Report on the Assessment of Government Evaluations

Evaluation Title:	Formative Evaluation of the Gauteng Young Women Development Programme
Evaluation Number:	608
Start Date:	Monday, 01 September 2014
Completion Date:	Saturday, 30 May 2015
Period of Evaluation:	September 2014 - May 2015
Submitted:	Monday, 29 January 2018 by Lerato Shai
Approved:	Wednesday, 31 January 2018 by Mike Leslie

Evaluation Details

Evaluation Title:	Formative Evaluation of the Gauteng Young Women Development Programme
Evaluation Number:	608
Evaluation Completion Date:	Saturday, 30 May 2015
Created:	Thursday, 02 November 2017 by Cara Hartley
Submitted:	Monday, 29 January 2018 by Lerato Shai
Approved:	Wednesday, 31 January 2018 by Mike Leslie
Period of Evaluation:	September 2014 - May 2015
Known Cost:	R 239000.00
Known Cost Type:	Estimate
Initiated By:	Gauteng Office of the Premier
Initiated By Internal:	Yes
Service Providers:	Unleash Potential Unlimited
Undertaken By Internal:	No

Assessors

Lerato Shai	lerato@pdg.co.za

Assessment Documents

Document Name:	Document Type:	Added By:	Added On:
Report Formative Evaluation of YWDP of Gauteng.pdf	Evaluation report	Cara Hartley	Thursday, 02 November 2017
YWDP Inception Report version 3.docx	An inception report	Lerato Shai	Monday, 22 January 2018
ToR_YWDP_25 10 2013 1.docx	Terms of Reference (ToR) for the evaluation	Lerato Shai	Monday, 22 January 2018
Assessment Report at Moderation.pdf	Assessment Report at Moderation	Cara Hartley	Wednesday, 24 January 2018

Quality Assessment Summary

This evaluation is scored 3,05 overall which is adequate. Its main strengths are in the planning and design, and implementation phases where scores are above 3. The follow up, use and learning that took place in the evaluation was also of an adequate level, however the poor quality of the reporting is a key weakness which negatively affects the overall score.

The planning and design of the evaluation scored 3,38 overall. The planning was more than adequately completed in this evaluation with strengths in project management and resourcing. The team demonstrated the necessary skills and a well represented steering committee was setup to manage the evaluation. A weak point is the failure to list and discuss the needs of intended users of the evaluation in the ToR. However, it is notable that beneficiary representatives were included in the steering committee.

The implementation phase score of 3,25 is the product of good beneficiary participation in data collection, and sub-optimal collaboration between stakeholders in the steering committee. The evaluation fully incorporated beneficiaries' inputs, not only in the data collection but also through engagement in the steering committee. The data collection process was informed by a set of ethical principles which are stated in the report with samples of the informed consent forms used provided in an appendix. Moreover, the collection was preceded by a pilot of the instruments that contributed to some improvements in the accessibility of the questions asked. Despite these strong points, the steering committee could have worked more collaboratively to facilitate creative and efficient problem solving throughout the evaluation.

With a score of 2,77, the reporting phase is the weakest aspect of this evaluation. The report suffers from major shortcomings in terms of a draft that required multiple revisions following extensive comments from stakeholders. The structure of the final report has room for improvement to make content more readable and enhance the flow. The report's consistent poor formatting, grammar and spelling mistakes detract from its accessibility and readiness for publication.

The follow-up, use and learning phase score of 3,07 suggests an adequate process. The evaluation was completed outside of the planned timeframes and the report was only presented to stakeholders directly involved in the evaluation. A positive, however, is that reflection did occur in the project close out meeting and the steering committee drew key lessons from the experience.

