



planning, monitoring & evaluation

Department:
Planning, Monitoring and Evaluation
REPUBLIC OF SOUTH AFRICA

Report on the Assessment of Government Evaluations

Evaluation Title:	Diagnostic Review of the State Response to Violence Against Women and Children
Evaluation Number:	527
Evaluation Completion Date:	31 March 2016
Period of Evaluation:	January 2015 - March 2016
Submitted:	28 October 2016 by Wilma Wessels-Ziervogel
Approved:	28 October 2016 by Mike Leslie

Evaluation Details

Evaluation Title:	Diagnostic Review of the State Response to Violence Against Women and Children
Evaluation Number:	527
Evaluation Completion Date:	31 March 2016
Created:	16 September 2016 by Mike Leslie
Submitted:	28 October 2016 by Wilma Wessels-Ziervogel
Approved:	28 October 2016 by Mike Leslie
Period of Evaluation:	January 2015 - March 2016
Known Cost:	R 1 986 902.12
Known Cost Type:	Referenced
Initiated By:	Department of Social Development, Department of Planning, Monitoring and Evaluation
Initiated By Internal:	Yes
Undertaken By:	KPMG
Undertaken By Internal:	No

Assessors

Wilma Wessels-Ziervogel

wilma@southernhemisphere.co.za

Assessment Documents

Document Name:	Document Type:	Added By:	Added On:
Diagnostic Review VAWC Final Full Report_16 03 14.pdf	Evaluation report	Mike Leslie	16 September 2016
Diagnostic Review VAWC Final Summary Report_16 03 14.pdf	Evaluation report	Mike Leslie	16 September 2016
DPME VAWC - Sector Review - Primary Survey Tools- Final.docx	Assessment tools	Wilma Wessels-Ziervogel	20 October 2016
DPME VAWC Phase II proposed scope change 28 07 2015 CL LB.doc	Minutes or notes of meetings with stakeholders	Wilma Wessels-Ziervogel	20 October 2016
DPME - Diagnostic review of VAWC - Inception Report.pdf	An inception report	Wilma Wessels-Ziervogel	20 October 2016
TOR-VAWC- Final 30 Jan 2014.pdf	Terms of Reference (ToR) for the evaluation	Wilma Wessels-Ziervogel	20 October 2016
22 Feb 2016 VAWC IMC Presentation_final.pptx	Presentations of evaluation findings and recommendations	Wilma Wessels-Ziervogel	20 October 2016
Assessment Report at Moderation.pdf	Assessment Report at Moderation	Mike Leslie	20 October 2016
TOR-VAWC- Final_14 10 24.doc	Any other relevant documentation pertaining to the evaluation process	Wilma Wessels-Ziervogel	28 October 2016

Quality Assessment Summary

The evaluation received an overall score of 3.04.

The scope of the diagnostic review shifted considerably. Initially, the purpose was to conduct 1) a sector review of VAWC and then 2) an evaluation of the effectiveness of particular programmes identified. Part two of the evaluation was dropped, as the scope was too broad.

The evaluation was guided by a well-constructed Terms of Reference (score of 4.25) which followed a very consultative process. This consultative or partnership approach between DPME, IMC TTT and the service provider was generally followed throughout the entire evaluation process (3.60). Although it was strategic to have the IMC TTT as the steering committee; this was also met with the practical challenges.

A mixed-method approach was used for this evaluation, including a literature review, document review (M&E records, budgets, Plans, etc.), online survey and semi-structured face to face interviews with Departmental staff. The literature and document review components were thorough and systematic. The literature review informed the evaluation framework extensively was a strength of the evaluation (3.80). The methodological integrity (overall score of 3.00) was challenged by the online survey which did not work well within the government context, and received low response rates. While the NGO sector was not included in the scope of the evaluation, they were still consulted, and played a role in shaping the evaluation process.

While the quality of the qualitative interviews may have met the necessary standards (it is difficult to determine as the transcripts were not reviewed), these were not systematically analysed or sufficiently used in the report. There are also some challenges with using multiple data sources to provide evidence for the arguments made. Data analysis in this regard was challenged in terms of its depth.

