



planning, monitoring & evaluation

Department:
Planning, Monitoring and Evaluation
REPUBLIC OF SOUTH AFRICA

Report on the Assessment of Government Evaluations

Evaluation Title:

Implementation Evaluation of the National School
Nutrition Programme

Evaluation Number:

528

Start Date:

10 December 2014

Completion Date:

16 September 2016

Period of Evaluation:

21 months

Submitted:

10 March 2017 by Cathy Chames

Approved:

10 March 2017 by Mike Leslie

Evaluation Details

Evaluation Title:	Implementation Evaluation of the National School Nutrition Programme
Evaluation Number:	528
Evaluation Completion Date:	16 September 2016
Created:	18 January 2017 by Mike Leslie
Submitted:	10 March 2017 by Cathy Chames
Approved:	10 March 2017 by Mike Leslie
Period of Evaluation:	21 months
Known Cost:	R 3 749 804.00
Known Cost Type:	Referenced
Initiated By:	Department of Planning, Monitoring and Evaluation; Department of Basic Education
Initiated By Internal:	Yes
Undertaken By:	JET Education Services
Undertaken By Internal:	No

Assessors

Cathy Chames	cathy@southernhemisphere.co.za
--------------	--------------------------------

Assessment Documents

Document Name:	Document Type:	Added By:	Added On:
NSNP report final 16092016.pdf	Evaluation report	Mike Leslie	18 January 2017
NSNP 1 5 25_report final 17092016.pdf	Evaluation report	Mike Leslie	18 January 2017
NSNP Evaluation_Inception report_17 December 2014 final updated.docx	An inception report	Cathy Chames	28 February 2017
NSNP Evaluation_Proposal_17 December 2014 final updated.docx	The proposal selected to conduct the evaluation	Cathy Chames	28 February 2017
RFP 14 229 NSNP TOR.pdf	Terms of Reference (ToR) for the evaluation	Cathy Chames	28 February 2017
NSNP Consolidated Improvement Plan 15 11 16.docx	Any other relevant documentation pertaining to the evaluation process	Cathy Chames	28 February 2017
Assessment Report at Moderation.pdf	Assessment Report at Moderation	Mike Leslie	01 March 2017

Quality Assessment Summary

The overall score for this evaluation is an above average 3.54. The evaluation was well-managed, particularly during the planning and inception phase where a wide range of stakeholders were meaningfully involved in the scoping of the TOR and choosing the purpose of the evaluation. This contributed towards a good quality TOR which also scored well particularly since it was well-structured and complete.

Key stakeholders were involved in the evaluation through a formalised Steering Committee and there was a good relationship between the service provider, the Steering Committee and the technical working group during the implementation phase. Another strength of during this phase was that the evaluation team was able to work freely without interference and was given access to all sought data and information sources and thus it was a free and open evaluation process. A weakness, however, was that there was limited capacity building of partners responsible for the evaluand.

In terms of methodological integrity, the alignment to the policy context and background literature was a particular strength and the evaluation was underpinned by an excellent, succinct literature review covering both national and international literature. The methods were implemented adequately, the instruments were piloted before roll-out and data was collected with most of the planned sample including beneficiaries who were meaningfully consulted as a primary source of data.

The data analysis was well executed and the final report scored well. The accessible content, robust findings and conclusions backed up by evidence are all particular strengths. However, it fell short of documenting and acknowledging ethical considerations even though evaluation ethics were adhered to throughout the study.

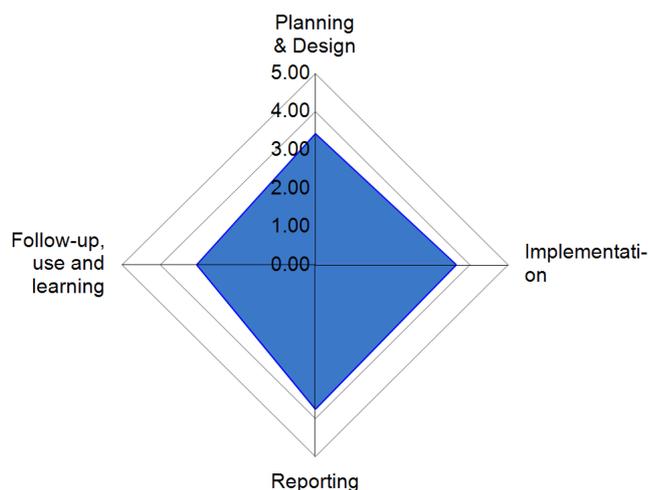
Whilst the evaluation was completed within the planned budget the time frames were not adhered to and the multiple delays resulted in the entire process taking 21 months to complete. A strength, however, was the partnership approach followed throughout which culminated in the results of the evaluation being presented to key government stakeholders who actively engaged in shaping the recommendations. Overall the evaluation is of conceptual value as it has deepened stakeholders' understanding of the NSNP, its implementation and proposed recommendations for improving the programme. The DBE has agreed to 80% of the recommendations which have been captured in an improvement plan and there appears to be good buy-in from the Provincial Departments of Education to implement the recommendations.