Quality Assessment Scores

Phase of Evaluation	Score
Planning & Design	3.38
Implementation	3.25
Reporting	2.77
Follow-up, use and learning	3.07
Total	3.05

Overarching Consideration	Score
Partnership approach	3.20
Free and open evaluation process	3.11
Evaluation Ethics	4.00
Alignment to policy context and background literature	3.27
Capacity development	3.00
Quality control	2.73
Project Management	3.14
Total	3.05

Scores: Phases of Evaluation
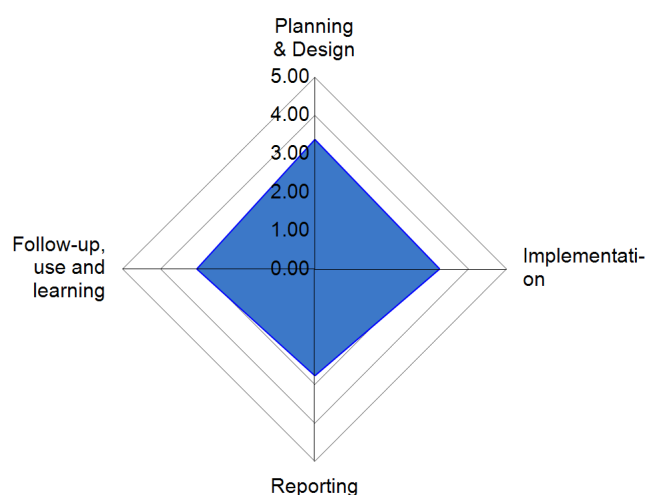
Scores: Overarching Considerations
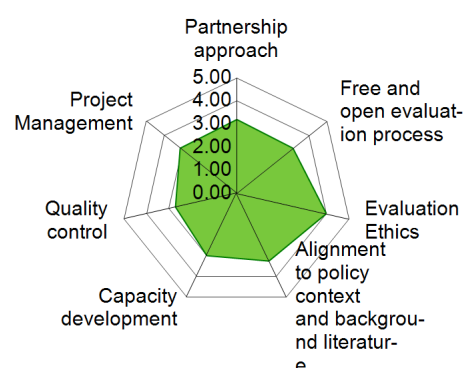
Phase of Evaluation	Area of Evaluation	Score
Planning & Design	Quality of the TOR	3.31
Planning & Design	Adequacy of resourcing	4.00
Planning & Design	Appropriateness of the evaluation design and methodology	2.91
Planning & Design	Project management (Planning phase)	4.00
Implementation	Evaluation ethics and independence	4.00
Implementation	Participation and M&E skills development	3.00
Implementation	Methodological integrity	3.21
Implementation	Project management (Implementation phase)	3.00
Reporting	Completeness of the evaluation report	2.00
Reporting	Accessibility of content	2.00
Reporting	Robustness of findings	2.60
Reporting	Strength of conclusions	3.00
Reporting	Suitability of recommendations	3.50
Reporting	Acknowledgement of ethical considerations	4.00
Follow-up, use and learning	Resource utilisation	2.00
Follow-up, use and learning	Evaluation use	3.33
Total	Total	3.05

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 submitted for assessment is a draft version. Based on this version, the ToR is rated as adequately structured, however some sections are missing. The background section details the programme sufficiently and covers its key objectives. The purpose of the evaluation is clearly articulated and the key evaluations questions for each aspect of the programme are listed. Moreover, the key skills and competencies required to successfully complete the evaluation are provided. The ToR omits a discussion of the intended audience and users of the output.
Rating:	3: The evaluation was guided by a well-structured and complete TOR or internal evaluation proposal of an adequate 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 purpose and scope of the evaluation was an assessment of the programme's delivery, performance against indicators, management arrangements and the documentation of the underlying theory of change. The draft ToR requests a formative evaluation of the Gauteng Young Women Development Programme. The inception report details a mixed methods approach to answer the key evaluation questions. The type is well suited to the programme's stage with the pilot phase completed and ready for evaluation. The approach is sufficient but depends on the availability of relevant data.
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:	The ToR only makes implicit mention of the users and their information needs in the description of the programme and the scope of the evaluation e.g. the background details who the participants are and which partners assisted with sourcing participants.
Rating:	2: The TOR made only implicit or indirect mention of the 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:	According to the programme manager, the development of the ToR was through stakeholder engagement inclusive of various directorates of the Office of the Premier, participating departments, mentors and beneficiary representatives.
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 inception report details a project timeline spanning 7 months with a budget of R239 000. The time allocated in the inception report is sufficient and also provides the evaluation team with some flexibility to complete data collection. The Office of the Premier also assisted with expediting the fieldwork by inviting and transporting beneficiaries to meetings where focus groups could take place. The evaluator noted that the resources were sufficient.
Rating:	4: The evaluation was well resourced in terms of the time and budget allocated (i.e. there was some room for flexibility)

