



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 Drug Master Plan 2013-2017
Evaluation Number:	525
Evaluation Completion Date:	12 May 2016
Period of Evaluation:	10 months
Submitted:	30 November 2016 by Cara Hartley
Approved:	30 November 2016 by Mike Leslie

Evaluation Details

Evaluation Title:	Implementation Evaluation of the National Drug Master Plan 2013-2017
Evaluation Number:	525
Evaluation Completion Date:	12 May 2016
Created:	14 September 2016 by Mike Leslie
Submitted:	30 November 2016 by Cara Hartley
Approved:	30 November 2016 by Mike Leslie
Period of Evaluation:	10 months
Known Cost:	R 1 588 875.00
Known Cost Type:	Referenced
Initiated By:	Department of Social Development and Department of Planning, Monitoring and Evaluation
Initiated By Internal:	Yes
Undertaken By:	Southern Hemisphere
Undertaken By Internal:	No

Assessors

Cara Hartley	cara@pdg.co.za
--------------	----------------

Assessment Documents

Document Name:	Document Type:	Added By:	Added On:
Final Full NDMP Report_16 05 11.docx	Evaluation report	Mike Leslie	14 September 2016
Final Summary NDMP Report_16 05 11.docx	Evaluation report	Mike Leslie	14 September 2016
Final Inception Report-NDMP-24.08.2015.docx	An inception report	Cara Hartley	27 September 2016
Annexure 3 Consent form_interview and focus group respondents.docx	Any other relevant documentation pertaining to the evaluation process	Cara Hartley	27 September 2016
Annexure 2 NDMP TOC_logframe.docx	Assessment tools	Cara Hartley	27 September 2016
National Master Plan Proposal_revised.v2.docx	The proposal selected to conduct the evaluation	Cara Hartley	27 September 2016
NDMP SCM A1a ToR Template.pdf	Terms of Reference (ToR) for the evaluation	Cara Hartley	27 September 2016
NDMP Lit Review 27.10.2015.docx	Any other relevant documentation pertaining to the evaluation process	Cara Hartley	24 November 2016
NDMP fieldwork report - 4 dec.docx	Progress reports	Cara Hartley	24 November 2016
Assessment Report at Moderation.pdf	Assessment Report at Moderation	Mike Leslie	30 November 2016

Quality Assessment Summary

This evaluation scores 3.71 overall, indicating an evaluation of good quality. The evaluation was well implemented throughout: With a score of 3 being adequate and 4 being good, all four of the main Evaluation Phases scored between 3.3 - 3.9. Looking at the evaluation in terms of 7 overarching considerations, 6 out of the 7 were rated between 3.7 - 4.0.

These factors contributed to a good evaluation:

- The evaluation benefited from guidance and inputs from DPME, a peer reviewer, some of the CDA's sectoral experts who served on the committee, and an ethical review board.
- The evaluation service provider brought a good deal of experience and flexibility to the evaluation.
- A diverse steering committee representing many of the key stakeholders of the NDMP was established during the conceptualisation of the evaluation and remained engaged throughout. Supporting the work of the steering committee, it is worth noting that the workshop on emerging findings was attended not just by national stakeholders (as is often the case with National Evaluation Plan evaluations) but also by Local Drug Action Committees (LDACs), structures that tend to include citizens personally affected by substance abuse. This reportedly gave a very concrete dimension to the discussion.

To single out three overarching considerations:

Project Management was effective throughout. The inception phase featured an "inception workshop" which was beneficial in developing a common agreement about how the evaluation would be implemented; and ensuring that resources matched the evaluation plan. The steering committee worked together effectively to oversee implementation and reporting, and the evaluation service provider drew on careful planning and past experience to implement the agreed methodology efficiently.

Furthermore, there was a relatively Free and Open Evaluation Process: The analysis was based on a commonly agreed and explicit Theory of Change; the evaluation team worked without interference (although some data sources were unavailable for other reasons); and the final evaluation report gives a clear and evidence-based judgment on how the NDMP is implemented.