Quality Assessment Scores

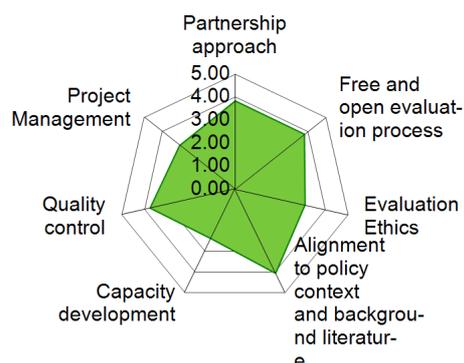
Phase of Evaluation	Score
Planning & Design	3.43
Implementation	3.65
Reporting	3.77
Follow-up, use and learning	3.07
Total	3.54

Overarching Consideration	Score
Partnership approach	3.84
Free and open evaluation process	3.83
Evaluation Ethics	3.10
Alignment to policy context and background literature	4.07
Capacity development	2.40
Quality control	3.75
Project Management	3.05
Total	3.54

Scores: Phases of Evaluation



Scores: Overarching Considerations



Phase of Evaluation	Area of Evaluation	Score
Planning & Design	Quality of the TOR	3.75
Planning & Design	Adequacy of resourcing	3.14
Planning & Design	Appropriateness of the evaluation design and methodology	3.00
Planning & Design	Project management (Planning phase)	4.00
Implementation	Evaluation ethics and independence	3.50
Implementation	Participation and M&E skills development	3.14
Implementation	Methodological integrity	3.95
Implementation	Project management (Implementation phase)	3.50
Reporting	Completeness of the evaluation report	3.50
Reporting	Accessibility of content	4.00
Reporting	Robustness of findings	4.00
Reporting	Strength of conclusions	4.00
Reporting	Suitability of recommendations	3.50
Reporting	Acknowledgement of ethical considerations	3.14
Follow-up, use and learning	Resource utilisation	2.00
Follow-up, use and learning	Evaluation use	3.33
Total	Total	3.54

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: The TOR is comprehensive, well-structured and clear. It contains the following elements: Background information and rationale; purpose of the evaluation; key evaluation questions; potential users of the evaluation and how they will use the results; scope of the evaluation; evaluation plan; methodology and evaluation approach; milestones; competencies and skills-set; service provider details; management arrangements; structure and contents of proposal; additional information required; and additional information for service providers. Whilst two out of the three interviewees felt the TOR was clear, one interviewee felt the TOR could have been more tightly defined.

Rating: 4: The evaluation was guided by a well-structured and complete TOR or internal evaluation proposal of a good standard

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: This is an implementation evaluation and the purpose as stated in the TOR is "to assess whether the National School Nutrition Programme (NSNP) is being implemented in a way that is likely to result in significant health and education benefits to primary school learners". The scope of the evaluation is immediately qualified here by stating that even though the programme is implemented in both primary and secondary schools, due to time and resource limitations, the evaluation will focus on primary schools only.

All eight of the questions are appropriate for an implementation evaluation and question 6 states: "is there evidence that the NSNP enhances learning behaviour? (likely impact of the programme)". This suggests elements of an impact evaluation even though the DPME Guideline on Implementation Evaluation states that, whilst an implementation evaluation does not include an evaluation of outcomes, it can be used to assess likely achievement of outcomes.

It was noted by interviewees that, even though the TOR stated that the evaluation scope was focused on 'likely impact' there was an expectation from the NSNP programme staff to answer questions around the impact of the NSNP. It was also indicated that the scope of the evaluation was too broad and that there could have been more value in being more 'focused' with fewer evaluation questions to obtain a more in depth understanding of certain aspects of the programme.

One interviewee also noted that, although not stated clearly in the TOR, there were elements of a design evaluation because there was a need to revisit the TOC and assess if it was still 'relevant' for the NSNP.

The section in the TOR which outlines the evaluation scope includes clear details on the time period under review, themes covered/not covered and geographic coverage (all 9 provinces).

Rating: 3: The approach and type of evaluation was suited to the purpose and scope of the evaluation TOR (or an internal evaluation proposal)

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

Comment and Analysis: The TOR clearly maps out the intended users of the evaluation (DPME, DBE, Provincial Departments of Education, DOH, National Treasury, DSD and civil society organisations) and how they will use it (their information needs).

