



planning, monitoring and evaluation

Department:
Planning, Monitoring and Evaluation
REPUBLIC OF SOUTH AFRICA

Report on the Assessment of Government Evaluations

Evaluation Title:

Expanding Opportunities for South African Youth
through Math & Science: The Impact of the Dinaledi
Program

Evaluation Number:

402

Evaluation Completion Date:

01 December 2010

Period of Evaluation:

2008-2010

Submitted:

26 November 2013 by Cathy Chames

Approved:

29 November 2013 by Mike Leslie

Evaluation Details

Evaluation Title:	Expanding Opportunities for South African Youth through Math & Science: The Impact of the Dinaledi Program
Evaluation Number:	402
Evaluation Completion Date:	01 December 2010
Created:	23 October 2013 by Mike Leslie
Submitted:	26 November 2013 by Cathy Chames
Approved:	29 November 2013 by Mike Leslie
Period of Evaluation:	2008-2010
Known Cost:	
Known Cost Type:	No Data
Initiated By:	The World Bank and National Department of Education
Initiated By Internal:	Yes
Undertaken By:	World Bank
Undertaken By Internal:	No

Assessors

Cathy Chames

cathy@southernhemisphere.co.za

Assessment Documents

Document Name:	Document Type:	Added By:	Added On:
State of Literacy 2012 NEEDU national report_final.pdf	Evaluation report	Mike Leslie	23 October 2013
DIME Report on SA DINALEDI Math and Science v2 docx (3).pdf	Evaluation report	Cathy Chames	14 November 2013
DINALEDI Math and Science v1 DIME Summary.pdf	Evaluation report	Cathy Chames	14 November 2013
CASE STUDY IMPACT OF THE DINALEDI PROGRAMME.pdf	Presentations of evaluation findings and recommendations	Cathy Chames	14 November 2013
Action Plan to 2014 FINAL.pdf	Any other relevant documentation pertaining to the evaluation process	Cathy Chames	14 November 2013
Assessment Report at Moderation.pdf	Assessment Report at Moderation	Mike Leslie	25 November 2013

Quality Assessment Summary

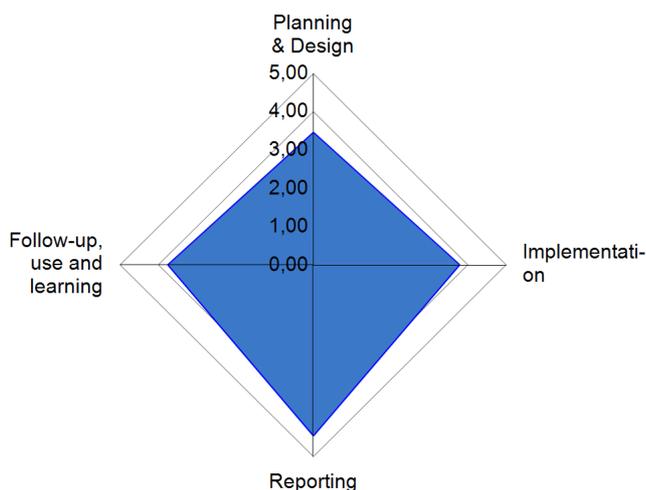
The overall score for this evaluation is 3.98. The planning and design phase scored lower at 3.46. A key weakness of this phase is that there was no TOR or internal proposal which documents and formalises all planned aspects of the evaluation. Despite this, extensive meetings were held between DOE and World Bank to discuss the choice of programme to evaluate and the overall design of the evaluation. Another challenge experienced during this phase was the length of time DOE took to decide on which programme to evaluate which then impacted on the financial resources and time which were allocated to the task. Strengths of this phase are the good alignment of the study to the policy context; and the fact that a truly participatory approach was followed to ensure that stakeholders reached a common agreement on how the evaluation would be implemented. A weakness in this participatory approach, however, is that the team worked with a small group of national level DOE stakeholders and did not include provincial level stakeholders. This would have strengthened the overall partnership approach to the evaluation which scored 3.50 - the lowest score amongst the overarching considerations. The Implementation phase of the study scored 3.79 and an aspect which contributed to this score was the independence of the evaluation team with no conflict of interest. This contributed to the high score of 5.00 for evaluation ethics as an overarching consideration. The methodological integrity for this phase scored 3.29. The main shortcoming noted here is that, the use of secondary data only to determine the impact of the programme meant that the study was more focused on the outputs of Dinaledi rather than on the outcomes. Thus many questions were left unanswered about the factors which are causally linked to the results. In addition, the methods and analysis used were deemed to be highly technical in nature resulting in very limited engagement with key stakeholders as part of the methodology. Despite this, the report is written in such a way that its content is accessible. Overall the report scores the highest at 4.45. It is well-structured (score of 4.63), and the findings are robust (score of 5) due to the well-executed data analysis strategy, frequent acknowledgement of limitations of the methods used for data analysis, and recognition of the possibility of alternative interpretations of data. With regard to the follow up use and learning phase, the main weakness was poor resource utilisation because both the time frames and budget had to be extended. Besides this, the evaluation use scores 4.63 and there is clear evidence from both interviewees and the DBE Action Plan towards the Realisation of Schooling (2014) that there has been instrumental use of the evaluation to improve the design of the programme for the next phase of implementation. The positive evaluation results have also been used by the DOE at the time to secure funding and upscale the Dinaledi programme.