The only overarching consideration on which the evaluation scored less than adequate is Capacity Development, which scored a 2. This is based on two standards namely the fact that there was little structured M&E capacity building of those responsible for the NDMP, and that there was no formal reflective session to reflect on what could be done to strengthen future evaluations. Nevertheless, the evaluation was implemented with a strong Partnership Approach (4.0) and as a result there were several activities where the steering committee was exposed to evaluation theory/concepts, methods and reasoning.

Overall the particular value of this evaluation is that it integrated multiple divergent data sources and engaged a variety of stakeholders from different sectors and parts of the country effectively in both data collection and discussion of the emerging findings, contributing to clarity and coherence for those involved in implementing this National Plan. It sheds light on how the NDMP has been implemented to address substance abuse as a bio-psycho-social issue in South Africa; but it also has broader lessons for other multi-sectoral initiatives driven by the South African public sector.

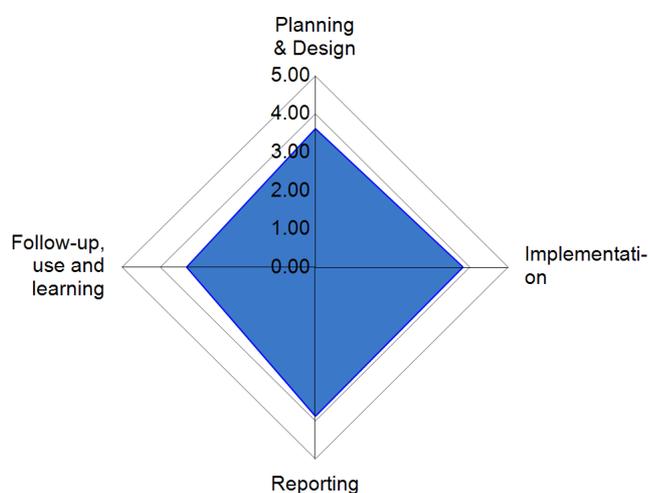
Quality Assessment Scores

Phase of Evaluation	Score
Planning & Design	3.62
Implementation	3.83
Reporting	3.89
Follow-up, use and learning	3.33
Total	3.71

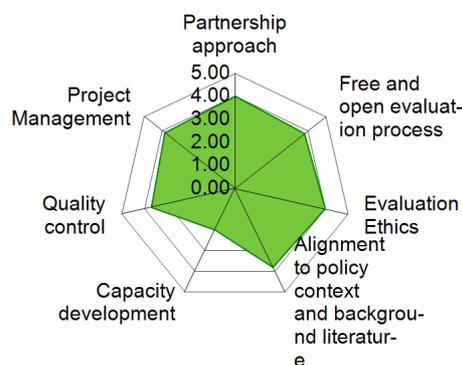
Overarching Consideration	Score
Partnership approach	4.00
Free and open evaluation process	3.83
Evaluation Ethics	4.00
Alignment to policy context and background literature	3.80

Capacity development	2.00
Quality control	3.71
Project Management	3.86
Total	3.71

Scores: Phases of Evaluation



Scores: Overarching Considerations



Phase of Evaluation	Area of Evaluation	Score
Planning & Design	Quality of the TOR	3.56
Planning & Design	Adequacy of resourcing	3.57
Planning & Design	Appropriateness of the evaluation design and methodology	3.36
Planning & Design	Project management (Planning phase)	5.00
Implementation	Evaluation ethics and independence	4.00
Implementation	Participation and M&E skills development	3.14
Implementation	Methodological integrity	3.95
Implementation	Project management (Implementation phase)	4.00
Reporting	Completeness of the evaluation report	4.00
Reporting	Accessibility of content	3.50
Reporting	Robustness of findings	3.85
Reporting	Strength of conclusions	4.00
Reporting	Suitability of recommendations	4.00
Reporting	Acknowledgement of ethical considerations	4.00
Follow-up, use and learning	Resource utilisation	3.00
Follow-up, use and learning	Evaluation use	3.42
Total	Total	3.71