Standard:	The team conducting the evaluation was adequately resourced in terms of staffing and skills sets
Comment and Analysis:	The three person team possessed the required skills and could demonstrate a track record of completing government evaluations. All team members hold doctorate qualifications in their respective fields. The team leader is a psychologist and has worked on gender issues, programme design and monitoring and evaluation. The other team members included another psychologist specialising in the psycho-social and cultural aspect and someone with experience in rural development and sustainable livelihoods. An intern was also on the team to provide administrative support.
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 inception report details an emerging theory of change and results chain for the programme and provides a graphic representation thereof. The version of the inception report assessed includes a theory of change that incorporates input of stakeholders implementing the Young Women Development Programme.
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:	The methodology comprises an inception phase, a desktop review, validation of the emerging theory of change, data collection, analysis and reporting. The analytical framework outlines which aspect of the methodology will be utilised to answer the evaluation questions (inclusive of sub-questions). The planned analysis includes coding the data and analysing it around the research themes. The inception report states that the focus of the analysis was the identification of emerging recurrent patterns and themes. Overall, the methodology is adequate for the questions asked, however it is unclear whether the data collection is planned to be representative of the stakeholders. This affects what claims can be made about the programme as a whole, which some of the questions require.
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:	The sampling planned in the inception report lists the groups of stakeholders that will be interviewed and the type of data that will be collected from them is provided. The planned sample size for programme staff and mentors includes all members in these two groups but it is not clearly stated in the inception report how many potential respondents this translates to. The planned method for both groups was an electronic survey with telephonic interviews conducted with mentors that dropped out of the programme. A focus group was also held with programme staff at inception to tailor data collection instruments to specific aspects of the programme. It is not clear how many and which staff participated in the focus group. The beneficiary sample was planned to include both those who completed the programme and those who dropped out. The Office of the Premier planned to organise meetings where beneficiaries would participate in the focus groups. No sampling was completed for this group prior to collecting data from them. This method of selecting beneficiaries to participate in the sessions lends itself to selection bias and a non-representative sample. Both these aspects have a bearing on the extent to which the purpose of the evaluation can be achieved in terms of drawing conclusions on the emerging outcomes for the programme's beneficiaries. Moreover, it is notable that external partners such as non-profit organisations were not included in the data collection process which would have provided further information on the arrangements in the evaluation.
Rating:	2: The sampling planned was not entirely appropriate given the focus and purpose of the evaluation

Project management (Planning phase)

Standard:	The inception phase was used to develop a common agreement on how the evaluation would be implemented
Comment and Analysis:	According to the programme manager, the inception meeting clarified what the evaluation was all about and ensured all members of the steering committee were on the same page with the service provider. Moreover, the evaluator noted that the objectives of the evaluation and the planned methodology were clarified. One change made to the methodology was the approach to data collection. The Office of the Premier took on the responsibility of organising beneficiaries to come to a central location where data collection can take place. The programme manager noted that the theory of change was a source of contention. The inception report notes that "reconstructing an explicit theory of change has evolved over time and this version is the result of inputs made by the designers of the YWDP. It is possible that as interaction with other stakeholders occurs, it may further be refined." There is therefore acknowledgement that this theory of change is still work in progress and further refinements will occur as more inputs are received.
Rating:	4: The inception phase was used to good effect to achieve a common agreement and understanding of how the evaluation would be implemented

