment and Analysis: The correct tools and statistical inferences were used to conduct the study. The graphical representations assisted with looking at percentiles, boxes and whiskers with bar graphs that described and explained the data analysis adequately. The data could have better described some focus areas more soundly.

Rating: 4

Standard: 3.3.6. Relevant limitations of the evaluation are noted

Comment and Analysis: Limitations are noted in that the "study does not seek to understand how the workbooks have impacted on the pedagogy of teacher in South Africa" (ACER, 2014: 24). It notes and substantiates that the number of sampled teachers was adequate to fulfil the evaluation. However, "the number of teachers who were in schools that were reported to not have workbooks and who report not using workbooks was 45. This means that there were very few schools available to the study to investigate what is happening in those schools without workbooks" (ACER, 2012: 24). The limitations could also further be extended

Rating: 3

3.4. Strength of conclusions

Standard: 3.4.1. Conclusions are derived from evidence

Comment and Analysis: Conclusions from the evidence were incorporated in each of the chapters. This also helped to ensure that information moved seamlessly from one end of the report to the next, to always "highlight a message" or make a conclusion about what was found. A summary of the various conclusions was also noted in the last chapter, as a result, recommendation were drawn from it.

Rating: 4

Standard: 3.4.2. Conclusions take into account relevant empirical and/or analytic work from related research studies and evaluations

Comment and Analysis: At each stage of the chapters, the evidence looks at collation of empirical evidence, that interacts with the literature review. This also helps to ensure the credibility and salience of the evaluation to allow for engagement and interrogation with high level literature interaction. The literature review assisted with the measuring the SA quality of work books and text books to those in the international space to conclude that they were of good quality standards. The information translated with providing sound conclusions that could help sum-up the findings and necessary inferences the study sought to identify.

Rating: 4

Standard: 3.4.3. Conclusions address the original evaluation purpose and questions

Comment and Analysis: The main purpose that was mentioned through three questions which were (A) How are the textbooks and workbooks being utilised? (b) What is happening in schools where the workbooks and textbooks are not being utilised? (c) Do the workbooks and textbooks have the characteristics of quality text/workbooks, (ACER, 2013). In looking at the conclusion is clear that "Nearly all schools are using the workbooks. Workbooks are being utilised in approximately 80 per cent of schools by at least some teachers. The biggest barrier to workbook use was reported as non-delivery of workbooks, suggesting that where the workbooks are available, they are quickly taken up by schools" (ACER, 2013: 141).

Rating: 5

Standard: 3.4.4. Conclusions are drawn with explicit reference to the intervention logic or theory of change

Comment and Analysis: Conclusions go back to understanding how the quality of the works books and their contribution aligns itself to the values of the Constitution as the driver for the workbooks intervention project has also been expressly tasked with reinforcing these fundamental values. The intervention logic drew issues around the fit and purpose of the evaluation of these books, which is explicitly mentioned in the conclusions as each of these work books and text books provide for specific learning objectives.

Rating: 4

3.5. Suitability of recommendations

Standard: 3.5.1. Recommendations are made in consultation with appropriate sectoral partners or experts

Comment and Analysis: On the basis of the evidence that surfaced out of the research, recommendations were passed onto the steering committee for the evaluation. through the consultation process of these officials, recommendations were discussed and also further refined to ensure that they focused on shaping worthwhile outcomes that could interact with the purpose of the evaluation.

Rating: 4

Standard: 3.5.2. Recommendations are shaped following discussions with relevant government officials and other relevant stakeholders

Comment and Analysis: The DBE was fully aware and made conscious of the recommendations before they took form on the final evaluations. A thorough consultative process was initiated and was maintained throughout the evaluation which made summing up of the evaluation easier as stakeholders were involved right from the inception phase. The ability to have relevant officials within the process also allowed for ownership of the evaluation to be taken into the DBE's hands as they needed to implement and roll-out the changes. Without proper engagement, the evaluation could go to waste. Having a champion like Ms Carol Nuga Deliwe at the DBE provided a point of entry that facilitated a good partnership approach.

Rating: 5

Standard: 3.5.3. Recommendations are relevant to the current policy context

Comment and Analysis: There is clear relevance of the recommendation to the policy context on both the improvement of quality and utilisation of workbooks and textbooks. Largely addressing the need to speak to providing schools with good teachers that can get better training on these books to get buy-in from them to effectively execute the use of these books. Not only does the report mention that there "needs to be a pilot towards a professional training program that might incorporate and compare teacher perceptions of those completing training and those who had not in order to compare pedagogic differences, but it also seeks to address the issue around alignment of workbooks and textbooks that the education space should provide" (ACER: 2013: 20).