The report also does not have clear parameters in terms of what is being assessed. The findings were not sufficiently/explicitly framed within evaluation questions and there was not an obvious alignment between evaluation questions and the model used for the analytical framework. Even within the model used for the analytical framework there was insufficient clarity on what was being assessed under each "dimension" and there are overlaps in the way that dimensions are reported. So while there is much data and useful information in the report, it is not sufficiently focused to create a clear argument. There was also insufficient balance between the use of the model and the evaluation questions to create an overall analytical framework for the evaluation. These challenges as well as the data collection limitations affects the robustness of the findings (2.55). The recommendations were however found to be useful (4.00).

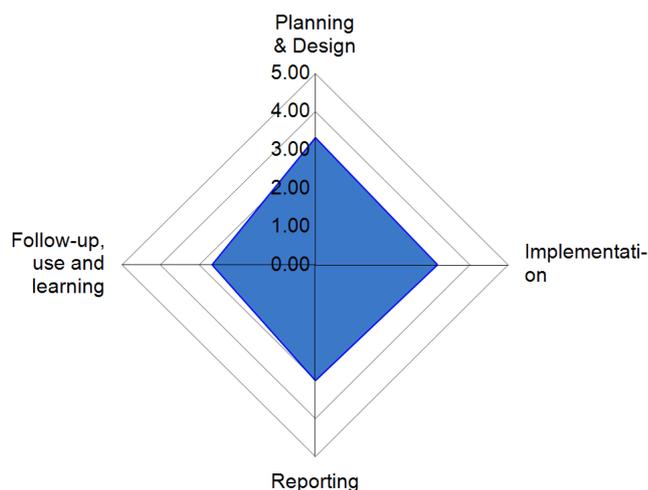
The evaluation was not well resourced in terms of time of budget (2.00).

Quality Assessment Scores

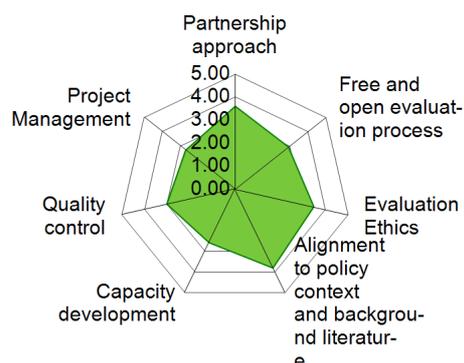
Phase of Evaluation	Score
Planning & Design	3.32
Implementation	3.18
Reporting	3.02
Follow-up, use and learning	2.67
Total	3.04

Overarching Consideration	Score
Partnership approach	3.60
Free and open evaluation process	2.94
Evaluation Ethics	3.50
Alignment to policy context and background literature	3.80
Capacity development	2.60
Quality control	3.00
Project Management	2.71
Total	3.04

Scores: Phases of Evaluation



Scores: Overarching Considerations



Phase of Evaluation	Area of Evaluation	Score
Planning & Design	Quality of the TOR	4.25
Planning & Design	Adequacy of resourcing	2.00
Planning & Design	Appropriateness of the evaluation design and methodology	2.91
Planning & Design	Project management (Planning phase)	3.00
Implementation	Evaluation ethics and independence	3.50
Implementation	Participation and M&E skills development	3.57
Implementation	Methodological integrity	3.00
Implementation	Project management (Implementation phase)	3.00
Reporting	Completeness of the evaluation report	3.00
Reporting	Accessibility of content	3.00
Reporting	Robustness of findings	2.55
Reporting	Strength of conclusions	3.00
Reporting	Suitability of recommendations	4.00
Reporting	Acknowledgement of ethical considerations	3.29
Follow-up, use and learning	Resource utilisation	2.00
Follow-up, use and learning	Evaluation use	2.83
Total	Total	3.04