Rating: 4: The TOR identified the intended users of the evaluation and differentiated between their information needs well

Standard:	Key stakeholders were involved in the scoping of the TOR and choosing the purpose of the evaluation
Comment and Analysis:	The request for this evaluation came from Cabinet because of the considerable amount of funding which has been allocated to the programme since its inception. Following this Cabinet instruction, the DBE worked in partnership with DPME in designing the evaluation purpose and TOR. Other stakeholders involved in developing the TOR included the DOH, Treasury and an NGO working in the sector - FUEL.
Rating:	4: A wider range of stakeholders (i.e. beyond government stakeholders) were meaningfully involved in the scoping of the TOR and choosing the purpose of the evaluation

Adequacy of resourcing

Standard:	The evaluation was adequately resourced in terms of time and budget allocated
Comment and Analysis:	Although Cabinet initially requested an impact evaluation, this could not have been achieved given the limited budget allocation for a study of this scale. Following an initial scoping study to find out if an impact evaluation was feasible, the focus changed to an implementation evaluation however, interviewees were still unanimous that the budget allocation was inadequate given its very broad scope and national scale. This resulted in certain aspects being 'cut back'. For example, the evaluation could only focus on primary schools even though the programme is run in both primary and secondary schools. Furthermore, a detailed study on the nutritional content of meals was deemed to be too expensive even though this was a core component expected to be covered. The way in which the service provider approached this was to build a tool developed by FUEL for assessing food groups, quantities and types of food, and the standard of food. In terms of timeframes, the TOR allocated 8 months for evaluation completion but this was felt to be unrealistic given its scope. More realistic timeframes were renegotiated with the service provider once they were contracted to ensure that quality of work was not compromised.
Rating:	2: The evaluation was resourced with tight timeframes and budget which were challenging from the outset
Standard:	The team conducting the evaluation was adequately resourced in terms of staffing and skills sets
Comment and Analysis:	The evaluation team from JET Education Services has a long standing history in research and education and overall the team was adequately resourced in terms of staffing and had a strong skills set. Although JET did not have experience as an organisation in conducting an evaluation of this nature, the individuals on the team each have considerable experience in the education sector. The revised proposal specifies that the team consisted of a Project Director, Project Manager, two Technical Advisors; a nutritionist; two M&E Specialists; and a big team of qualitative and quantitative researchers. However, in contrast to the team specifications in the proposal, some interviewees pointed out that there was no nutritionist on the evaluation team. This was highlighted as a gap in the skill set particularly since the focus of the NSNP is on nutrition and its contribution towards the cognitive development of children. One of the reasons given for this gap was that the delays in contracting may have affected the availability of all the proposed team members.
Rating:	4: The evaluation was well resourced in terms of staffing and skills sets

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:	There is explicit reference to the intervention logic of the NSNP in the TOR and the theory of change (TOC) is referred to briefly in the revised proposal. However, even though considerable time was spent on clarifying the TOC during the implementation phase, there is no evidence that the TOC was utilised to inform the planning and costing of the evaluation at the proposal and inception phase.
Rating:	3: There was explicit reference to the intervention logic or the theory of change of the evaluand in the TOR or the Inception Report

Standard:	The planned methodology was appropriate to the questions being asked
Comment and Analysis:	In preparation for this study an initial scoping study was undertaken by DBE to understand whether an impact evaluation was feasible for the NSNP. The study revealed that it would not be possible to conduct an impact evaluation of the NSNP given the difficulty in identifying a comparison group of children who did not receive the programme and yet are similar to programme beneficiaries and the absence of baseline data on indicators such as academic or nutrition outcomes. An implementation evaluation was thus decided upon and the six core questions centred largely on the NSNP implementation. Overall the proposed methodology was thought to be well suited to these questions. It planned to adopt a mixed method approach combining a literature review of past evaluations and international best practices for school nutrition programmes; a document review of all the various guidelines and programme documents; interviews with key stakeholders at various levels of the system (national, provincial, district, community); a quantitative survey with school stakeholders, food handlers, NSNP coordinators and food management teams, and learners to ascertain what they think of meals and services. An observation instrument was also used to assess food preparation, serving, feeding and other aspects such as availability of food, infrastructure and utensils. One aspect not included in the original TOR was that of a cost analysis. To undertake a true cost effectiveness analysis requires outcome data which was not available, it was therefore agreed that a basic financial analysis would be included in the methodology, i.e. analysis of cost and output data where data was available.
Rating:	4: The planned methodology was well suited to the questions being asked and considered the data available

Standard:	The sampling planned was appropriate and adequate given the focus and purpose of evaluation
Comment and Analysis:	Overall the sampling planned was not entirely appropriate given the purpose and scope of the evaluation. The NSNP is a large scale national programme and the TOR specified that a 'nationally representative sample of all public primary schools in quintiles 1-3, including special schools' should be included in the evaluation. However, the final selection of 30 schools per province (270 schools in total) was not nationally representative - the revised proposal states that a minimum of 107 schools per province would have been the ideal number. The limited number of schools surveyed due to budgetary constraints means that the confidence intervals are wider than would have been preferred, especially when dis-aggregating per province. This meant that the findings per province could not be generalised. One interviewee confirmed that the small sample at provincial level made it difficult to draw out the provincial differences which was an expectation from some members of the evaluation steering committee.
Rating:	2: The sampling planned was not entirely appropriate given the focus and purpose of the evaluation