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 included all of the listed components, including some elaboration in terms of subquestions, milestones, sources of data, and the stakeholder population which would be useful to potential service providers seeking to conceptualise the evaluation.
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:	The TOR argued that a programme-theory based, mixed methods approach would be most appropriate given the multifaceted nature of the intervention (a plan with national scope). This is appropriate for an implementation evaluation.
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 explicitly stated the main questions to be answered through the evaluation. It also listed the intended users and distinguished between their possible uses of the evaluation, from which it is possible to extrapolate how the evaluation questions might be of relevance to them.
Rating:	3: The TOR (or an internal evaluation proposal) identified the intended users of the evaluation and their information needs

Standard:	Key stakeholders were involved in the scoping of the TOR and choosing the purpose of the evaluation
Comment and Analysis:	The steering committee including non-governmental stakeholders was formed upfront and key members of this committee, including from DPME, DSD and civil society, were involved in scoping the TOR and choosing the purpose of the evaluation, including a preliminary discussion of what the NDMP's intervention logic is, what type of evaluation would be most useful to NDMP stakeholders at present, and formulation of the evaluation questions to put in the TOR.
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:	The TOR proposed a 10 month time frame (June 2015 to March 2016) but the project only got underway in August 2015. During inception phase, the budget and the scope of the fieldwork were negotiated downward and the agreement was that the project would still end in March 2016. Stakeholders were comfortable that the project plan as per the agreement at the end of the inception phase was adequately resourced in terms of time and budget
Rating:	3: The evaluation was adequately resourced in terms of time and budget allocated
Standard:	The team conducting the evaluation was adequately resourced in terms of staffing and skills sets
Comment and Analysis:	The team consisted of 8 people, including several experienced evaluators, with several with a background in social development in general and community development and youth programmes in particular. An expert consultant was involved to support the budget analysis component of the evaluation. The one area of expertise that some interviewed stakeholders believe may have further benefited the team is in medicine / public health, and the steering committee gave some inputs to improve the evaluators' handling of the medical side of the intervention in one of the evaluation outputs. However, overall the team was well resourced to conduct the evaluation and interviewed stakeholders expressed appreciation for their background in developmental interventions in particular.
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:	The Steering Committee developed a theory of change (TOC) as part of preparing for the TOR, and the TOR indicated that the service provider would need to produce a (revised) TOC as one of the evaluation deliverables. The Inception Report, with the revised proposal as attachment, indicated that the evaluation team would host a TOC workshop with the steering committee and that this TOC would serve as the basis for the evaluation. As for stating what the actual intervention logic of the NDMP is, the planning documents did not do this. The TOR's background section could be interpreted as giving an implied, but not explicit, statement of the NDMP's intervention logic.
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:	<p>The planned methodology was theory-driven i.e. the evaluation would be based on a TOC. Although the TOC was still to be developed at this stage, the evaluation questions already showed a consideration for the NDMP's intervention logic and they were well suited to being situated and unpacked into a TOC. They were phrased in such a way that both quantitative and qualitative data could be used to answer them.</p> <p>A mixed-methods approach to data analysis was planned, with specific qualitative methodologies including: an assessment of coordination between systems; application of a continuum of care approach to the programmatic aspects of the NDMP. Both of these proposed methods held clear value for interpreting and responding effectively to specific evaluation questions. Thematic analysis of qualitative data was also proposed.</p> <p>For quantitative analysis - budgets would be analysed descriptively and comparatively over time to address the evaluation question on resources. A structured survey of private treatment centres would also be quantitatively analysed.</p> <p>Based on the above it was plausible that the planned methodology would appropriately address the evaluation questions.</p>
Rating:	3: The planned methodology was appropriate to the questions being asked