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 for the evaluation clearly expressed the background to VAWC in South Africa. The purpose was clear (being to strengthen government interventions to address VAWC), and intended users listed with the likely ways in which they would use the results. Evaluation questions and sub-questions were listed in the ToR, according to two major areas of interest: 1) Sector review; 2) Assessment of relevant government programmes. Ideas for the methodology and sample were also provided in the ToR, without being too prescriptive. Deliverables and time frames were also set out. In addition, information on the evaluation team requirements and submission requirements were also provided. What could have been made more explicit was the scope of the document review analysis as this had a few components: review of literature on VAWC, legislation relevant to VAWC, international best practice, budgetary analysis, mandates and policies and plans at national and provincial levels.
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:	There was a good fit between the type of evaluation and the purpose and scope of the ToR, although the evaluation, as originally conceptualised in the ToR, was also to some degree implementation focused (as its interest was in "assessing the effectiveness of a number of national Programmes").
Rating:	4: The approach and type of the evaluation was well-suited to the purpose and scope of the evaluation TOR
Standard:	The TOR (or an internal evaluation proposal) identified the intended users of the evaluation and their information needs
Comment and Analysis:	Each potential information user (including civil society) was identified, and the potential use of the evaluation findings explored in the ToR.
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 Inter-ministerial committee technical task team (IMC TTT) for VAWC comprises some of the key departments responsible for addressing VAWC. IMC TTT initiated this evaluation and determined its objectives. DPME and UNICEF played a key role in drafting a detailed ToR, which was based on a number of consultations with the IMC TTT. The ToR was thus very much driven by the key stakeholders through the IMC TTT, and was supported by DPME. The scope of this TOR also considered other sector research on the direct causes of VAWC (also conducted during the evaluation period). Furthermore, after contracting, two workshops were conducted with sector experts (including civil society) to determine the scope and purpose of the evaluation. While this was done after the ToR had been developed, their input was used to shape the evaluation process. Stakeholders were therefore highly engaged in scoping the ToR and evaluation itself. With this level of engagement the ToR took 5 months to develop and finalise.
Rating:	5: A wider range of key stakeholders (i.e. beyond government stakeholders) were meaningfully involved in the scoping of the TOR and chose the purpose of the evaluation with clear consideration of other relevant sector work

Adequacy of resourcing

Standard:	The evaluation was adequately resourced in terms of time and budget allocated
Comment and Analysis:	The plan was to conduct the evaluation over 6 months for both the sector review and the programme evaluation components. Given the scope of the evaluation and the nature of data collection within government sector, this time frame was not realistic. The budget allocation was also reportedly insufficient given the scope of the evaluation.
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:	<p>The evaluation was challenging in terms of methodology and so it went through 2 bidding processes. The evaluation was awarded to a consulting firm, and the contract was subsequently terminated due to methodological challenges which the team could not manage. KPMG was then selected to conduct the diagnostic review.</p> <p>There were a number of subject matter experts on the team, which was a strength. It was felt, however that these were not always sufficiently utilized. The DPME thus encouraged the increased involvement of these sector experts (e.g. to attend meetings) to ensure that they influenced the evaluation process.</p> <p>The service provider (and particularly the core team, excluding experts) did not have a track record in working with government evaluations. Some of the resulting challenges was that plans for implementation did not properly consider government processes; and there were also challenges around navigating the government system particularly around data collection.</p> <p>While the team was comfortable with the quantitative data analysis methods needed for this evaluation, qualitative data analysis skills (explored later) were not as strong.</p>
Rating:	2: The evaluation was under-staffed or lacked some skills sets appropriate for the type and sector of the evaluation

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:	The Theory of Change was not referred to in the ToR or inception report. The review process however made provision for developing a Theory of Change (under the first contractor). This was expressed as a narrative summary as well as a diagram. There was alignment between the Theory of Change and what was explored in the review.
Rating:	4: The intervention logic or theory of change meaningfully informed and shaped the TOR or the Inception Report, including a visual representation

Standard:	The planned methodology was appropriate to the questions being asked
Comment and Analysis:	<p>The data collection methodology included an extensive literature review, document review (M&E records, budgets, Plans, etc.), online survey and semi-structured face to face interviews with Departmental staff. Overall, this was deemed appropriate for the evaluation questions.</p> <p>The online survey, within the government context, was not the most appropriate as it was anticipated that the surveys would not receive a good response. Also, this format did not allow the complexities around VAWC programming to be explained by the interviewer in order to appropriately answer questions. So, while using a survey was appropriate, it perhaps would have been more effective to administer this survey.</p> <p>The planned data analysis methodology was not entirely articulated in the Inception Report or proposal.</p>
Rating:	2: The planned methodology was not entirely appropriate for addressing all of the questions being asked