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:	The inception phase was used to develop a common agreement on how the evaluation would be implemented. Following an inception meeting JET presented a revised proposal to DPME with a reduction in budget and scope of work. This was reviewed by two reviewers appointed by the DPME to review the proposal and discussed at a Steering Committee meeting before the revised proposal was agreed to by all parties. The comments of the reviewers and Steering Committee were taken into account wherever feasible. The inception report specifies the key changes in the final proposal. These include the following changes: changes to the scope of the literature review and of secondary data; removal of questions about job creation in relation to the NSNP; removal of sampling as an activity; details on the piloting processes; removal of majority of qualitative fieldwork in 10% schools and changes to the targeted stakeholders for the qualitative sample; and inclusion of a cost analysis.
Rating:	4: The inception phase was used to good effect to achieve a common agreement and understanding of how the evaluation would be implemented

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:	<p>The inception report provided specific details on how the evaluation team planned to adhere to research ethics and confidentiality stating that all team members engaging with stakeholders through interviews, focus groups or surveys will adhere to strict legal and ethical guidelines without hampering their creativity and delivery to the clients. The interview with the service provider confirmed that fieldworkers were trained in this regard and that they followed the principle of informed consent.</p> <p>The proposed evaluation was taken through the national DBE's internal research approval process which was mirrored in each province.</p> <p>An area of concern is that, even though the evaluation included interviews with school children who are minors (under the age of 18 years), a formalised ethical clearance process was not adhered to. It was explained that, in order to deal with parental consent, the school principals were deemed to be 'locus parentis' and were thus able to give consent for children to participate in the survey.</p>
Rating:	3: There was clear evidence that ethical protocols were observed for some data collection instances including: informed consent agreements; confidentiality; documenting and storing data notes, recordings or transcripts; and ethics review board approvals where appropriate
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:	<p>There is no evidence to suggest that there was any significant interference to the work of the evaluation team and they were given access to existing data. It was mentioned by one interviewee that the service provider was required to report to DPME and not the custodian department. This mechanism is put in place to ensure there is no 'intimidation' by the custodian department. The use of peer reviewers throughout the process are also important to ensure impartiality and credibility.</p> <p>The DBE developed the sampling frame and selected schools which could have the potential for interference from the department, however, the schools were randomly selected and JET was fully involved in decisions about the variables for sampling and the approach. The process was also overseen by the Steering Committee.</p> <p>Another area with potential for interference was the planned visits to schools. Concern was raised that the accuracy of information being collected would be affected if schools knew the fieldwork team would be visiting on a particular day (i.e. not a 'typical day' of feeding but a day 'staged' by the school in anticipation of the fieldwork visit). The matter was discussed in detail at the Steering Committee whilst taking into consideration that unannounced visits may have resulted in challenges such as researchers being turned away or respondents being unavailable which would impact on the completeness of data. In the end it was agreed that: all schools would be notified about the evaluation; and JET would contact the sampled schools to inform them that they would be visited 'at some point' within the fieldwork period but not the specific day.</p>
Rating:	4: The evaluation team was able to work freely without interference and was given access to all sought data and information sources

Participation and M&E skills development

Standard: Key stakeholders were involved in the evaluation through a formalised mechanism or institutional arrangement

Comment and Analysis: A multi-sectoral Steering Committee was established for the evaluation and there was a TOR in place for this structure which outlined the roles and responsibilities of each member. The committee was made up of DPME, DBE, National Treasury, representatives from two provinces (Eastern Cape and Gauteng) and an NGO with experience in the education/nutrition sector (FUEL). The representatives from DBE included those from the directorates of monitoring and evaluation, school nutrition and a researcher from the DG's office. The DOH was also involved in the early stages of the evaluation (development of the TOR). The programme manager from DBE chaired the committee to ensure ownership and the committee was actively involved in all phases of the evaluation including the development of the TOR, inputs into the methodology, comments on all reports, and development of an implementation plan. All provinces were invited to provide comments on the draft report and they also took part in the development of the implementation plan.

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)

Standard: Where appropriate, an element of capacity building of partners responsible for the evaluation and evaluators was incorporated into the evaluation process

Comment and Analysis: The revised evaluation proposal indicates that two junior researchers from DPME and DBE had been identified and would form part of the fieldwork team for evaluation. It was planned that their involvement of the team would be addressed in an agreed upon Skills Transfer Plan and that there would be several opportunities during the various research phases. However, interviewees revealed that these two interns were not involved in the capacity building. On the other hand, JET had interns and junior researchers working on the evaluation who were given the opportunity and exposure to work on a high level evaluation. They were involved in many different aspects including the design of instruments, data collection, analysis and the improvement plan workshop.