Standard:	The sampling planned was appropriate and adequate given the focus and purpose of evaluation
Comment and Analysis:	<p>Sample size in the primary data collection activities was reduced during the inception stage, but the inception report showed how the sample size would still allow for coverage of the main stakeholders. Instead of data collection in all nine provinces, four provinces (Western Cape, Northern Cape, Limpopo and Gauteng) were purposively selected, with considerations spelled out in the inception report. Furthermore the interviews with private treatment facilities was reduced to be "more realistic" based on discussion in the inception meeting. A broad range of others were also included in the sampling; on the whole the sample appears to have been well selected in terms of covering the breadth of government, government entities, civil society, sector specialists and private stakeholders.</p>
Rating:	4: The sampling planned was good given the focus, purpose and context 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:	<p>On the suggestion of the evaluation team, the inception discussion took the form of a workshop. Since the Steering Committee was already composed by this time and most members took a keen interest in the evaluation (the Committee was described as "very engaged" by an interviewed stakeholder), there was good involvement and discussion at the workshop. This resulted in agreement on changes to some aspects of the evaluation - i.e. time frames; and a reduction of the sample size commensurate with a requested reduction to the budget. It also created the opportunity to go beyond the high level questions provided in the TOR, with the evaluation team working with the Steering Committee to elaborate on their information needs. The good level of agreement achieved through this process was articulated in a clear inception report and confirmed by interviewed stakeholders, who pointed out that it supported a smooth process further down the line.</p>
Rating:	5: The inception phase was used to maximum effect to ensure a common agreement between all stakeholders and a shared 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: The evaluation was approved by an ethical review board. Documentary evidence annexured to the final full report shows that interview / focus group participants were given clear and detailed information about the study and their role in it and then asked for their consent before continuing. This is a high level of ethical behaviour, especially considering that none of the participants were minors, or patients of drug treatment services, who could be seen as potentially vulnerable. Interviewed stakeholders further indicated how data was managed to ensure confidentiality.

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: Interviewed stakeholders were all confident that the evaluation team was able to work without interference, and was given access to all sought data and information sources. There were some challenges and delays in accessing existing data and information but these were not caused by deliberate obstruction but rather related to poor data management and some stakeholders not responding promptly to requests for information.

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 Steering Committee was involved throughout the evaluation, consisting of: Department of Social Development (DSD), Department of Planning Monitoring and Evaluation (DPME), the Central Drug Authority (CDA), Department of Health (DOH), Department of Community Safety (DCS), National Treasury, and other identified experts in the field.

The committee was described as active and engaged, with the key stakeholders taking a keen interest in the evaluation process, sessions, and products, including commenting in detail on the draft reports. There were no representatives of the NDMP's ultimate beneficiaries (people with drug addictions, their families and communities) on the steering committee, but an interviewed stakeholder expressed the view that this was not so crucial given that the evaluation had a strong process-oriented, institutional focus. (See below for discussion of how such beneficiaries, as represented in community-based or non-governmental organisations, were nevertheless engaged at drafting stage.)

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:	There was no formal capacity building of partners responsible for the evaluand, but those on the steering committee were exposed to evaluation concepts and methods, especially during the conceptualisation of the evaluation and drafting of the TOR; the inception workshop during which questions and methods were discussed; and the half-day TOC workshop. Some emerging evaluators also received training and fieldwork experience. However none of these were structured capacity building sessions (i.e. there were no set learning outcomes or assessment of the achievement of these outcomes).
Rating:	2: There was some evidence of capacity building of partners responsible for the evaluand 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:	The literature review was extensive, covering population factors, different approaches to addressing substance abuse, application of a systems approach to substance abuse, and examples from other countries where relevant. At the end of sections as well as in a concluding section at the end, the document explicitly noted the implications of the reviewed literature for how the evaluation should be approached. Based on this the literature review clearly provided insights that benefited the analytical framework and provided good context for the findings.
Rating:	4: A good quality literature review was developed which was insightful in terms of the analytical framework and provided good context for the findings