The recommendations also speak to revision of some of the educational policies in place to better adapt to the context for the need of the workbooks for pupils in the relevant grades. Recommendations speak to some of the pressing and headed areas that need to be addressed in the basic education space for these books in SA.

Rating: 4

Standard: 3.5.4. Recommendations are targetted at a specific audience sufficiently - are specific, feasible, affordable and acceptable

Comment and Analysis: Each of the recommendations broadly touch base in a relevant area that contributes to the improving the use of the workbooks, not only with pupils but with teachers too. The report gives 9 recommendations to DBE to better employ the use of the books. Firstly, they touch on creating and targeting a professional training pilot programme for teachers. Developing a system for teachers, pupils and parents to understand why there is a need for using these books. DBE also needs to employ mechanisms that better enable understanding how CAPS and these books can be better aligned and used to inform one another. This highlights that all stakeholders involved in this process are aware of their specific tasks, and that the DBE truly has an opportunity to improve the current state for the use of these books. The recommendations made to each audience are specific and seem to be doable if all resources are fully fledged.

Rating: 4

3.6. Consideration of reporting risks and ethical implications

Standard: 3.6.1. Peer review of the draft evaluation report occurred prior to finalisation of the evaluation report

Comment and Analysis: Initial findings based on the results of the data analysis were discussed with the DBE and UNICEF as soon as conclusions were drawn. ACER drafted a draft report for stakeholders including recommendations. It was not mentioned that there was a peer review done on the draft evaluation report. However, ACER established a group of national experts who were either consulted individually or as a group, face-to-face, by e-mail or teleconference, to advise the ACER Project Team on key issues throughout the life of the project. Before the finalization of ACER provided stakeholders with an eight-page illustrated summary document in electronic form, cleaned and fully referenced electronic datasets, full transcripts of all interviews and focus group discussions, and a list of data collection sites.

Rating: 3

Standard: 3.6.2. The full report documents procedures intended to ensure confidentiality and to secure informed consent where necessary (in some cases this is not needed - e.g. evaluation synthesis - in which case N/A should be recorded)

Comment and Analysis: The report did not mention how procedures of confidentiality would be informed. Professional format in contractual obligations followed to ensure all documentation that isn't allowed for public knowledge remains secure. None of the interviewed schools/pupils or teachers were named in the written report. Only the DBE, UNICEF and ACER members would have had all that information.

Rating: 2

Standard: 3.6.3. There are no risks to participants in disseminating the original report on a public website

Comment and Analysis: The documentation was all in ACERs hands and no report was at any risk of harming the public or those who participated in the evaluation.

Rating: 4

Standard: 3.6.4. There are no unfair risks to institutions in disseminating the original report on a public website

Comment and Analysis: No unfair risks that could associate any kind of liability to institutions without permission to do so in publicly making the document available.

Rating: 4

3.7. Project management (Reporting phase)

Standard: 3.7.1. A project closure meeting that reflected on the challenges and strengths of the evaluation process occurred

Comment and Analysis: Meeting was held to close the project and also done to discuss the way forward on ways in which DBE can also use the evaluation further. Matters like how to use the evaluation as a tool to improvement if the workbooks and textbooks was discussed. Also various challenges within the project life cycle were discussed.

Rating: 4

4. Follow-up, use and learning

4.1. Resource utilisation

Standard:	4.1.1. The evaluation was completed within the planned timeframes
Comment and Analysis:	The originally envisioned timeframe for the evaluation in the ToR was 12 months, but in the original contract it was shortened to six month. The evaluation team delivered the assessment outside of this six month period following an extended deadline by DBE. With the new timeframe everything was completed as planned, but this was subject to multiple shifts.
Rating:	3

Standard:	4.1.2. The evaluation was completed within the agreed budget
Comment and Analysis:	Due to the extension and change of the scope a new budget needed to be considered, since the original budget did not account for the extra additions.
Rating:	2

4.2. Evaluation use

Standard:	4.2.1. Results of the evaluation have been presented to all relevant stakeholders
Comment and Analysis:	The final evaluation results were well presented to all stakeholders on the steering committee (representatives on the committee and those within the department that the evaluation would have an impact on) of the evaluation. It was not mentioned whether the results were shared to district representatives, SGBs and teachers and there was limited evidence that the results had been communicated down to this level.
Rating:	3

Standard:	4.2.2. A reflective process has been undertaken by the steering committee (if no steering committee exists then by the evaluation management team or the involved department officials) to reflect on what could be done to strengthen future evaluations
Comment and Analysis:	Although there was strong engagement from the steering committee throughout the process there was not any specific evidence that a dedicated reflective process was undertaken to determine what could be done to strengthen future evaluations. Regular engagement and process feedback does motivate for some minor in process reflection, but not at the conclusion of the assessment.
Rating:	2