Rating: 2: There was some evidence of capacity building of partners responsible for the evaluation or evaluators but this was either unstructured or incomplete

Methodological integrity

Standard: A literature review was developed which informed the analytical framework and findings of the evaluation

Comment and Analysis: An excellent, succinct literature review covering both national and international literature was conducted. The review covers national and international research regarding the health and nutritional status of school-age children in South Africa and the outcomes of school feeding programmes. Of note here is that research from developing countries is reviewed which is relevant to the South African context including India, Brazil and Chile; and also from other African countries including Kenya, Malawi and Uganda. The review also covers aspects related to implementation of school nutrition programmes including the cost of scaling up such programmes and the findings from previous evaluations of the NSNP are also presented and discussed. It concludes by extracting key contextual factors and characteristics that typically determine the effectiveness of school nutrition programmes. These lessons learnt have been used to inform the conceptual framework, analytical framework, the development of the theory of change, and to triangulate the findings of the evaluation.

Rating: 5: An excellent literature review was developed covering international and national literature, a diversity of view points, which informed the analytical framework and interpretation of issues relevant to the findings

Standard: The methods employed in the process of the evaluation were consistent with those planned and implemented adequately

Comment and Analysis: The evaluation methods reported on in the final report are consistent with those planned in the original revised evaluation proposal and inception report. Some challenges were encountered with entry into a few of the schools resulting in a difference in the number of planned (270) versus actual targeted schools (267) but this did not impact on the validity of findings. All of the planned qualitative interviews at national, provincial and district level took place.

Interviewees highlighted some challenges with data collection. Firstly, the financial information that was provided to the evaluation team was found to be limited in terms of its usefulness and this impacted on the validity of findings of the cost analysis. Secondly, the data on actual quantities of food being served was difficult to assess because schools don't have weighing scales and thus estimates of quantities had to be made. Furthermore, the nutritional content of meals also proved difficult to assess. This was done by examining the extent to which schools prepared and served the correct number of food groups (quality) and the correct amount of starch, protein, vegetables/fruit for the number of learners approved for the NSNP (quantity). However, the number of learners approved for the NSNP is based on enrollment in the previous school year and it became apparent during fieldwork that some schools were feeding fewer and some were feeding more than the approved learners.

Rating: 3: The methods employed in the process of the evaluation were consistent with those planned and implemented adequately

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: The qualitative interview and survey instruments were piloted prior to being finalised to check that the fieldwork could be done in the timeframe available and to make sure the instruments would generate useful information. After this changes were made to the instruments. Piloting took place in six schools in Limpopo (rural) and Gauteng (urban) to provide feedback on administration in a variety of contexts.

One challenge highlighted for the piloting phase was that the instruments were discussed with the Steering Committee after piloting instead of beforehand. The result is that additional questions were added to the instruments which were then changed further and could not be piloted again.

Rating: 3: A pilot of basic data collection instrumentation occurred prior to undertaking data collection and it was used to inform the research process

Standard: Data was collected from key stakeholders (e.g. implementers, governance structures, indirectly affected stakeholders) as data sources

Comment and Analysis: A broad range of stakeholders were included in the sample from all levels of the system at national, provincial, district and school level. Key stakeholders at national, provincial and district level were interviewed and then at each school interviewees included: principals, NSNP coordinators, volunteer food handlers, school governing bodies, learners and service providers. In terms of planned versus actual data collection, a total of 267 out of the planned 270 sampled schools were reached (98% of target); for the learner surveys, a total of 4999 out of 5400 learner instruments were completed (93%); for the other instruments (food handler, SMT, NSNP co-ordinator, SGB and observation instruments), a total of 1309 out of 1350 were completed (97%).

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)

Standard:	The methodology included engaging beneficiaries appropriately as a key source of data and information
Comment and Analysis:	Grade 6 and grade 7 learners who are beneficiaries of the NSNP were engaged as primary sources of data and information in the evaluation. They were asked about key aspects of programme including their perceptions of the quality of the meals and services and its effect. As mentioned previously, secondary school learners who are also intended beneficiaries of the programme were not included in this evaluation but this was due to budgetary constraints.
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)

Project management (Implementation phase)

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:	Overall the interviewees confirmed that there was a good working relationship between the service provider, the Steering Committee and the technical working group. It was noted that the Steering Committee is a 'robust structure which works well'. However, it was noted that sometimes there is almost too much involvement in some aspects of the evaluation such as the instrument design which resulted in the learner instrument, in particular, being too long.
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
Standard:	Support provided by the evaluation secretariat (e.g. the administrators responsible for the evaluation) facilitated achievement of the objectives of the evaluation (eg turnaround times, addressing problems, preparation for meetings etc)
Comment and Analysis:	DBE provided access to database of schools, assisted with sampling and with gaining access into schools. The role of the DPME was much appreciated as they fulfilled their role of facilitator well and 'did a good job in terms of bringing together the many stakeholders'. One major challenge experienced during the feedback and report finalisation phase is that DBE did not stick to their deadlines in terms of providing comments to the report and this delayed the finalisation of the report. It was noted that DPME should have played a stronger role in terms of enforcing these deadlines.
Rating:	3: Support provided by the evaluation secretariat facilitated achievement of the objectives of the evaluation