Standard:	The sampling planned was appropriate and adequate given the focus and purpose of evaluation
Comment and Analysis:	<p>The sample, which covered 11 national departments (Departments of: Basic Education, Correctional Services, Health, Higher Education and Training, Social Development, Home Affairs, Treasury, National Prosecuting Authority, Department of Women in the Presidency, South African Police Services, Justice and Constitutional Development), and 7 provincial departments (Departments of: Education, Health, Social Development, Home Affairs, Treasury, Community Safety and Premiers Office) (for each of the provinces).</p> <p>Although not originally planned, 4 municipalities were selected for the review process as it emerged (through discussions with sector experts through the stakeholder workshops) that implementation of VAWC services also happened at this level. These municipalities were: City of Cape Town Metropolitan Municipality (Western Cape), City of Johannesburg Metropolitan Municipality (Gauteng), OR Tambo District Municipality (Eastern Cape) and Vhembe District Municipality (Limpopo). This was chosen as it was thought to provide a good mixture of rural and urban sites.</p> <p>A purposeful sampling method for both the online survey and face to face interviews allowed for all relevant Departments and individuals to be included in the sample.</p> <p>Two online surveys were distributed to national and provincial departments, although it is not clear whom this was to. 39 respondents completed the Institutional Survey, and 10 completed the programme survey.</p> <p>For face to face interviews: As the responsibility for VAWC varied within each department, the evaluation team relied on the Head of Departments to advise who should be interviewed; all these interviewees were at Director level or above. A total of 13 national interviews were conducted. A total of 73 interviews were conducted at provincial level. Overall, representatives of 91% of national departments were interviewed and 80% of the provincial representatives sampled were interviewed.</p> <p>The diagnostic review could have been enriched through involving some NPO sector interviews to provide their insights (as key partners of government) into the various aspects reviewed; but this however was beyond the scope of this evaluation.</p> <p>As originally planned, the second part of this evaluation was an implementation evaluation. For this part it would have been beneficial to include programme beneficiaries to provide their perspectives on the effectiveness of programme implementation.</p>
Rating:	3: The sampling planned was appropriate and adequate given focus and purpose of 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:	<p>The inception phase was used to develop a common agreement about how to proceed with the evaluation. An inception report documented the issues raised and decisions taken.</p> <p>The scope of the evaluation however did change quite significantly following this inception phase. Initially the evaluation included: 1) A sector review and 2) a programme evaluation (of about 10 VAWC programmes within government). Part 2 was dropped, as it became apparent that the scope of the evaluation was too broad. The sector review was thus expanded (to include, for example Municipal level interviews).</p>
Rating:	3: The inception phase was used to develop a common agreement on 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: Appropriate ethical procedures were followed. Since no vulnerable groups were interviewed, ethical approval was not necessary. Ethics around anonymity and informed consent was adhered to; this was clearly documented in the data collection tools reviewed. The anonymity of interviewees was upheld in the report; where quotes were used, only the Department name or titles were recorded. Evaluation ethics was not documented in the evaluation report.

Rating: 4: There was clear evidence that ethical protocols were observed for most data collection instances including: informed consent agreements; confidentiality; documenting and storing data notes, recordings or transcripts; Where data was gathered in contexts where ethical sensitivity is high, appropriate clearance was achieved through an ethics review board; e.g. in evaluation involving minors, institutions where access usually requires ethical or bureaucratic clearance, and situations where assurances of confidentiality was offered to participants

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 could work quite independently, while still getting the input from key stakeholders in the process. The DPME as well as the Chair of the Inter-ministerial Committee TTT played a key role in protecting the independence of the evaluator.

There were challenges with accessing data particularly on budgets; but this was mainly because these budgets were not packaged (i.e. clearly aligned with VAWC activities) as initially anticipated.