Quality Assessment Scores

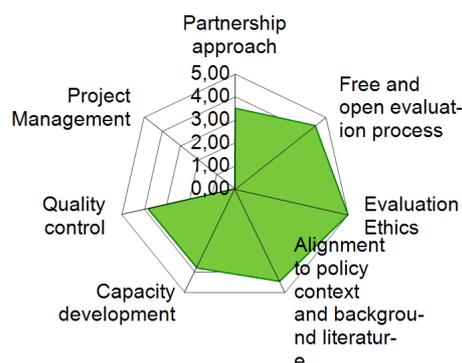
Phase of Evaluation	Score
Planning & Design	3,46
Implementation	3,79
Reporting	4,45
Follow-up, use and learning	3,76
Total	3,98

Overarching Consideration	Score
Partnership approach	3,50
Free and open evaluation process	4,42
Evaluation Ethics	5,00
Alignment to policy context and background literature	4,45
Capacity development	3,80
Quality control	3,86
Project Management	
Total	3,98

Scores: Phases of Evaluation



Scores: Overarching Considerations



Phase of Evaluation	Area of Evaluation	Score
Planning & Design	Quality of the TOR	
Planning & Design	Adequacy of resourcing	2,82
Planning & Design	Alignment to policy context and background literature	4,00
Planning & Design	Appropriateness of the evaluation design and methodology	3,61
Planning & Design	Project management (Planning phase)	4,00
Implementation	Evaluation ethics and independence	5,00
Implementation	Participation and M&E skills development	4,00
Implementation	Methodological integrity	3,29
Implementation	Project management (Implementation phase)	2,00
Reporting	Completeness of the evaluation report	4,63
Reporting	Accessibility of content	4,53
Reporting	Robustness of findings	5,00
Reporting	Strength of conclusions	3,71
Reporting	Suitability of recommendations	3,77
Reporting	Acknowledgement of ethical considerations	4,70
Reporting	Project management (Reporting phase)	4,00
Follow-up, use and learning	Resource utilisation	1,00
Follow-up, use and learning	Evaluation use	4,63
Total	Total	3,98

Planning & Design

Quality of the TOR

Standard:	The evaluation was guided by a well-structured and complete TOR or a well-structured and complete internal evaluation proposal (e.g. Background, Purpose, Evaluation Questions, Design & Methodology, Deliverables & Timeframes, Resource requirements, Intended Audience & Utilisation, etc).
Comment and Analysis:	There was no TOR or internal evaluation proposal available for review. The project was initiated by the World Bank who entered into a partnership with The Presidency in 2006 to conduct impact evaluations of major government programmes. An initial workshop was held with government departments to introduce the project and then departments entered into a process of selecting appropriate programmes to evaluate. The National Department of Education chose to evaluate the impact of the Dinaledi Programme. Since The World Bank financed the agreement it was felt that a TOR was not necessary. However, it was indicated from the programme manager at the time that a internal proposal would have been developed. A copy of this document was not available at the time of this assessment despite attempts to locate it from the current DBE Research Unit.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted

Standard:	The purpose of the evaluation stated in the TOR (or an internal evaluation proposal) was clear and e
Comment and Analysis:	Since there was no TOR or internal evaluation proposal, the purpose of the evaluation was not clearly and explicitly documented at the outset. However, interviewees indicated that a considerable amount of time was spent deciding on which programme to evaluate and the purpose of the evaluation.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted

Standard:	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:	There was no TOR or internal proposal for this evaluation. However, interviewees indicated that a participatory approach was used the outset of the study where the National Department of Education expressed a need to improve their understanding of whether the Dinaledi programme should be expanded or not. Thus it is felt that the evaluation approach was well suited to the intended purpose and scope.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted

Standard:	The TOR (or an internal evaluation proposal) identified the intended users of the evaluation and their information needs
Comment and Analysis:	There was no TOR or internal proposal for this evaluation. The interviewees indicated that the main intended user would be the National Department of Education which would use the results of the evaluation to determine whether it should continue to expand the Dinaledi Programme.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted

Standard:	The evaluation questions in the TOR (or an internal evaluation proposal) were clearly stated and ap
Comment and Analysis:	There was no TOR or internal evaluation proposal for this task. These key evaluation questions are reflected in the evaluation report.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted

Standard:	Key stakeholders were involved in the scoping of the TOR and choosing the purpose of the evaluation
Comment and Analysis:	There was no TOR for this task. There was, however, extensive consultation between the World Bank and National level DOE officials when deciding on the purpose, scope, approach and method to be used for this evaluation. However, these were not formally captured in a TOR or evaluation proposal. Furthermore, no provincial level stakeholders were consulted.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted

Adequacy of resourcing

Standard:	The evaluation was adequately resourced in terms of time and budget allocated
Comment and Analysis:	Once the data was received from DOE the evaluation took about one and a half years to complete - from data analysis to dissemination of findings. It was reported that the analysis of data took much longer than anticipated. The DOE interviewee confirmed that the report was produced in reasonable time frames.
Rating:	3: The evaluation was adequately resourced in terms of time and budget allocated
Moderation:	Accepted
Approval:	Accepted

Standard:	The evaluation was adequately resourced in terms of original budget
Comment and Analysis:	The study was funded by the World Bank. However, due to the extra time it took for the Department of Education to select the appropriate programme for the evaluation and the extended time it took to analyse the data, the evaluation team had to apply for additional financial resources from the Nordic Trust Fund for Secondary Education.
Rating:	1: 1
Moderation:	Reconsider
Moderation Comment:	Rating is fine, just revise grammar and punctuation.
Approval:	Accepted

Standard:	The team conducting the evaluation was adequately resourced in terms of staffing and skills sets
Comment and Analysis:	The evaluation was well resourced in terms of staff and skills. The team leader included the Head of the Development Impact Evaluation Initiative (DIME) at the World Bank. The technical team consisted of four experts from the World Bank and three officials from the National Department of Education.
Rating:	5: The staffing and skills sets required for the evaluation were ideal for the evaluation purpose, sector and incorporated high quality international expertise
Moderation:	Reconsider
Moderation Comment:	Please revise comment to remove extra "Head of the..."
Approval:	Accepted