Reporting

Completeness of the evaluation report

Standard: The first draft evaluation report was of a sufficient quality to go to stakeholders and did not require major changes

Comment and Analysis: The first draft evaluation report produced by the service provider was shared with all stakeholders prior to finalisation. Although there were some weaknesses with the report this did not require any major changes prior to sharing with stakeholders. Following the submission of the first draft there was a very long period of feedback and comments and four drafts were submitted before the final report was submitted incorporating all comments and inputs. Some gaps highlighted by interviewees were: some of the original evaluation questions were not answered; limited interpretation of the findings and their significance for programme improvement; recommendations not being adequately linked to findings. However, the final report that was produced managed to address most of these issues.

Rating: 3: A first draft of the evaluation report was of a sufficient quality to go to stakeholders and did not require major changes prior to sharing

Standard: The final evaluation report is well-structured and complete in terms of the following: executive summary; context of the development evaluation; evaluation purpose, questions and scope; methodology; findings and analysis; conclusions and recommendations

Comment and Analysis: The final report is well-structured and comprehensive. It contains an executive summary; introduction and background to the NSNP, the evaluation and the methodology and approach; the theory of change and logframe; literature review; presentation of evaluation findings and analysis; conclusion and recommendations/

Rating: 4: The final evaluation report is well-structured, complete and presents the following report components well: executive summary; context of the development evaluation; evaluation purpose, questions and scope; methodology; findings and analysis; conclusions and recommendations

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: The final report is user friendly and is written in accessible language. This was confirmed by interviewees who said that it was 'easy to read'. There are no grammatical errors and the style is consistent throughout and has an appropriate level of formality. References are complete and all acronyms are listed. One weakness is that the final report is very long - 189 pages long (excluding appendices). It was mentioned by one of the interviewees that it contained a 'huge amount of information' making it difficult to organise and present in a succinct way.

Rating: 4: The final report is well written, accessible to the common reader and ready for publication with only minor spelling, grammar or formatting mistakes

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:	Figures and tables are used extensively throughout the report to visually present the quantitative survey findings. The tables are easy to read, clearly labelled and the source of data is provided for each figure and table (i.e. learner survey, VFH survey, KPI instrument etc). In most instances the data is dis-aggregated per province to allow for comparison across provinces. Some of the data is also dis-aggregated according to other variables where relevant, for example when comparing centralised and decentralised service delivery models.
Rating:	4: Figures, tables and conventions are well used for a variety of types of data presentations and supporting explanations make them accessible to readers

Robustness of findings

Standard:	Data analysis appears to have been executed to an adequate standard
Comment and Analysis:	<p>Overall the data analysis appears to have been well executed for all datasets. Qualitative data was coded using qualitative codebook. The quantitative survey data was captured, cleaned and analysed in R language version 3.2.1 and Stata version 14.0 and tables generated to present the overall disaggregated results.</p> <p>For the cost analysis the main sources of data used were NSNP Quarterly Reports covering the period 2009/10 - 2013/14 and Business Plans for the 2012/13 financial year, which were provided by the DBE. A subset of indicators was identified within the Quarterly Reports for which comparable data were available across provinces and over time.</p>
Rating:	4: Data analysis appears to have been well executed for all datasets
Standard:	Findings are supported by evidence which is sufficiently and appropriately analysed to support the argument, integrating sources of data
Comment and Analysis:	<p>An evaluation matrix was developed as an analytical framework for the evaluation and the evaluation questions and sub-questions were related to this framework using the 6 DAC criteria of relevance, efficiency, effectiveness, impact, sustainability. Although not one of the DAC evaluation criteria, "additionality" was added, as one of the key evaluation questions which relates to investigating spinoffs of the NSNP. and additionality. This analytical framework provided an overarching conceptual framework for structuring the data analysis work and report writing. Following an initial round of data analysis an implementation index which summarises performance in the key aspects of implementation measured via the school surveys was created to guide analysis of successful implementation of the NSNP across schools, provinces and per implementation model (centralised/decentralised). Additionally, statistical tests were conducted to check the relationships between categorical variables.</p> <p>Overall the evidence gathered is well-analysed particularly in relation to the variations in implementation between provinces and between the centralised and decentralised models of service delivery. All findings are backed up by the qualitative and quantitative evidence including fieldworker observation notes. The findings from the literature review and other evaluations of the NSNP are woven into the findings to further strengthen the key arguments made.</p>
Rating:	4: The evidence gathered is well analysed, integrated and supports the argument in key sections of the report, without presenting data which are not used in the argument

Standard:	There is appropriate recognition and exploration of the possibility of alternative interpretations
Comment and Analysis:	Care is made to present the findings in a tentative manner and alternative interpretations are presented where necessary. For example, where the survey data reveals that learners do not like soya, it is argued that their not liking soya may be related to the high proportion of schools (41.7%) serving soya which does not meet the NSNP soya mince specification. In another example, when comparing the centralised and decentralised models, it is stated that there was more variation between different provinces using the same model than between models, indicating that province specific factors account for the greatest part of the differences.
Rating:	4: There is clear recognition of the possibility of alternative interpretations and these are concisely presented without detracting from other findings