Standard:	The methods employed in the process of the evaluation were consistent with those planned and implemented adequately
Comment and Analysis:	The data collection methods employed were consistent with those planned. From a reading of the report and stakeholder interviews, methods were well implemented in terms of time (they stayed largely on schedule despite some challenges in the field), coverage (which was as planned, with at least one other stakeholder interviewed where deemed valuable), and content (as planned, including follow-up interviews where needed to complete the content in the interview guide). Interviewed stakeholders and the comments of the peer reviewers support the conclusion that methods were well implemented.
Rating:	4: The methods employed in the process of the evaluation were consistent with those planned and implemented well (in terms of time, coverage, and content)

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:	According to the fieldwork report, key instruments were piloted in the field on 9 and 10 November. This resulted in minor adjustments (as expected there were no major adjustments, since instruments had already been reviewed by three experts before being piloted). While key instruments were piloted, not all of them were as this was not practical given tight time frames and the several types of qualitative data collection instruments developed to engage several types of respondents. The one issue that might have been picked up in piloting and was not, is that the survey of treatment centres might not be entirely appropriate. This was picked up later (see below) and the approach was changed to semi-structured interviews.
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: As per the plan, data was collected from a range of key stakeholders: Provincial Substance Abuse Forums (PSAFs); Local Drug Action Committees (LDACs); CDA members and stakeholders including national departments and public entities; Interministerial Committee task team members; Sector specialists; Civil Society Organisations; Private and non-profit treatment centres. In engaging treatment centres, the approach was changed from a structured questionnaire to semi-structured interviews when the team realised this would be more appropriate.

As detailed in the Fieldwork Report and summarised in the full Final Report, the data collection took place very much as planned, with all the intended groupings engaged (including treatment centres, just with a different data collection instrument), with only small variations in the sample size per grouping (still reaching more than 90% of those intended).

The report notes one set of stakeholders that could have provided important insight but had been overlooked until the final stage of the evaluation ie. members of the judiciary. This is noted but since it was not part of the (consultatively developed) research plan and intended sample, it is not taken to influence the rating on this standard.

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: The beneficiaries of the NDMP are people affected by substance abuse - particularly people with addictions, their families and communities. LDACs, community-based organisations focused on substance abuse, and treatment centres often include people who have overcome addictions or were personally affected by addiction. Therefore engaging these structures in interviews and focus groups allowed for the beneficiary perspective to be included in the evaluation - especially the perspective of beneficiaries that have in-depth knowledge of how NDMP is being implemented on the ground.

The involvement of some LDAC members in the stakeholder workshop also reportedly brought a very concrete dimension to the discussion of draft findings and ensured that realities of local implementation informed the report and recommendations.

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:	<p>According to all interviewed stakeholders, there was a positive and constructive working relationship between the steering committee and the service provider.</p> <p>A high score of 5 is appropriate given some examples of creative and flexible cooperation between these parties - for instance, both parties agreed to workshop the evaluation questions at the inception stage; furthermore as the service provider's understanding of NDMP and the needs of the evaluation users unfolded there was a decision to innovate around the format of the Theory of Change (TOC) as this would better serve the evaluation and this was supported by all; also when the originally planned two-day TOC workshop could not be held the service provider worked creatively to ensure a quality product was nevertheless developed and had the steering committee's full support. The service provider also accepted more than one unanticipated invitation to present about the evaluation, including at a conference at very short notice, and overall these presentations proved beneficial to achieving the objectives of the evaluation especially securing the buy-in of important stakeholders who had concerns about the evaluation findings. The steering committee was described as very engaged and making an effort to assist the service provider to have a good conceptual understanding of the intervention.</p>
Rating:	5: The steering committee, technical working group and service provider worked together in a creative, flexible and constructive manner facilitating achievement of a high quality 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>The secretarial support had some strengths and weaknesses - strengths included the ability to secure the participation of a wide variety of stakeholders in the stakeholder workshop; and timely feedback on drafts to ensure timely completion of evaluation phases. Some challenges were encountered in securing and communicating with the peer reviewer; and some sets of comments on draft reports were not consolidated making it harder for the service provider to respond efficiently. There was also the challenge of slow or non-availability of documents and data (in many cases outside the secretariat's control). On the whole though, these challenges could be overcome, and secretarial support was considered adequate by interviewed stakeholders.</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:	The first draft evaluation report was of sufficient quality to go to stakeholders. According to interviewed stakeholders, the draft evaluation report was satisfactory as a draft. Many comments were received but the essence of the findings were acceptable and remained fundamentally unchanged; changes required were mostly about clarifying statements, improving precision and terminology, and showing more clearly what evidence and data analysis underlay the findings.
Rating:	4: A first draft of the evaluation report was of a good quality and required only minor changes prior to finalisation