Standard:	4.2.3. The evaluation study is seen by interviewed stakeholders as having added significant symbolic value to the policy or programme (e.g. raised its profile)
Comment and Analysis:	The interviewees saw value in the study as it helped highlight the different subject needs in order to better articulate where improvement for the workbooks would need to be focused. Because the DBE saw this an important task to better understanding the functioning of these workbooks and textbooks, the evaluation helped solidify and emphasize that there is great use and value embedded in these books for pupils around the country.
Rating:	4
Standard:	4.2.4. The evaluation study is of conceptual value in understanding what has happened and possibly in shaping policy and practice
Comment and Analysis:	The evaluation identified factors contributing to teachers' and pupils' use of workbooks and textbooks. In particular, how teachers should utilise the textbooks was one example. Issues where pupils either felt that some pages were too content heavy and not illustrative enough or where books were not fully addressing the issues were limiting efficacy. The study highlighted that issues that limited the use of the books effectively were: late or no delivery; lack of parent, teacher or pupil communication regarding aims and objectives of these books; and lack of monitoring on how they can be used. The study was of significant conceptual value in identifying the need for the department to take more of an active role in ensuring that this occurs.
Rating:	4
Standard:	4.2.5. Development of a draft improvement plan has been started, but not completed, based on the findings and recommendations set out in the evaluation
Comment and Analysis:	There was no draft improvement plan shared with the supporting documentation given by the DBE, although some recommendations have been acted upon. Nevertheless, there was no indication that this evaluation directly informed any specific improvement planning that followed on the assessment.
Rating:	2
Standard:	4.2.6. The report is publicly available (website or otherwise published document), except where there were legitimate security concerns *Note: only apply if sufficient time has elapsed since completion of the evaluation
Comment and Analysis:	The evaluation report can be found on the DBE website. It has been shared amongst the department with other employees who were not necessarily directly involved in the programme but who deal with the content delivery.
Rating:	4
Standard:	4.2.7. There is clear evidence of instrumental use - that the recommendations of the evaluation were implemented to a significant extent *Note: only apply if sufficient time has elapsed since completion of the evaluation
Comment and Analysis:	Yes, enough time has elapsed since the completion of the evaluation to show instrumental use. Most of the recommendations were taken up by the department. What the evaluators realized is that maybe 3 out of the 9 recommendations that came out of the evaluation were already in the process or DBE was already putting these in place.
Rating:	4

Standard:	4.2.8. There is clear evidence that the evaluation has had a positive influence on the evaluand, its stakeholders and beneficiaries over the medium to long term *Note: only apply if sufficient time has elapsed since completion of the evaluation
Comment and Analysis:	This evaluation shows strong results that the DBE is truly engaged with using most of the recommendation to make the most of what was conducted. However, it has only been about 2 years since the final evaluation report was given to the department, so one cannot fully attest to how the evaluation is making an influence over the medium - long term.
Rating:	N/A

References

- Australian Council for Educational Research (2011). Technical Report: An independent formative evaluation of textbooks and workbooks UNICEF, South Africa. RFP/SAFB/2011/073
- Department of Basic Education (2011). Statement of Work and Terms of Reference: Overview of Evaluation of Workbooks and Textbooks.
- The Australian Council for Educational Research (2011). Inception Report:Formative evaluation of text books and work books in South Africa. RFP/SAFB/2011/073.
- Department of Basic Education (2011). Formative evaluation of text/work books,Summary of screening of the technical proposals.
- Department of Basic Education (2012). Terms of Reference for Critical Readers. Evaluation of workbooks and textbooks. Framework for M&E curriculum implementation. 3 April, 2012.
- The Australian Council for Educational Research (2013). Formative Evaluation of Textbooks and Workbooks . RFP/SAFB/2011/073.
- Department of Basic Education (2012). Evaluation of Workbooks and Textbooks: Request for Extension of Project.
- The Australian Council for Educational Research (2013). Desk-Based Evaluation of Learning Materials in South Africa. Department of Basic Education, South Africa and The United Nations Children’s Fund, South Africa. 43139017.

List of Interviewees

- Ms Nompumelelo Mohohlwane, Assistant Director: Research Coordination, Monitoring and Evaluation, Department of Basic Education, In person (interview), 11 February 2014
- Mr Gerrit Maritz, Education Specialist (Quality Education), UNICEF South Africa, Telephone Interview, 10 February 2014
- Ms Rachel Outhred, Senior Research Fellow: ACER, Policy Analysis and Program Evaluation, Skype Interview, 24 February 2014
- Ms Carol Nuga Deliwe, Director Policy Support, Department of Basic Education, In person (interview), 11 February 2014