Standard:	The report appears free of significant methodological and analytic flaws
Comment and Analysis:	The report appears free of significant methodological analytic flaws. The report clearly describes the methods used to collect and analyse data and provides extensive detail on the sampling procedures used to arrive at a 'nationally representative sample of schools' via a representative sample through probability sampling of 270 primary schools from quintiles 1 to 3. The report also provides extensive detail on the weighting process during the sampling - two weighting variables were required for the purposes of making estimations of population characteristics - learner weight (for calculations such as X% of learners are fed) and school weight (for calculations such as Y% of schools served a nutritious meal). These were adjusted after data collection.
Rating:	4: The report documents some of the methodological and analytical processes used to ensure that it is free of methodological and analytic flaws

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:	<p>The limitations of the methodology are clearly articulated in the report. With regards to the scope and design the report reiterates that this is not an impact evaluation, economic evaluation or an audit of the NSNP and explains that the the impact evaluation originally planned for this evaluation was discarded based on the findings from the scoping study.</p> <p>There are also limitations related to the investigation of the nutritional content of meals; the limited number of schools surveyed and that care should be taken when interpreting results; the risk of recall bias of stakeholders recounting their experiences of NSNP business processes; and the implementation index which is tentative and requires further refinement and validation.</p>
Rating:	4: Limitations of all aspects of the methodology and findings are clearly articulated and distinguish between different kinds of limitations

Strength of conclusions

Standard:	Conclusions are derived from evidence
Comment and Analysis:	All of the conclusions made are backed up by evidence which have been well-analysed and triangulated with multiple sources of data including the primary quantitative and qualitative data (including observations); the findings from the literature review; and findings from previous evaluation studies of the NSNP.
Rating:	4: Conclusions are derived from evidence and well supported by multiple sources of data that has been well analysed

Standard:	Conclusions address the original evaluation purpose and questions
Comment and Analysis:	The overall purpose of the evaluation is to assess whether the National School Nutrition Programme (NSNP) is being implemented in a way that is likely to result in significant health and education benefits to primary school learners. Whilst the concluding chapter of the report does not revisit this overarching question explicitly it does answer it implicitly by addressing each of the original evaluation questions in detail.
Rating:	4: Conclusions address the original evaluation purpose and questions well
Standard:	Conclusions are drawn with explicit reference to the intervention logic or theory of change
Comment and Analysis:	There are multiple explicit references to the theory of change and intervention logic of the NSNP in the concluding chapter. In particular, when answering question 3 - Is the programme implemented as planned? - the conclusion revisits the four outcomes and pathways of change for the TOC and explores whether they are being upheld by the findings on programme implementation, highlighting the barriers to the potential achievement of outcomes. When answering question 7 - is there evidence that the NSNP enhances learning behaviour (likely impact)? the TOC is revisited again and it is concluded that if the programme is implemented as planned, and the change theory is plausible, impact is more likely. It then goes on to summarise the key challenges which need to be addressed to increase the likelihood that the NSNP will enhance learning behaviour and improve retention. With regards to long term nutritional and health benefits it is concluded that these are difficult to ascertain because there are many other factors outside of the school meals and school which influence child health and nutrition.
Rating:	4: Conclusions are drawn with an explicit reference to, and provide a clear judgement on, the intervention logic or theory of change

Suitability of recommendations

Standard:	Recommendations are made in consultation with relevant government officials, stakeholders and sectoral experts
Comment and Analysis:	Recommendations were made in consultation with the Steering Committee including DBE. A workshop was held to present the findings of the first draft report and the preliminary recommendations were discussed and refined. Two peer reviewers reviewed the report with one of them being a nutrition advisor who gave input into the recommendations.
Rating:	3: Recommendations are made in consultation with relevant government officials, stakeholders and sectoral experts
Standard:	Recommendations are useful- they are relevant, specific, feasible, affordable and acceptable
Comment and Analysis:	The Interviewees reported that initially there were too many recommendations (41) with some described as not being 'implementable', thus the consultation process was critical in ensuring that the key recommendations be extracted, refined and accepted by all involved. The final report thus presents a total of 18 recommendations which focus on improvement of the NSNP in relation to the programme relevance and appropriateness; programme effectiveness; fidelity and efficiency; and funding, sustainability, upscaling and impact. Overall they appear to be specific and relevant to the evaluation findings. Interviewees also reported that a lot of the recommendations were already being implemented by the DBE by the time the report had been finalised. This is evidence that they are regarded as feasible, affordable and acceptable.
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

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 full report does not document any of the procedures put in place to ensure confidentiality and informed consent even though there is evidence from the interviews and the revised proposal that these procedures were followed.