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 evaluation report of 127 pages includes all of the above listed components - it is complete but not excessively detailed. The findings are structured according to the main evaluation questions (except two which it makes more sense to deal with differently) and each section ends with a concluding summary - this assists a reader to trace the reasoning on each question. The report also makes good use of appendices to supplement the main report content structure. Appendices that supplement the sections on methodology and findings include the theory of change and log frame; literature review; consent form and ethical clearance letter; and a summary of provincial reports with achievements and activities. Three of the appendices provide further detail on the recommendations: a proposal for a national clearing house; an outline of what should be included in the TOR of PSAFs and LDAC; and an example of a LDAC TOR which was developed by the Western Cape PSAF.
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 easy to read; text is not too dense and generally the point of a sentence, paragraph, list, table or quote is clear. Interviewed stakeholders specifically expressed satisfaction with the user-friendliness and readability of the report. The small grammatical and typing errors do not detract from the report.
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:	<p>Tables and figures are presented appropriately. Where quotes are used their relevance is clear and they are usually accompanied by an indication of whether they represent a strong theme.</p> <p>Two matters motivate for assigned a rating of 3 here instead of 4:</p> <p>(1) The report frequently reports the exact number of respondents in qualitative engagements who "responded positively" to a question, or indicated that they are observing a particular tendency. This was probably done in an attempt to help the reader gauge whether the finding is general or exceptional instead of using vague terms like "most of" or "usually". This can seem to imply that all respondents not "responding positively" or not indicating that they are observing a tendency, were indeed asked the question - which according to the section on limitations was not the case for all the questions in the semi-structured interviews. There is also a risk that the reader may interpret such findings as statistically representative of the whole population from which the sample was drawn, which is not the case in a purposively selected sample. It would have been good to state this caveat.</p> <p>(2) In some tables, monetary figures are presented with no "thousands separator" (R1 000 000 is presented as R1000000) which does not aid interpretation, especially in the discussion of budgets.</p>
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:	The data analysis required for this evaluation was fairly straightforward and appears to have been well executed. The main challenge would have been to consolidate findings across multiple datasets which is discussed below. In reading the report there are no concerns about inadequate analysis of qualitative or quantitative data, nor were any raised by the interviewed stakeholders.
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>The evaluation report flows well, drawing on multiple sources of data as relevant to each point. For most statements, the source of evidence is directly mentioned or can be easily inferred.</p> <p>The only information that is presented somewhat superfluously is a set of budget tables in the section on resources; for some of these the in-text discussion is not extensive leaving these hard to interpret. Some colours, bold text etc. could have assisted, or the tables could have been summarised here to bring out the main point and the full tables moved to the appendix. However this is not a common issue in the report.</p> <p>Overall the authors have very effectively integrated multiple sources of data to present the findings and arguments throughout the report. Given that a very wide range of documents and data was synthesised (see the 9-page reference list), this is an achievement.</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:	<p>Key philosophical divides in the context of the NDMP are between psycho-social and bio-psycho-social approaches to substance abuse; and between different drug control paradigms. These distinctions were taken into account early on with the literature review and subsequently was integrated into the discussion of findings.</p> <p>The evaluation report had the benefit of drawing on a large qualitative sample as well as a range of other sources that served to surface alternative interpretations of issues. This is reflected in the frequent juxtaposition in one paragraph of sources that reflect possible alternative interpretations of a matter. In other cases alternative interpretations are not drawn directly from sources but are nevertheless presented and discussed in the text - this is probably where the evaluation benefited from the engagement of a diverse evaluation stakeholder committee and peer review. (Interviewed stakeholders also praised the evaluators for engaging directly with some important role players when there was discomfort with the findings, and making sure that the report reflects any valid arguments raised in these discussions.)</p>
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:	<p>From a reading of the report, it appears free of significant methodological and analytic flaws. The issue mentioned above, of reporting the exact number of qualitative respondents expressing a particular view, might be somewhat misleading to a reader unfamiliar with qualitative analysis, but it is not misleading to the evaluators, as they do not use these findings as statistically representative.</p> <p>The report also notes some of the methodological and analytical processes used to ensure it is free of flaws, including review of data collection instruments by several parties and piloting, oversight by the steering committee, and for thematic analysis a check of inter-coding reliability by the lead evaluators.</p>
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:	A section on limitations notes challenges with data collection and data availability. In assessing the quality of the evaluation, no other major limitations have been apparent from a reading of the evaluation report and documentation, nor mentioned by stakeholders. The limitations are therefore considered adequately articulated.
Rating:	3: Limitations of all aspects of the methodology and findings are articulated