Rating: 3: The evaluation team was able to work without significant interference and was given access to existing 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: The Inter-ministerial committee TTT for Violence against Women and Children acted as the Steering committee for this evaluation process. This was very strategic as all members of the committee had the authority to make decisions and were "politically attuned". Though there were competing issues that needed attention at this level, it was worth getting these important stakeholders involved in the evaluation process. The IMC TTT played an active role in shaping the evaluation process from its inception through to report writing. All decisions were made collaboratively between DPME and the IMC TTT.

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 evaluand and evaluators was incorporated into the evaluation process
Comment and Analysis:	<p>There was no purposeful capacity development with the partners responsible for the evaluand; rather this happened informally.</p> <p>A strength however, was the capacity development efforts made within the evaluation team. Six unemployed graduates were targeted for capacity development. Graduates underwent training to introduce the programme and familiarize them with the analytical framework to be used. These graduates assisted with the document review process, and accompanied skilled fieldworkers to interviews, where they took the responsibility for note taking. Each graduate was assigned a mentor/buddy who would review their work and provide feedback. The ToR did not request for capacity development.</p>
Rating:	3: An element of capacity building of partners responsible for the evaluand and evaluators was incorporated into the evaluation process

Methodological integrity

Standard:	A literature review was developed which informed the analytical framework and findings of the evaluation
Comment and Analysis:	<p>An extensive literature review was conducted and a number of models (mostly international) were considered for the analytical framework. The WHO readiness assessment model was used as this was thought to be the most comprehensive, relevant (to low and middle income countries and for conducting assessments at national, provincial and local government levels) and generally contained the elements suggested in other models. The report was then structured according to this framework. Findings and conclusions were based on the WHO model.</p> <p>Differentiating between the dimensions was not always obvious. For example, issues around "leadership and political will" is not easily differentiated from "funding"; Likewise a key issue around indicators is covered under "mandates and policies", but this would have been more appropriate under "data, monitoring and evaluation". Although the dimensions are interlinked, it would have been useful to define each category, in terms of what should be covered, and report on this as appropriate.</p> <p>Another limitation is that there was not always a clear alignment between the evaluation questions and the WHO readiness model (to create a framework for the evaluation). This was evident in that the table showing alignment between the model and the evaluation questions omitted few evaluation questions and few dimensions of the model.</p> <p>It is thus evident that there was a strong emphasis on the literature to inform the framework and interpretation of findings; but perhaps not sufficient alignment of the model and its findings to the evaluation questions.</p>
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:	<p>The methods were consistent with those planned, however the challenge was that some of these methods, particularly the online survey, was not appropriate to start off with. This affected the adequacy of the data.</p> <p>The online survey did not receive a good response (44% and 29% response rate for the two surveys), considering the participants were purposefully selected and officially required to respond to the survey. This again suggests that the data collection approach was perhaps not appropriate.</p>
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 schedule was piloted in the first interviews conducted, and minor adjustments were made. The online survey was not piloted; this may have helped test the assumptions about the feasibility of self-completion as well as test the clarity and interpretation of questions.
Rating:	2: A pilot of data collection instrumentation occurred but not in a way that could meaningfully test or improve upon instrumentation
Standard:	Data was collected from key stakeholders (e.g. implementers, governance structures, indirectly affected stakeholders) as data sources
Comment and Analysis:	Interviews were conducted with key stakeholders at government level. Interviewees were identified by Heads of Departments, and interviewees mostly comprised of Heads of Departments, Senior Managers and Directors. In some instances the correct data source (person to be interviewed) was not identified in the evaluation process. Institutional survey data was not obtained (due to non-responses) from Department of Home Affairs, Department of Correctional Services and South African Police services. The study could have benefited from interviews with NGO sector, but this was beyond the scope of the evaluation.
Rating:	3: Data was collected from key stakeholders (e.g. Implementers, governance structures, indirectly affected stakeholders) as data sources
Standard:	The methodology included engaging beneficiaries appropriately as a key source of data and information
Comment and Analysis:	In its original form, that is the objective being to conduct a sector review and a programme evaluation, it would have been appropriate to interview beneficiaries or use secondary data from beneficiaries. However, as the scope of the review shifted to focus on a sector review only, (i.e. the institutional mechanisms), it was not crucial to interview beneficiaries. What would however have been useful is include the views of the Non-governmental sector, particularly those receiving state support.
Rating:	2: The methodology included beneficiary representative perspectives but did not include beneficiaries directly as a key source of data