Rating: 2: The full report does not acknowledge whether confidentiality was ensured or informed consent secured but there is some evidence that this is the case

Standard: There are no risks to participants or institutions in disseminating the evaluation report on a public website

Comment and Analysis: The anonymity of all evaluation participants has been protected and there is no evidence to suggest that there would be a risk to participants or institutions through dissemination of the report on a public website.

Rating: 4: There are no risks to participants or institutions in disseminating the original full evaluation report on a public website

Follow-up, use and learning

Resource utilisation

Standard:	The evaluation was completed within the planned timeframes and budget
Comment and Analysis:	The evaluation was completed within the planned budget but not within the planned time frames. The time frames and deliverable schedule in the inception report states that the first draft report would be submitted by end of May 2015 and approved by 30 June 2015. However, the first draft was submitted three months later, on 21 August 2015 and the final report was only approved by the evaluation steering committee one year later on 16 September 2016. Reasons for this were the following: delays in appointment of the service provider; inception phase took much longer than expected due to lengthy negotiations around budget and scope; delays in fieldwork due to Easter holiday; delays in commenting and input from DBE into the draft report.
Rating:	2: The evaluation was completed outside of the planned timeframes and over budget, but with approval of the commissioning organisation

Evaluation use

Standard:	Results of the evaluation have been presented to relevant stakeholders
Comment and Analysis:	The results of the evaluation were shared with the Steering Committee and provinces were also invited to give inputs which were incorporated into the findings and recommendations. The report has been widely distributed to staff in the DBE and has been discussed at inter-provincial meetings.
Rating:	3: Results of the evaluation have been presented to relevant stakeholders in government

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:	A close-out meeting was held with the service provider and some members of the Steering committee, including the DBE to reflect on the evaluation process. Overall it was felt that all parties had learnt a lot from the process, particularly the DBE staff members who had not been involved in an evaluation such as this and indicated that they had learnt a lot, particularly though the technical support provided by the DPME. It is unclear whether this reflective process looked at ways in which evaluations could be strengthened in the future.
Rating:	3: A reflective process has been undertaken by the steering committee with the service provider to reflect on what could be done to strengthen future evaluations

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:	There were mixed responses with regards to the perceived symbolic value of the evaluation study. Some respondents felt disappointed with the findings stating that there was insufficient analysis of the nutritional content of food and insufficient evidence on likely impact of the NSNP. It was also noted that the evaluation was not of real symbolic value to the NSNP particularly since most of the recommendations made were already being addressed by the DBE. On the other hand, it was felt that, with the development of an improvement plan, there has been renewed support from senior management for implementing changes and allocating funding for these improvements.
Rating:	3: The evaluation study is seen by interviewed stakeholders as being of symbolic value to the policy or programme

Standard:	The evaluation study is of conceptual value in understanding what has happened and possibly in shaping future policy and practice
Comment and Analysis:	This evaluation is of conceptual value as it has deepened stakeholders' understanding of the NSNP, its implementation and recommendations for improving the programme. These insights have informed the development of an improvement plan for the NSNP by the DBE in November 2016. The DBE has agreed with 80% of the recommendations and indicate that some are already being implemented. There also appears to be good buy in from the Provincial Departments of Education to implement the recommendations. This is evidence that the evaluation study has strong potential to shape and influence implementation of the NSNP in the near future.
Rating:	4: The evaluation study is of conceptual value in understanding what has happened and some interviewed stakeholders indicated the likelihood of it constructively shaping policy and practice

References

Department of Planning, Monitoring and Evaluation and Department of Basic Education. 16 September 2016. Report on the Implementation Evaluation of the National School Nutrition Programme. Full Report.

JET Education Services. 17 December 2017. Inception Report submitted to the Department of Planning Monitoring and Evaluation in the Presidency relating to the Implementation Evaluation of the National School Nutrition Programme.

Department of Planning, Monitoring and Evaluation and Department of Basic Education. 16 September 2016. Report on the Implementation Evaluation of the National School Nutrition Programme. 1,5,25 Report.

Department of Basic Education. 15 November 2016. Improvement Plan and Reporting Template for Implementation Evaluation of the National School Nutrition Programme.

Department of Basic Education and Department of Planning, Monitoring and Evaluation. August 2014. Terms of Reference for the Implementation Evaluation of the National School Nutrition Programme.

JET Education Services. 10 December 2014. Revised Proposal to the Department of Planning Monitoring and Evaluation in Presidency to conduct an Implementation Evaluation of the National School Nutrition Programme.

List of Interviewees

Mr Jabu Mathe, Monitoring and Evaluation Advisor: DPME. Telephonic interview, 1 February 2017.

Ms Eleanor Hazel, Service Provider: JET Education Services. Telephonic interview, 3 February 2017.

Mrs M. Mashigo, Deputy Director Monitoring, Research and Knowledge Management and Mrs S Behane, Chief Education Specialist: DBE. Telephonic interview, 8 February 2017