Standard:	Where appropriate, the evaluation planned to incorporate an element of capacity building of partners
Comment and Analysis:	According to the World Bank interviewees, an overall strength of this evaluation was its capacity building approach which was planned from the outset. The World Bank works very closely with clients in a joint evaluation process to build capacity on how to make more systematic use of evidence to inform policy design and decision making. In practice, this process entails working with stakeholders on the ground in order to build understanding and capacity on how to collect, analyse and use data for decision making. For this particular study the transfer of skills was around the dimension of organising and collecting administrative data so that it can be used for evidence. The DOE interviewee neither confirmed nor refuted this aspect of the evaluation but did state that a key learning from the study was the importance of gathering evidence for policy making. It was further noted by one interviewee that since the evaluation methodology was complex, the programme staff were more interested in the evaluation findings rather than the evaluation process.
Rating:	3: 3
Moderation:	Accepted
Approval:	Accepted

Alignment to policy context and background literature

Standard:	There was evidence that a review of the relevant policy and programme environments had been conducted
Comment and Analysis:	The World Bank team did review the relevant policy and programme documents in order to familiarise themselves with the programme context. Evidence of this review is contained in the evaluation report.
Rating:	4: 4
Moderation:	Reconsider
Moderation Comment:	The comment does not substantiate a rating of 5 as 'excellent' as it is merely acknowledges that a review took place and that there was evidence of this. If the World Bank team did do an excellent job of this, consider providing a comment as to what distinguishes this as better than good, or else consider revising the rating down.
Approval:	Accepted

Standard:	There was evidence of a review of appropriate literature having been conducted and used in planning
Comment and Analysis:	The evaluation team reviewed relevant literature in planning the research particularly around deciding on the research methodology required in order to estimate the effect of the Dinaledi Programme on physical sciences learning outcomes. The team also did a review of international benchmarks for maths and science results in other African countries at similar economic and human development levels in order to compare South Africa's performance. Evidence of this is contained in the evaluation report.
Rating:	4: 4
Moderation:	Reconsider
Moderation Comment:	The same comment as the above applies here. The above comment indicates that the standard was met, yet the rating is indicative of excellence. Consider substantiating what makes it excellent or revising down the rating.
Approval:	Accepted

Appropriateness of the evaluation design and methodology

Standard:	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:	Interviewees confirmed that during the initial planning stage of the study, discussions were held around the theory of change and intervention logic of the Dinaledi programme and time was spent analysing some of the issues affecting implementation.
Rating:	4: The intervention logic or theory of change meaningfully informed and shaped the TOR or the Inception Report, including a visual representation
Moderation:	Reconsider
Moderation Comment:	Remove second "and" from "and and time was spent" in middle line.
Approval:	Accepted

Standard:	Key stakeholders were consulted on the design and methodology of the evaluation
Comment and Analysis:	A participatory approach was used when designing the evaluation study and The World Bank consulted extensively with key stakeholders from the National DOE in this regard. Interviewees indicated that, since the study only made use of secondary data from the DOE, it was not necessary to consult with other stakeholders, however, it may have been useful to get input from provincial level departments.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Standard:	The planned methodology was appropriate to the questions being asked
Comment and Analysis:	The study made use of administrative data from the DOE in order to establish whether the Dinaledi programme has had an impact on the number of students in secondary schools studying and passing maths and science examinations. The DOE interviewee confirmed that the use of secondary data was the most appropriate method to use. The evaluation team from the World Bank indicated that, whilst the use of administrative data only was appropriate, it did have limitations because the the evaluation focused on outputs only rather than on outcomes.
Rating:	3: The planned methodology was appropriate to the questions being asked
Moderation:	Accepted
Approval:	Accepted

Standard:	The sampling planned was appropriate and adequate given the focus and purpose of evaluation
Comment and Analysis:	Sampling was appropriate to the purpose of the evaluation. All 350 Dinaledi schools were selected for the evaluation and thus purposive sampling. In addition, each school was matched to a control school with similar observed characteristics.
Rating:	4: The sampling planned was good given the focus, purpose and context of the evaluation
Moderation:	Reconsider
Moderation Comment:	Please consider revising the comment because the second sentence is unclear. If the assessment provides a 5 for this rating please be explicit as to why it is excellent.
Approval:	Accepted

Standard:	There was a planned process for using the findings of the evaluation
Comment and Analysis:	Interviewees indicated that the DOE intended to use the findings from the evaluation to inform their decision making around whether to expand the Dinaledi programme or not. However, it appears that there was no planned process for doing this.
Rating:	3: 3
Moderation:	Accepted
Approval:	Accepted

Project management (Planning phase)

Standard:	The inception phase was used to develop a common agreement on how the evaluation would be implemented
Comment and Analysis:	Interviewees indicated that a considerable amount of time was taken during the inception phase of the evaluation to develop a common agreement with DOE stakeholders on what would be evaluated; how it would be evaluated; why it would be evaluated and what data could be used. Whilst it was good to develop this common agreement it was noted that too much time may have been spent on this phase of the programme.
Rating:	4: The inception phase was used to good effect to achieve a common agreement and understanding of how the evaluation would be implemented
Moderation:	Accepted
Approval:	Accepted

Implementation

Evaluation ethics and independence

Standard:	Where data was gathered in contexts where ethical sensitivity is high, informed consent, assurances of confidentiality and appropriate clearance were achieved; e.g. through an ethics review board, in evaluation involving minors, institutions where access usually requires ethical or bureaucratic clearance
Comment and Analysis:	The evaluation made use of secondary data only and therefore ethical issues were not applicable.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted

Standard:	Where external, the evaluation team was able to work without significant interference and given access to existing data and information sources
Comment and Analysis:	The evaluation team confirmed that there was no interference from the DOE - once they had received the data there was no scope for interference. The DOE interviewee corroborated this indicating that the evaluation team analysed the administrative data that was handed to them and that they worked independently.
Rating:	5: The evaluation team was able to work freely and independently without interference and significant efforts were documented to ensure unfettered access to all existing data and information sources
Moderation:	Accepted
Approval:	Accepted