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:	The Inter-ministerial committee TTT on VAWC fulfilled the role of the evaluation steering committee. This meant that they were not only mandated to oversee the evaluation process, and so it was sometimes challenging to have the evaluation agenda prioritized. There were also instances where Department officials did not facilitate access (in time) to their respective departments. The constantly changing participants also sometimes hindered buy in to the evaluation process. So while the IMC TTT was a strategic chosen steering committee, there were practical challenges with using this forum to successfully achieve the evaluation objectives.
Rating:	3: The steering committee, technical working group and service provider worked together adequately to facilitate 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:	<p>A strength of the support provided for the evaluation process was that both the DPME and the IMC TTT chair protected the independence of the evaluator and in this way was supportive.</p> <p>Although progress meetings were constantly held, it was felt that the DSD programme representatives (other than M&E) should have attended these meetings and assisted with navigating around the problems experienced and assisting with the specific details on the sample (who to be included/targeted) much earlier on the process.</p> <p>There were varied levels of support within the IMC TTT. While some participants reviewed the report, were engaged in sessions, and provided feedback to the report; others were not as involved, and even lacked constructive inputs into the reporting process. For example, some participants criticized the evaluation report due to gaps, however when asked to provide additional information, those offering criticisms would not be forthcoming.</p> <p>Another challenge was that the IMC TTT did not always work efficiently to assist with securing interviews within their own Departments or make decisions about the evaluation process (e.g. scope change).</p>
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: There were a few iterations of the draft report (about 5). The IMC and DPME needed to provide guidance towards getting the report to the required standard. This was in relation to using evidence adequately to support findings (for example triangulation of data sources and inclusion of qualitative data/quotes) as well as framing the report appropriately for government.

Rating: 2: A first draft of the evaluation report was of a poor quality and required major changes

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 key components of the report were included in the final report (although appendices were not found in the same report). This included: executive summary, introduction, background to the evaluation, structure of the report, methodology, definitions, analytical framework, evaluation findings (according to the analytical framework), conclusions and recommendations. The data analysis methods and ethics were not however covered in the report.

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 evaluation report overall is user-friendly, particularly the summary report. The language is accessible and the formatting is consistent. Although a list of references was referred to in the Table of Contents of the report, this was not part of the report (and was therefore not reviewed). References were noted in the footnotes of the report. Source documents were not referenced in full on the first occasion that it was mentioned in the report.

An introduction to each subsection, particularly the key components addressed in the subsection, evaluation questions answered and aspects of the theory of change covered in the subsection, would have provided better guidance to the reader and improved the reader-friendliness of the report. The readiness summary at the end of each chapter helped to consolidate the findings.

Rating: 3: The final evaluation report is user-friendly, written in accessible language and adequate for sharing (e.g. some spelling, grammar or formatting mistakes but these do not seriously detract from the report)

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:	Graphs and tables were mostly used for quantitative data, and figures were used for existing models referenced from other sources. Section 5.5 (Funding and budgets) made extensive use of tables and graphs, and appropriately so. There could have been more use of tables and diagrams for summarizing qualitative data (for example, tables 11, 12 and 18). Also using numbers in tables presenting qualitative data would have provided good snapshots of data. The appropriate language was used for reporting data.
Rating:	3: Figures, tables and appropriate conventions are used in presentation of data and are readily discernible to a reader familiar with data presentation conventions