Strength of conclusions

Standard:	Conclusions are derived from evidence
Comment and Analysis:	Conclusions flow logically from the findings presented in the report, and as already discussed, these are drawn from multiple sources of data that have been well analysed.
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 section on conclusions starts by reminding the reader what the evaluation's purpose was. It then presents the conclusions theme by theme in a way that relates clearly back to the original evaluation questions.
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:	Specific aspects of the Theory of Change are cited in the discussion of conclusions, with some assumptions shown to have proven false, and some links in the causal chain highlighted as having proven particularly important for the achievement of intended impacts. It is particularly useful that the authors acknowledge the difference between "theory failure" and "implementation failure" by arguing that only if LDACs are sufficiently resourced in future and their roles clearly defined, and they still do not achieve the intended outcomes, then the suitability of the structures themselves can be questioned.
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:	When draft findings were presented to stakeholders (which included non-government stakeholders, see below), recommendations were discussed in small groups and debated in plenary. Recommendations were then refined through rounds of comments from the steering committee (which included sectoral experts) as well as inputs from an external peer reviewer.
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:	<p>There are many recommendations - 15 main recommendations with multiple sub-recommendations. Based on a reading of the report they are relevant; interviewed stakeholders are also confident that the recommendations are specific, feasible, affordable and acceptable. Where there was not clear agreement about how practically to increase the autonomy, independence and authority of the CDA (recommendation 3) this was well dealt with by providing options for consideration of DSD and CDA, without being prescriptive.</p> <p>Recommendations also distinguish between different users. The main concern with their feasibility is that, in naming the multiple, specific stakeholders who would need to implement them, they show just how dependent the NDMP is on intergovernmental and intersectoral cooperation to be effective; in this sense the recommendations may fall prey to the same challenges as the Plan at large. However, interviewed stakeholders expressed satisfaction with the recommendations.</p>
Rating:	4: Recommendations are well-formulated for use- they begin to differentiate by user and are relevant to the current policy context, specifically targeted, 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 includes as annexure a full informed consent statement with consent form, which includes information about procedures intended to ensure confidentiality such as secure storing of transcripts and not quoting respondents by name.

Rating: 4: The full report documents all procedures to ensure confidentiality and to secure informed consent and provides some examples in appendices

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

Comment and Analysis: No concerns or risks were flagged in the interviews with stakeholders nor in a reading of the final full report.