Standard:	The evaluation team was impartial and there was no evidence of conflict of interest
Comment and Analysis:	The World Bank team was impartial and there is no evidence of conflict of interest.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

Participation and M&E skills development

Standard:	Key stakeholders were involved in the evaluation through a formalised mechanism or institutional arrangement
Comment and Analysis:	An internal technical team was set up within the DOE to oversee the study. It consisted of members from the research and planning branch; and the curriculum branch. They worked together with the two members of the World Bank technical team.
Rating:	4: Key stakeholders were regularly, actively involved in the evaluation and contributed through a formalised mechanism or institutional arrangement (e.g. a steering committee or reference group)
Moderation:	Accepted
Approval:	Accepted

Standard:	Where appropriate, an element of capacity building of partners responsible for the evaluand and evaluators was incorporated into the evaluation process
Comment and Analysis:	The evaluation team indicated that the process of transferring skills and knowledge was mainly around the dimension of organising and collecting administrative data in DOE. Time was taken to understand how data was collected and what needs to be done in order to make data more 'usable' for evidence based programming and policy making in the future.
Rating:	4: Structured capacity building of evaluators and partners responsible for the evaluand was incorporated into the evaluation process
Moderation:	Accepted
Approval:	Accepted

Methodological integrity

Standard:	The methods employed in the process of the evaluation were consistent with those planned
Comment and Analysis:	All evidence suggests that the planned method of collecting and analysing secondary administrative data from the DOE was adhered to in the process of the evaluation. However, the absence of an existing ToR or internal project proposal against which to corroborate this means that certainty of methodological fidelity to planning cannot be triangulated
Rating:	4: 4
Moderation:	Reconsider
Moderation Comment:	Consider revising. Although all evidence would appear to suggest this to be the case, the absence of an existing ToR or internal project proposal against which to corroborate this means that certainty of methodological fidelity to planning cannot be triangulated and thus may keep this from being a 5.
Approval:	Accepted

Standard:	The methods employed in the process of the evaluation were consistent with those planned and implemented adequately
Comment and Analysis:	The evaluation made use of secondary administrative data only - no empirical work was done. The data was sourced from the DOE data management systems and included the following: EMIS, NEIMS and PERSAL data. The data was easily available and the DOE was fully collaborative and assisted the evaluation team to access the information systems. It was noted by the evaluation team that the use of administrative data was within the scope of this study, however, it meant that the focus of the study was more on outputs than on outcomes. Thus there were still many questions left unanswered about which factors are causally link to the results.
Rating:	3: The methods employed in the process of the evaluation were consistent with those planned and implemented adequately
Moderation:	Accepted
Approval:	Accepted

Standard:	A pilot of basic data collection instrumentation occurred prior to undertaking data collection and it was used to inform the research process
Comment and Analysis:	A retrospective analysis approach was used to analyse the administrative data. Propensity score matching and difference in difference strategy were used to measure student maths and science uptake and pass rates in Dinaledi schools relative to student behaviour and achievements in similar non-Dinaledi schools. The study also analysed the impact of Dinaledi in former Bantustan schools. Overall this combined strategy was well-suited to the purpose of this evaluation because it served to distinguish between programme effects and those arising from pre-existing differences between Dinaledi and non-Dinaledi schools. Furthermore, the propensity score matching minimises observable differences by matching each Dinaledi school with a control school that had similar observed characteristics based on an aggregate indicator of similarity (the propensity score). The evaluation team noted that data cleaning was an issue in this study and they also spent much longer than expected (3 months) on cleaning up the data, merging the data sets and trying to find the schools that "matched".
Rating:	4: All components of the data collection instrumentation were piloted which led to some improvements in the data collection instrumentation or affirmation of the instruments
Moderation:	Accepted
Approval:	Accepted

Standard:	Data collection was not compromised by fieldwork-level problems or unplanned diversions from original intentions
Comment and Analysis:	The evaluation team gathered two sets of secondary administrative data from the DOE. Whilst the datasets were fairly easy to access, the evaluation team spent much time and effort in trying to understand the DOE's data collection process and some challenges were experienced around the quality and completeness of data. For example, when the evaluation team attempted to merge the two datasets based on national school identification numbers and exam centre numbers, it was found that the quality and completeness of data on school characteristics varied significantly between provinces and this presented an obstacle during the study. Furthermore, whilst the outcome data on student test participation for science was detailed and comprehensive, the pre-treatment data (for 2004) for mathematics was unreliable and inaccurate.
Rating:	3: 3
Moderation:	Reconsider
Moderation Comment:	Although the project made use of secondary data, there was clearly an effort made to understand the data collection process with the Department and to access administrative data that was produced. I think this could still allow for unplanned diversions from original intentions. For instance, maybe there is clearly incorrect or incomplete data available? Or possibly the administration of the data was more complex? Consider revising the comment in line with a broader understanding of how secondary data collection could have been compromised in this instance, especially in light of the challenges experienced with data cleaning acknowledged in the later standard.
Approval:	Accepted

Standard:	Data was collected from key stakeholders (e.g. implementers, governance structures, indirectly affected stakeholders) as data sources
Comment and Analysis:	The executive summary is two pages long and provides a clear and concise summary of the key components of the report including the programme context, the approach and method, evaluation results, and policy recommendations.
Rating:	5: Data was collected from all of the key stakeholder groupings identified in the research plan and the intended sample was well achieved (approx. 90-100% of those intended)
Moderation:	Accepted
Approval:	Accepted