Robustness of findings

Standard:	Data analysis appears to have been executed to an adequate standard
Comment and Analysis:	<p>Quantitative data was analysed using Survey Monkey and data was often disaggregated according to province. This was appropriate.</p> <p>The WHO dimensions were used as a basis for analysis. The sector review measured each dimension on a readiness scale: not ready, partially ready, ready (it seems from the findings). Readiness in the report is explained as "how 'ready' a country, province, or community is to implement evidence based child maltreatment prevention programmes on a large scale" (Report, p. 24). This scale was not further explained in the report and it is not quite clear how this measure of readiness described above (which seems to focus on children only, and looks at prevention programs specifically) was appropriate for the VAWC. It seems that "readiness" question should rather have been framed within the context of this sector review. It is therefore not clear how the "readiness" status was derived, and also what this in fact means for the sectors readiness/response.</p> <p>Qualitative data analysis was conducted using a thematic analysis. A fieldworker debriefing workshop was facilitated where transcripts were analysed to support or refute some of the theories. Data was not however systematically coded.</p>
Rating:	2: Data analysis was executed to an extent but it appears inadequate or significantly lacking for some 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>A challenge is that the data in the findings section is not framed within the evaluation questions or parameters of what is being explored under the relevant dimension (that is, as the dimension is defined in the earlier sections of the report). So it is not always clear what kind of evidence is necessary, and whether the evidence provided is sufficient.</p> <p>The evidence/data sources within the findings is also not always clear as sometimes statements are made without referencing the source (particularly for primary data).</p> <p>Qualitative data was not always used sufficiently to explain the findings. One example is that interviewees stated that VAWC is not a priority, but the reasons justifying these perceptions are not provided, which then does not provide good evidence.</p> <p>Data on perceptions were gathered through the surveys. Examples of this includes: perception of whether VAWC is prioritized, perception of whether a budget is allocated, etc. The relevance of asking perception questions is not always clear, particularly when more reliable evidence is available. An example of this is that participants were asked whether they thought that VAWC outcomes and targets were included in departmental plans; however this would be better determined through a review of departmental plans (which was conducted). In this example, the relevance of a "perception" question is not made explicit; and also where there are contradictions in data on perceptions (interviews) vs. actual data (plans), this is not properly analyzed for meaning.</p> <p>While there is a lot of data in the report, it is not always analyzed and presented systematically (excluding the unnecessary data) to provide a clear arguments. Having clear parameters of what is being explored under each section (covering both the evaluation questions and WHO dimensions) would have structured the arguments more.</p>
Rating:	2: The evidence gathered has been analysed to support the argument to an extent but this is not entirely sufficient or appropriate, and different data sources may be presented separately rather than integrated
Standard:	There is appropriate recognition and exploration of the possibility of alternative interpretations
Comment and Analysis:	This was not observed in the report, mostly because the analysis was rather descriptive . There were opportunities to explore more explicitly, why for example, VAWC is perceived to not be prioritized.
Rating:	2: There is an implicit or indirect recognition of alternative interpretations
Standard:	The report appears free of significant methodological and analytic flaws
Comment and Analysis:	<p>The data collection methodology and its limitations were well documented in the report, however data analysis was not. Both the data collection methodology (online survey particularly) and analysis methods were not always appropriate, but on the whole, the key findings were well supported.</p> <p>A peer review for this evaluation was undertaken.</p>
Rating:	3: The report appears free of significant 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:	Limitations are well detailed in the report for each method of data collection. The potential reasons for the challenges and limitations and also how these were addressed (where possible) are also articulated. Different types of limitations are considered: challenges with data collection (getting the right information, securing interviews, etc.), limitations to the appropriateness of the evaluation design (e.g. data collection methods - online survey), implications for data quality and interpretation.
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:	Conclusions were quite succinct and based on the key issues presented in the report. Although there were challenges with the methods used to analyse and present data (particularly qualitative data), the data summarised in the conclusions focused on only the key findings that had clear, sufficient evidence.
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:	There is not explicit reference made to the evaluation purpose or questions in the conclusion. It is however clear that some questions are answered through the conclusion, and others not (geographical/demographical coverage, international best practices and effectiveness of the approaches for collaborative planning). There seemed to generally be a focus on the WHO model as the analytical framework, however the evaluation questions should have been integrated to provide an overall framework for the evaluation process.
Rating:	2: Conclusions address the original evaluation purpose and questions in implicit or indirect terms to an extent