Implementation

Evaluation ethics and independence

Standard:	Where data was gathered in contexts where ethical sensitivity is high, informed consent, assurances of confidentiality and appropriate clearance were achieved; e.g. through an ethics review board, in evaluation involving minors, institutions where access usually requires ethical or bureacratic clearance
Comment and Analysis:	There is clear evidence that all ethical protocols were duly observed in the collection of all primary data. Informed consent forms were explained to and signed by beneficiaries, mentors and programme staff prior to interviews. The documentation and storing of data notes is however not addressed.
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 bureacratic 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 evaluator noted that the team could work without any interference and received all data that was requested. The programme manager further noted that the programme office was responsible for ensuring that the evaluator received all required data. The office also facilitated information from other departments and collated this before sending to the evaluator.
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:	The key stakeholders were involved in the evaluation through a well represented and formalised steering committee. The steering committee met at different milestones linked to key deliverables. The evaluator noted that attendance of stakeholders was not always consistent.
Rating:	3: Key stakeholders were involved in the evaluation 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:	Capacity building of the programme office took place in meetings with the head of strategic planning, monitoring and evaluation informally training staff on key concepts at different stages of the evaluation.
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:	The literature review presented appears to confuse a literature and document review as these are both presented in one chapter and at times in one section e.g. the review of why organisations implement mentoring programmes is immediately followed by a discussion of the targeting approach in the programme being evaluation. The literature review chapter focuses on mentoring (what it is, what makes it work, how to design a prgramme) but does not provide any insight into success criteria of similar programmes. However, these insights do appear towards the end of the report under lessons learned which suggests that these were not sufficiently incorporated into the analytical framework and findings.
Rating:	2: A literature review was undertaken but was not well developed eg a limited set of literature, not sufficently analysed, or not used to inform the analytical framework or findings

Standard:	The methods employed in the process of the evaluation were consistent with those planned and implemented adequately
Comment and Analysis:	The methods employed were adequately consistent with those planned and implemented well in some respects (e.g. collection of data from programme staff and mentors) and adequately in others (e.g. beneficiaries). The beneficiary data collection experienced some delays due to challenges with language and the arrival of participants who did not partake in the programme arriving at the focus group sessions.
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:	A pilot of all data collection instruments was undertaken and it informed the research process. The evaluator noted that following the pilot, the instruments were adjusted to be more user-friendly and accessible. The content did not change. One drawback with how the pilot was undertaken was the selection of participants for the pilot. They were drawn from the pool of interns at the Department of Health and did not adequately represent the overall beneficiaries. These interns were all proficient in English. As a result, critical challenges experienced with beneficiaries' English language proficiency were not identified at this stage and hence not planned for ahead of data collection.
Rating:	4: All components of the data collection instrumentation were piloted which led to some improvements in the data collection instrumentation or affirmation of the instruments