Standard:	The methodology included engaging beneficiaries appropriately as a key source of data and information
Comment and Analysis:	The context of the development intervention is presented in the report. The background section provides a good overview of Maths and Science Education in South Africa. It covers the impact of apartheid on the inequalities in Maths and Science performance between different ethnic groups and then presents some international benchmarks against which South Africa fares very poorly. The background to the Dinaledi programme provides a good overview of the programme objectives, intended results and process of implementation. The report does not refer to previous evaluations or studies conducted on the Dinaledi programme.
Rating:	4: The methodology included meaningfully engaging beneficiaries as a primary source of data and information (or if based on secondary data, includes data from beneficiaries and beneficiaries consulted on emerging findings)
Moderation:	Reconsider
Moderation Comment:	Revise last sentence "where were".
Approval:	Accepted

Standard:	Key stakeholders were significantly engaged as part of the methodology
Comment and Analysis:	Whilst much time was spent consulting key stakeholders in the DOE around the planning and design of the study methodology, the data analysis and report writing was done by the World Bank technical team. Thus DOE engagement in the methodology was limited. It was noted by interviewees that one stakeholder in the DOE statistics department was involved in cross-referencing the data and thus new skills were acquired in this regard. DOE interviewees indicated that the methodology was highly complex which few people understood which made it a fairly disempowering process.
Rating:	2: 2
Moderation:	Accepted
Approval:	Accepted

Standard:	The methodology included engaging beneficiaries appropriately as a key source of data and informatio
Comment and Analysis:	The evaluation made use of secondary administrative data only. No primary data collection was undertaken by the evaluation team.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted

Project management (Implementation phase)

Standard:	The evaluation was conducted without significant shifts to scheduled project milestones and timefram
Comment and Analysis:	There were significant shifts to the planned time-frames for the study and the evaluation team highlighted two key factors which impacted on these shifts. Firstly, the DOE took a long time to decide on what should be evaluated and what data they were could share; and secondly, the data cleaning and analysis took longer than expected (3 months). Thus the time-frames for the study were extended. This had budget implications for the World Bank who had to access additional funding to complete the study. The DOE interviewee, however, did not highlight any issues around shifts to milestones and time frames.
Rating:	2: 2
Moderation:	Reconsider
Moderation Comment:	Consider revising the score. The comment and analysis would suggest there were significant shifts to time-frames, especially around the availability of a data-st ready for analysis. Although this might not have had serious implications for DOE, shifting time-frames and additional costs for the World Bank would suggest this was less than adequate, especially because not actual primary data collection took place.
Approval:	Accepted
Standard:	The steering committee, technical working group and service provider worked together adequately to facilitate achievement of the objectives of the evaluation
Comment and Analysis:	The main evaluation question for this study is: did and can a programme like Dinaledi significantly contribute to expanding the number of students in secondary schools study and passing maths and science exams? The rationale for the study is clearly stated in the introductory section of the report - that very little rigorous evidence exists on the impact of policies aimed at increasing labour skills in maths and science which are of paramount importance for a country's growth in the global economy. Thus the purpose of the study is to provide the government with a measure of programmes effectiveness in meeting its objectives that can guide programme expansion in the future.
Rating:	4: The steering committee, technical working group and service provider worked together in a flexible and constructive manner facilitating achievement of the objectives of the evaluation
Moderation:	Accepted
Approval:	Accepted

Reporting

Completeness of the evaluation report

Standard:	The scope or focus of the evaluation is apparent in the report
Comment and Analysis:	The scope and focus of the evaluation is clearly stated in the introductory section of the report: to exploit administrative data to estimate the effect of enrolment in the Dinaledi programme for about 350 schools added to the programme in 2004.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Standard:	A detailed methodology is outlined in the relevant section of the report to the point that a reader
Comment and Analysis:	The methodology used for the study is well presented in the report. It starts with the Data section of the report which provides a thorough description of the data that was used for the study and is then followed by Identification Strategy which explains the choice of the DID-matching identification strategy for estimating the Dinaledi programme's effect and discusses underlying assumptions. Whilst the content is highly technical in nature the authors use simple, clear language which makes the content more accessible. The report includes four technical appendices which discuss methodological issues in further detail.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

Standard:	Acknowledgement of limitations of all aspects of the methodology and findings are clearly and succinct
Comment and Analysis:	The report provides extensive detail on the methodological limitations of retrospective analysis of administrative data and two of the technical appendices discuss major threats to the internal and external validity of the estimated programme effects and the plausibility of the identification strategy.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

Standard:	Key findings are presented in a clear way; they are made distinct from uncertain or speculative find
Comment and Analysis:	The results are clearly and succinctly presented and include average treatment effects on Dinaledi schools (for maths and science); and heterogeneous treatment effects (by former department under apartheid and by province) . Unused data is included in the appendices.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

Standard:	Conclusions and recommendations are clear and succinctly articulated
Comment and Analysis:	The concluding section of the report is clear and brief with a limited number of well constructed recommendations.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

Accessibility of content

Standard:	The final evaluation report is user-friendly, written in accessible language and adequate for publication (e.g. adequate layout and consistent formatting; complete sentences and no widespread grammatical or typographical errors; consistency of style and writing conventions; levels of formality; references complete and consistent with cited references in reference list and vice versa; etc.)
Comment and Analysis:	Data has been sufficiently and appropriately analysed using a combination of propensity score matching and difference in difference strategy (DID) in order to measure the causal linkage between the Dinaledi intervention and learning outcomes, i.e. the programme effect. Propensity score matching minimises observable differences by matching each Dinaledi school with a control school that had similar observed characteristics based on an aggregate indicator of similarity. DID seeks to eliminate bias from unobserved pre-existing differences (such as school management) by comparing Dinaledi and non-Dinaledi schools with regard to the change in learning outcomes over time.
Rating:	5: The final evaluation report balances an impressive depth of work with excellent writing that is accessible to the common reader and reflects an excellent publishing standard
Moderation:	Accepted
Approval:	Accepted