Standard:	Conclusions are drawn with explicit reference to the intervention logic or theory of change
Comment and Analysis:	The WHO dimensions are elements of the theory of change that was developed in the initial evaluation stages. This formed the basis of the conclusion and provide a good sense of the strengths and gaps that would inform future VAWC programming. Besides reporting against the elements of the theory of change, there was no unpacking of the design of the Theory of Change to determine its appropriateness. This may also have benefited the further development of a Theory of Change for VAWC.
Rating:	3: Conclusions are drawn with explicit reference to 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:	The recommendations were initially formulated by the Evaluation Team (which included sector experts), and then presented, discussed and further developed with the IMC TTT. UNICEF, was also consulted around the recommendations. It was thus a consultative process.
Rating:	4: Recommendations are made with relevant government officials, stakeholders including beneficiary representatives and sectoral experts beyond the project steering committee, making a significant contribution

Standard:	Recommendations are useful- they are relevant, specific, feasible, affordable and acceptable
Comment and Analysis:	Although the feasibility of the recommendations is not yet known, they have been formulated in detail with guidelines for how recommendations could be implemented. Who is responsible is not always communicated within the recommendations. The recommendations were found to be useful within the current policy context; the section on budgeting was found to be particularly important for the IMC.
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:	Ethical procedures were not documented in the report.
Rating:	1: The full report fails to document any procedures to ensure confidentiality and secure informed consent where appropriate.

Standard:	There are no risks to participants or institutions in disseminating the evaluation report on a public website
Comment and Analysis:	Stakeholders interviewed did not mention any concerns or risks associated with disseminating the report.
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

Follow-up, use and learning

Resource utilisation

Standard:	The evaluation was completed within the planned timeframes and budget
Comment and Analysis:	<p>The plan was to conduct the evaluation over 6 months for both the sector review and the programme evaluation components. The evaluation (sector review only) was conducted over about one year and 2 months; thus significantly exceeding the original time frames. Reasons for this include: given the scope of the evaluation and the nature of data collection within government (i.e. permissions required, setting up interviews at Senior levels, access to documentation, turnaround time for feedback, etc.) this time frame was not realistic to start off with; the challenges with data collection that were experienced; the organizational structure changes within the Department of Women, Children and People with Disabilities further challenged sourcing the right interviewees.</p> <p>The consultant spent more than the allocated budget allocation. The value for the service provider though was this it allowed them to make some inroads to the sector. These shifts to the original evaluation design was discussed and approved by the IMC.</p>
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 first and second drafts of the report was presented to the IMC TTT.
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:	Reflection on the evaluation process and learnings took place during implementation. This did not seem to be formal though, and thus lessons for future evaluations were not captured.
Rating:	2: The steering committee undertook a meeting in which some form of reflection occurred, but not in a clear, reflective process
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:	<p>A number of requests for the presentation of the Diagnostic Review has been requested suggesting that the evaluation has brought attention to VAWC sector. UNICEF also played a role in elevating the status of the evaluation internationally.</p> <p>The evaluation was felt to be useful overall by the DSD representative interviewed.</p>
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:	The evaluations was found to be useful, and supported much of what was already known in the sector. The findings around the budget and funding as well was found to be particularly useful.
Rating:	3: The evaluation study is of conceptual value in understanding what has happened and possibly in shaping policy and practice

References

KPMG, 2015, Phase II proposed scope change

VAWC Theory of Change. 2015

KPMG, Survey instrument.

KPMG, 2015. Inception report: Diagnostic Review of the Country's Response to Violence Against Women and Children. 3 February 2015.

KPMG, 2016, Report on the Diagnostic Review of the State Response to Violence against Women and Children.

DSD, DPME, UNICEF, 2014, Terms of Reference for the Diagnostic review fo the country's response to violence against women and children.

KPMG, 2016, Report on the Diagnostic Review of the State Response to Violence against Women and Children (summary report).

KPMG, 2016, Evaluation findings presentation.

List of Interviewees

Conny Nxumalo, National Department of Social Development, Deputy Director-General: Welfare Services (and Chairperson of the Interministerial Committee Technical Task Team). Telephonic interview conducted on: 11 October 2016.

Lisa Gahan, KPMG: Head of Human and Social Services. Telephonic interview conducted on: 25 October 2016.

Laura Brooks, KPMG Human and Social Services: Senior Consultant. Telephonic interview conducted on: 20 October 2016.

Matodzi Amisi, Department of Performance Monitoring and Evaluation: Director of Evaluation & Research. Telephonic interview conducted on: 11 October 2016.