Standard:	Data was collected from key stakeholders (e.g. implementers, governance structures, indirectly affected stakeholders) as data sources
Comment and Analysis:	Three key stakeholder types formed part of the primary data collection: beneficiaries, mentors and programme staff. The evaluation team were able to collect data from all stakeholders as planned. All programme staff responded to the electronic survey. Table 6 in the final report outlines the participation of mentors in the survey. The data provided in this table suggests that only 15 of 55 mentors participated. The sample of mentees was 78 and the final report states that this is 37% of the total. The inception report stated that the target sample for mentees depended on how many attended the meetings, hence no target sample size was set. Hence, while data was collected from all intended stakeholders, the desired sample size was either not achieved or never set at inception.
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:	The methodology included data collection from beneficiaries through focus group interviews. These sessions enabled beneficiaries to share their experience of the programme's different phases.
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 beneficaries 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:	The programme manager described the relationship between the different stakeholders as dynamic and noted that there were often cases of tension between the service provider and the internal team at the Office of the Premier. When asked what would have been different without the tension, the programme manager noted that the collective knowledge of wider team would have contributed better to the evaluation. The evaluator's view is that the commissioning organisation was sufficiently strong headed, however the working relationship remained cordial.
Rating:	2: The relationship between the steering committee, technical working group and service provider was inadequate with some challenges to the 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:	The programme office played the role of secretariat in the evaluation. The evaluator noted that this function was performed well and facilitated dialogue between the service provider and the steering committee.
Rating:	4: Good support was provided by the evaluation secretariat and facilitates timely and constructive 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 of the evaluation report required significant changes from respective stakeholders. The programme manager noted the following set of comments with respect to the theory of change, referencing and accessibility: – the theory of change was a bone of contention. Internal stakeholders were not satisfied with the standard of the theory of change and stated that it was not well defined within the context of the programme. This resulted in further revisions. – there were extensive comments around referencing and formatting of the report. The poor quality of referencing and lack of editing caused tension as internal stakeholders found this to be of substandard quality ,especially when compared to reports from evaluations in other units of the Office of the Premier.
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 final report contains all the key sections, however it is poorly structured. There is scope to structure the report better to facilitate flow from one section to the next e.g. the theory of change, analytical framework and detailed methodology are all subsumed in chapter one which is titled background to the Gauteng Young Women Development Programme. Moreover, given that the theory of change is a deliverable of the project, it is better followed by a discussion of the methodology and analytical framework. Findings from a meta-analysis of similar mentorship programmes is provided in a late chapter of the report on lessons learned instead of the literature review. The limitations of the methodology are also discussed prior to concluding the report instead of directly after the methodology.
Rating:	2: The final evaluation report is poorly structured and does not address all of the following components: executive summary; background/context of the 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 writing in the report is largely accessible, however the report is characterised by frequent formatting, spelling and grammar mistakes which detract from the accessibility of the report e.g. "The fluctuations in attendance figures given in the programme records , for both mentors and mentees and mentor as mentee attendance patterns as captured in the programme records suggests that that both mentors and mentees dropped out without notifying the programme office." More than a handful of such cases are present in the report. More importantly, the referencing in-text and in the list is not consistent, often key statements and data are not referenced in the body of the report. E.g the paragraph in section 1.1.1 discussing youth unemployment reports nine different percentages from external sources, none of which are referenced.
Rating:	2: The final evaluation report is characterised by either inaccessible language or frequent formatting, spelling and grammar 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:	The report uses simple tables, bar and pie charts to present data. The tables follow appropriate conventions and are easy to discern, however the charts often lack legends to describe what the colours signify and/or they do not show the data labels. In such cases such as the latter, the text explaining the figure does not state the percentages or frequency.
Rating:	2: Some figures, tables and conventions are used in presentation of data but not entirely appropriately or consistently

Robustness of findings

Standard:	Data analysis appears to have been executed to an adequate standard
Comment and Analysis:	The analysis is of an adequate standard. The quantitative data adequately describes the programme components e.g. mentor and mentee profiles and contact session attendance. There is scope for the analysis to link back to programme design and address aspects of the theory of change. E.g. figure 13 is pie chart of mentee core needs and the results suggest that priorities emerging from the document review do not align with mentees self-reported needs. The analysis of this data fails to comment on the fact that this finding is despite an initial session held with prospective mentees prior to programme conceptualisation. The qualitative data analysis adequately captures the process and experiences of various stakeholders, however limited synthesis of the data takes place in the report. In some cases a single quote from a respondent is used to comment on the views of all respondents without stating explicitly that the quote captures the sentiment held by a majority of staff interviewed.
Rating:	3: Data analysis appears to have been executed to an adequate standard for most datasets