Standard:	The final evaluation report is user-friendly, written in accessible language and its content follows
Comment and Analysis:	The report is well structured and follows a clear logic. It begins with a background section, followed by description of data and strategies used for identification of the Dinaledi programme's effect and then goes on to present the results and the conclusions. The Data section of the report uses statistical language which is technical in nature, however, the authors use clear, simple explanations for some of the terminology which makes the content more accessible. The authors have also produced three-paged 'brief' of the report which is easy to read. A DOE interviewee confirmed that the abridged version of the report was very user friendly, 'readable', and useful for sharing the findings with all stakeholders.
Rating:	5: 5
Moderation:	Reconsider
Moderation Comment:	Edit third and fourth sentences.
Approval:	Accepted

Standard:	Figures, tables and 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.) and are readily discernible to a reader familiar with data presentation conventions
Comment and Analysis:	The administrative data for the study was subject to rigorous analysis which provides good evidence to support the two main findings around maths and science enrolment and examination results.
Rating:	5: A wide variety of figures, tables and different data presentation conventions are used to an expert standard with explanations and additional references supporting comprehension by all readers
Moderation:	Accepted
Approval:	Accepted

Standard:	Quality of writing and presentation is adequate for publication including: adequate layout and consi
Comment and Analysis:	Some spelling mistakes and a few grammatical errors were noted throughout the report. However, this does not impact on the quality of writing which is generally good. The layout is adequate and formatting is consistent throughout. The report contains a bibliography.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Standard:	Appropriate conventions are used in presentation of data (e.g. use of appropriate statistical langua
Comment and Analysis:	The bulk of the report focuses on the presentation of data using appropriate statistical language. Issues with data quality and biases in data are noted throughout and the validity of data and reliability of findings are also acknowledged.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

Standard:	The use of figures and tables is such that it supports communication and comprehension of results; a
Comment and Analysis:	The structure of the figures and tables are good with clear headings for all variables. They provide a good visual representation of the data. All except one of the figures and tables are contained in the report annexure making it difficult to refer to them whilst reading.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Robustness of findings

Standard:	Data analysis appears to have been executed to an adequate standard
Comment and Analysis:	The report appears free of significant methodological and analytic flaws and authors frequently emphasise the limitations of the methodology and analysis used throughout the report.
Rating:	5: Data analysis is thorough and well executed to an exceptional standard for all datasets
Moderation:	Accepted
Approval:	Accepted

Standard:	Findings are supported by available evidence
Comment and Analysis:	The administrative data, which included information from 3,347 schools in South Africa, was subject to a number of levels of analyses. The Data section of the report presents a description of the observed differences between Dinaledi and non-Dinaledi schools and highlights the disparities in maths and science learning outcomes by province and by schools' former attribution to the different racially segregated education administrators under apartheid. Trends in the data are also presented including how the number and pass rates of maths and science for Senior Certificate examination takers have evolved before and after the 2004 baseline year; and the contribution of Dinaledi 2005 schools to these trends. This initial findings here are that South Africa's legacy of apartheid, in combination with its geographic and socio-economic diversity, is responsible for great variance in maths and science learning outcomes and in access to quality schooling. The next layer of analysis is presented in the section on Identification Strategy which details the the twin identification strategy used to estimate the Dinaledi programme effect on learning outcomes. This strategy is applied because it is difficult to estimate the Dinaledi programme's effect due to unobserved differences between Dinaledi and other schools before the inception of the programme. These different layers of analysis provide sufficient evidence for the overall results presented in the final section of the report.
Rating:	5: 5
Moderation:	Reconsider
Moderation Comment:	The comment would appear to speak to the main conclusions of the report, rather than the initial findings arising from the data. Consider revising to speak more to the findings and analysis as supported by evidence, rather than the key conclusions.
Approval:	Accepted

Standard:	Findings are supported by evidence which is sufficiently and appropriately analysed to support the argument, integrating sources of data
Comment and Analysis:	The conclusions do not refer to other relevant research studies and evaluations which have been undertaken in the education sector.
Rating:	1: There is little to no analysis of the evidence gathered in relation to the positions taken in report (purely descriptive)
Moderation:	Accepted
Approval:	Accepted

Standard:	There is appropriate recognition and exploration of the possibility of alternative interpretations
Comment and Analysis:	The conclusions to upscale the programme and the importance of targeting particular schools going forward addresses the original purpose of the evaluation which was to provide evidence of programme effectiveness to guide future expansion of the programme. The main evaluation question posed at the outset was: did and can a programme like Dinaledi significantly contribute to expanding the number of students in secondary schools study and passing maths and science exams? The conclusions clearly address this original question.
Rating:	5: Alternative interpretations of data are presented and their validity is convincingly dispelled through critical analysis
Moderation:	Accepted
Approval:	Accepted

Standard:	The report appears free of significant methodological and analytic flaws
Comment and Analysis:	Although the conclusion does not explicitly refer to the intervention logic, it is implicit in the main conclusion of the study which is that the Dinaledi programme was effective in expanding enrolment rates in passing rates in maths and science.
Rating:	3: The report appears free of significant methodological and analytic flaws
Moderation:	Reconsider
Moderation Comment:	Consider revising the score downwards as references are implicit rather than explicit as the standard states.
Approval:	Accepted

Standard:	There is appropriate recognition of the possibility of alternative interpretations
Comment and Analysis:	At the outset, the authors are careful to highlight issues related to the validity of the findings and point out that the positive findings of this evaluation should be treated with "cautious optimism" because the analysis was conducted using the best available methods given the time and data constraints. It is also noted that the estimated programme effect on mathematics learning outcomes are likely to be biased as they are based on strong assumptions due to lack of baseline controls. Hence the results are not conclusive.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