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 largely within the time frames and budget. It was expected to run from August 2015 to mid March 2016; by April 2016 there was an essentially final report with only minor edits made before deliverables were finalised in early May. Interviewed stakeholders expressed satisfaction that overall the evaluation process ran "smoothly". In terms of budget the evaluation ended up costing the commissioning departments what had been agreed at inception, and the service provider was satisfied that they did not internally exceed this budget by much.
Rating:	3: The evaluation was completed within the planned timeframes and budget

Evaluation use

Standard:	Results of the evaluation have been presented to relevant stakeholders
Comment and Analysis:	The results of the evaluation have been presented to relevant stakeholders in several ways: <ul style="list-style-type: none">- Draft findings were presented at a Stakeholder Workshop that was attended by a very wide collection of stakeholders, from various government departments nationally and provincially to the provincial PSAFs and LDACs, and civil society organisations. This engagement shaped the recommendations (see above) and a section on Lessons Learned about the NDMP (see below).- The service provider presented the draft findings to the CDA (including members outside government).- After the report was finalised, a member of the steering committee again presented the evaluation findings to the CDA.- CDA members and DSD staff were also given electronic copies of the summary reports and the presentation.
Rating:	4: Results of the evaluation have been presented to all relevant stakeholders, inside and outside of 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:	There was no formal reflective session in which the steering committee and service provider discussed what could be done to strengthen further evaluations. However there was some discussion of the evaluation process and methodology with the steering committee in the final handover of the report.
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:	The evaluation has raised the profile of the National Drug Master Plan. According to an interviewed stakeholder, the sharing of the evaluation findings has "improved the visibility" of the CDA and has made stakeholders more aware of the potential value and relevance of the CDA and NDMP to the problems they are also facing. In this way the evaluation may have promoted better cooperation and support among stakeholders, for example, partner departments which might not previously have grasped the relevance of the NDMP and CDA are reportedly placing a higher priority on reporting back to the CDA than they previously did.
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:	<p>Within the steering committee, most members already understood the NDMP and associated issues quite well at the outset. But from interviews with key evaluation stakeholders it is clear that steering committee members grew in their agreement about what the key challenges are, which is an important step towards appropriate decision making.</p> <p>Furthermore, in this evaluation, not only the steering committee but also the broader network of stakeholders reportedly grew in their understanding of how the NDMP has been implemented, and the realities faced by those involved. The workshop of draft findings with such a wide network of stakeholders, reportedly helped those involved with local and provincial implementation to gain better insight into realities at national level as well as vice versa. A section in the final report, Lessons Learned, was jointly developed in that workshop, showing stakeholders' learning about how NDMP was developed and implemented.</p> <p>It is also reported that beyond the steering committee and those that attended the draft findings workshop, the DSD and CDA staff more broadly have gained clearer understanding and insight about NDMP from the dissemination of evaluation findings.</p>
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

Southern Hemisphere, (2015), Proposal to undertake an implementation evaluation of the National Drug Master Plan (NDMP) of South Africa. 10 June 2015.

Southern Hemisphere (2016), Final Report on the Implementation of the National Drug Master Plan 2013-2017. 11 May 2016. (Including Annexures 1, 2, 3, and 7).

DPME, (2015), Request for proposals for Implementation evaluation of the 2013-2017 National Drug Master Plan (NDMP).

Southern Hemisphere, (2015), Final inception report for the National Drug Master Plan. 24 August 2015.

Southern Hemisphere, (2015), Fieldwork Report for the Implementation Evaluation of the National Drug Master Plan. 4 December 2015.

Southern Hemisphere, (2016), Letter to Peer Reviewer.

List of Interviewees

Dr Evodia Mabuza-Mokoko, Manager: Central Drug Authority, Department of Social Development. Telephonic interview, 12 October 2016.

Matodzi Amisi, Director: Evaluation, Department of Planning, Monitoring and Evaluation. Telephonic Interview, 29 September 2016.

Nana Davies, Senior Consultant, Southern Hemisphere. Telephonic Interview, 5 October 2016.