Standard:	Limitations of all aspects of the methodology and findings are clearly articulated (e.g. limitations of scope or evaluation design, recommendation for additional research, data collection challenges, etc)
Comment and Analysis:	The DOE stakeholders did not play a role in shaping the recommendations. Instead they were presented by the evaluation team at a workshop following the finalisation of the report. Interviewees noted that, since the findings on the impact of the programme were positive they were readily accepted by DOE stakeholders together with the recommendations.
Rating:	3: Limitations of all aspects of the methodology and findings are articulated
Moderation:	Reconsider
Moderation Comment:	Consider expanding on the comment to give an indication as to whether the DOE stakeholders accepted these recommendations, further discussed and shaped them, or rejected them outright. At present the comment is more a statement of how it happened, than whether it was meaningful consultation.
Approval:	Accepted

Strength of conclusions

Standard:	Conclusions are derived from evidence
Comment and Analysis:	In general there is no evidence to suggest that a participatory approach was used to shape recommendations. The recommendations were made independently by the evaluation team and then presented at a workshop hosted by DOE. It was attended by government officials, academics and other interested parties. Interviewees confirmed that the recommendations were straightforward and DOE was keen to upscale the Dinaledi programme, thus they were readily adopted.
Rating:	3: Conclusions are derived from evidence
Moderation:	Reconsider
Moderation Comment:	The standard assumes a participatory approach with regards to recommendations and seeks to determine whether the recommendations were influenced by participatory inputs or whether they were made independent of this. Could the recommendations be said to be shaped with official and stakeholder engagement, or were these purely technical recommendations that were presented to government and adopted?
Approval:	Accepted

Standard:	Conclusions are derived from evidence
Comment and Analysis:	The authors conclude that the positive results from the data analysis provide support for scaling up the Dinaledi programme and targeting schools that have fewer resources and that service disadvantaged populations. The quantitative results from the data analysis provide sufficient evidence for these conclusions.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

Standard:	Conclusions address the original evaluation purpose and questions
Comment and Analysis:	The recommendations are targeted at the Department of Education. The recommendation to upscale the programme and improve targeting were feasible and acceptable. Recommendations which highlighted the need to continue measuring the effectiveness of the programme impact and to incorporate a learning strategy into the design of the next phase of Dinaledi were particularly relevant as they will provide an evidence-base for improving policies and programming in the future.
Rating:	5: The conclusions are exceptional in the manner that they address the evaluation purpose and questions
Moderation:	Reconsider
Moderation Comment:	Revise grammar in the last sentence.
Approval:	Accepted

Standard:	Conclusions are drawn with explicit reference to the intervention logic or theory of change
Comment and Analysis:	The report covers the limitations of conducting an impact evaluation using administrative data only. It acknowledges that many questions remain unanswered such as: what are the mechanisms through which the Dinaledi programme delivered these results? Why was the Dinaledi programme particularly effective in former Bantustan schools? Will the programme be effective in the remaining schools? Are these short-term effects or will they be sustained over time?
Rating:	5: The conclusions are exceptional in the manner that they provide a judgement on the intervention logic or theory of change and are clearly linked to design recommendations
Moderation:	Accepted
Approval:	Accepted

Suitability of recommendations

Standard:	Recommendations are made in consultation with relevant government officials, stakeholders and sectoral experts
Comment and Analysis:	The evaluation made use of secondary data only and therefore issues related to confidentiality and informed consent were not applicable.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted

Standard:	Recommendations are useful- they are relevant, specific, feasible, affordable and acceptable
Comment and Analysis:	The results were presented to a workshop hosted by DOE and attended by a range of stakeholders including government officials and academics. Furthermore, they were shared with the inter-provincial committee for maths and science and shared at the committee meeting attended by the Heads of Education and also at the HEDCOM which is a committee made up of Heads of Departments from various provincial departments.
Rating:	4: Recommendations are well-formulated for use- they begin to differentiate by user and are relevant to the current policy context, specifically targetted, feasible to implement, affordable and acceptable to key stakeholders
Moderation:	Accepted
Approval:	Accepted

Standard:	Recommendations are relevant to the current policy context
Comment and Analysis:	The policy recommendations are relevant to the current policy context since the results of the study provide evidence to upscale the programme and to target Dinaledi to former Bantustan schools in order to be more effective.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Acknowledgement of ethical considerations

Standard:	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 evaluation was fully funded by the World Bank. The evaluation was not completed within the planned budget. Due to the extended time-frames for the study, the World Bank had to raise more resources from the Nordic Trust Fund for secondary Education.
Rating:	1: The full report fails to document any procedures to ensure confidentiality and secure informed consent where appropriate.
Moderation:	Accepted
Approval:	Accepted

Standard:	Peer review of the draft evaluation report occurred prior to finalisation of the evaluation report
Comment and Analysis:	The draft report was subject to an internal peer review process within the World Bank. It was also submitted to DOE for review and input particularly into the more descriptive sections of the report.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Standard:	There are no risks to participants or institutions in disseminating the evaluation report on a public website
Comment and Analysis:	The full report and DIME brief is publicly available on the World Bank website.
Rating:	5: All participants and institutions to the evaluation were formally informed that the original report would be disseminated on a public website and no risks exist
Moderation:	Accepted
Approval:	Accepted

Standard:	There are no risks to participants in disseminating the original report on a public website
Comment and Analysis:	The focus of this study was on secondary data and no schools or learners are identified in the report, thus there is no risk to disseminating the report on a public website.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

Standard:	There are no unfair risks to institutions in disseminating the original report on a public website
Comment and Analysis:	The results around the impact of the Dinaledi programme are positive and there is no unfair risk to the DOE disseminating the original report which is already available on a public website.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

Project management (Reporting phase)

Standard:	A project closure meeting that reflected on the challenges and strengths of the evaluation process o
Comment and Analysis:	All interviewees confirmed that a meeting was held at the end of the evaluation in order to discuss the key learnings from the evaluation process with a particular focus on the type of methodology that was used.
Rating:	4: 4
Moderation:	Reconsider
Moderation Comment:	Consider revising to distinguish and substantiate what made such a meeting 'excellent' as opposed to 'good'.
Approval:	Accepted

Follow-up, use and learning

Resource utilisation

Standard:	The evaluation was completed within the planned timeframes
Comment and Analysis:	The evaluation was not completed in the planned timeframes. This was due to the unanticipated length of time it took the evaluation team to clean and analyse the data.
Rating:	1: 1
Moderation:	Accepted
Approval:	Accepted

Standard:	The evaluation was completed within the planned timeframes and budget
Comment and Analysis:	The positive findings of this impact evaluation has provided the evidence required to motivate for the funding and upscaling of the Dinaledi programme. It paved the way for the continuation of the programme.
Rating:	5: The evaluation was completed within the planned timeframes and budget with much more value and insight achieved as a result of exceptional project management
Moderation:	Accepted
Approval:	Accepted

Evaluation use

Standard:	Results of the evaluation have been presented to relevant stakeholders
Comment and Analysis:	Interviewees indicated that, as a result of this evaluation, the DOE was able to motivate for more funding for the Dinaledi programme from Treasury. The conditional grant has been used to further fund schools in order to improve outputs in maths and science amongst learners. Whether this impact has been sustained still requires further investigation over the long term.
Rating:	4: Results of the evaluation have been presented to all relevant stakeholders, inside and outside of government
Moderation:	Accepted
Approval:	Accepted

Standard:	A reflective process has been undertaken by the steering committee with the service provider (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:	Due to the nature of the work conducted with government at national level, there was no skills development undertaken amongst evaluators.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted

Standard:	A reflective process has been undertaken by the steering committee (if no steering committee exists)
Comment and Analysis:	Interviewees indicated that the evaluation team reflected on the Impact Evaluation methodology used. The key learnings were that: an impact evaluation cannot tell you much about how the project is being implemented and that it cannot explain its findings. In addition the team learnt that impact evaluations on their own are not enough - that programmes need to have process or implementation evaluations in order to understand how the programme is being implemented so that corrective action can be taken.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted
Standard:	The evaluation study is seen by interviewed stakeholders as having added significant symbolic value
Comment and Analysis:	The DOE interviewees indicated that, because the evaluation was conducted by the World Bank, this raised the profile of the programme and improved credibility of findings. The DOE's Action Plan 2014 Towards the Realisation of Schooling 2025 places emphasis on the value of the study which is reflected in the following quote: "The 2009 impact evaluation of the Dinaledi programme, completed in 2010 by the World Bank on behalf of the Department of Basic Education, is arguably the most thorough evaluation of its kind ever undertaken in relation to an education intervention programme in South Africa".
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted
Standard:	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 World Bank has an extensive internal peer review system where regular meetings are held with peer reviewers prior to decision making for each step of an evaluation. Although no primary data was collected for this study, the overall design and methodology was subject to peer review.
Rating:	4: The evaluation study is seen by interviewed stakeholders as being of substantial symbolic value to the policy or programme and has noticeably raised its profile amongst stakeholders
Moderation:	Accepted
Approval:	Accepted
Standard:	The evaluation study is of conceptual value in understanding what has happened and possibly in shaping future policy and practice
Comment and Analysis:	The evaluation made use of secondary administrative data only. No primary data collection was undertaken by the evaluation team.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted

Standard:	Development of a draft improvement plan has been started, but not completed, based on the findings a
Comment and Analysis:	Interviewees noted that the study provided little information about the implementation of the programme and, thus the findings were not deemed to be useful for the development of an improvement plan. However, they were used during the development of the plan for the next phase of the programme. The action plan of the DOE's sector plan (until 2025).
Rating:	: N/A
Moderation:	Reconsider
Moderation Comment:	Consider revising this to Not Applicable. A draft improvement plan is a new requirement following on evaluations that has only come into existence of late and applied to National Evaluation Plan evaluations, and this is not one of those.
Approval:	Accepted

Standard:	There is clear evidence of instrumental use - that the recommendations of the evaluation were implem
Comment and Analysis:	The study has been used by the DOE to improve the design of the Dinaledi programme. Evidence of this is contained in the DOE's Action Plan 2014 Towards the Realisation of Schooling 2025 which states the following: "A recent evaluation of the programme by the World Bank revealed that, despite certain design flaws, Dinaledi had been successful in increasing the mathematics and physical science pass rate. Yet, the overall improvements in the Grade 12 mathematics and science results, in particular with regard to historically disadvantaged learners, were disappointingly low. To confront this challenge, the national department is currently improving the design of the Dinaledi programme and ensuring that lessons from this programme are cascaded to non-Dinaledi schools".
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

References

Blum J, Krishnan N, Legovini A, 2010, "Expanding Opportunities for South African Youth through Math & Science: The Impact of the Dinaledi Program", World Bank

"Case Study: Impact of the Dinaledi Programme", Presentation to the M&E Learning Network on Evaluations in Government, 19 July 2011 by Hersheela Narsee, Department of Basic Education, South Africa

" DIME Brief: Expanding Opportunities for South African Youth through Math & Science: The Impact of the Dinaledi Program", Development Impact Evaluation Initiative, World Bank

"Action Plan to 2014: Towards the Realisation of Schooling 2025", October 2011, Department of Basic Education, South Africa

List of Interviewees

Dr Hersheela Narsee, Director for Monitoring and Evaluation, Research and Policy Unit, Department of Education, self-completed questionnaire via email, 13/10/2013

Edward Mosuwe, Chief Director for FET schools, Department of Education, telephonic interview, 8/11/2013

Ariana Legovini and Nandini Krishan: project managers, Development Impact Evaluation, World Bank, joint telephonic interview, 30/10/2013