Standard:	Findings are supported by evidence which is sufficiently and appropriately analysed to support the argument, integrating sources of data
Comment and Analysis:	The findings are the product of a triangulation process from all the data and literature reviewed and are presented as themes based on the evaluation questions. These are adequately supported by the evidence presented in preceding chapters, however the synthesis of the evidence is not wholly used to answer all aspects of the evaluation questions and sub-questions provided in the analytical framework. The theory of change was also a key deliverable of the evaluation. It's adequately supported by evidence, however it is limited in mapping out the different causal pathways inherent in the programme. All interventions are combined and it would have been beneficial to map out different causal pathways for the different components of the programme.
Rating:	3: The evidence gathered is analysed to support the argument to an adequate standard and integrates sources of data

Standard:	There is appropriate recognition and exploration of the possibility of alternative interpretations
Comment and Analysis:	There is limited recognition of alternative interpretations in the report itself, however the interviews with both the programme manager and evaluator suggest that findings were extensively discussed and alternative interpretations from other stakeholders were considered.
Rating:	3: There is appropriate recognition of the possibility of alternative interpretations

Standard:	The report appears free of significant methodological and analytic flaws
Comment and Analysis:	The report documents all the methodological components and approach followed and appears to be free of significant methodological and analytical flaws. However, the reporting of sample size for each of the stakeholders interviewed is inconsistent. It is not clear how many programme staff participated in the focus group and the number of mentors that completed the email survey does not add up to the different categories (participated through out, dropped out and never participated). Moreover, the report states that "the mentee sample (78) of individuals who had been part of the YWDP pilot phase was deemed representative, as it is about 37% of the total mentees reflected on the YWDP attendance records." This suggest an inaccurate understanding of a representative sample which rather refers to all potential participants having an equal probability of participating in the focus groups.
Rating:	2: The report appears to include some minor methodological and analytic flaws, but these are not significant

Standard:	Limitations of all aspects of the methodology and findings are clearly articulated (e.g. limitations of scope or evaluation design, recommendation for additional research, data collection challenges, etc)
Comment and Analysis:	The discussion of limitations presents a list of challenges experienced during the study and states that these prohibited the development of an accurate baseline for the pilot phase. This section could have elaborated more on particular aspects of the findings that are compromised by the limitations. E.g. the low response rate from mentors means that the findings presented are not necessarily representative of the full cohort of mentors in the programme.
Rating:	2: There is some acknowledgment of the limitations of the methodology and findngs but these are not clear or exhaustive

Strength of conclusions

Standard:	Conclusions are derived from evidence
Comment and Analysis:	The section of the report for conclusions is a short paragraph of three sentences providing an overall perspective on the programme. It appears the findings and recommendations section draws the main conclusions with key evaluation questions used as themes. Based on these, the conclusions appear to be adequately derived from evidence, however this section is limited in its synthesis of the various sources of data to present an argument or position for each theme.
Rating:	3: Conclusions are derived from evidence

Standard:	Conclusions address the original evaluation purpose and questions
Comment and Analysis:	The section of the report for conclusions is a short paragraph of three sentences providing an overall perspective on the programme. It appears the findings and recommendations section draws the main conclusions with key evaluation questions used as themes. This choice of structure allows for the original questions to be addressed, however most of the sections only address the evaluation questions broadly and more detail could be provided to answer the sub-questions as well.
Rating:	3: Conclusions adequately address the original evaluation purpose and questions

Standard:	Conclusions are drawn with explicit reference to the intervention logic or theory of change
Comment and Analysis:	The section of the report for conclusions is a short paragraph of three sentences providing an overall perspective on the programme. It appears the findings and recommendations section draws the main conclusions with key evaluation questions used as themes. Based on these, there is some evidence that the conclusions are drawn from the theory of change. The section does not explicitly refer to the theory of change, however findings address the validity of some of the mechanism between the outputs and the outcomes. E.g. "While the quality of mentors involved in the pilot phase of the YWDP was high in terms of rank, credentials and experience and they were from all sectors, many mentors did not live up to expectations." This finding disproves one of the outputs listed in the theory of change that the intervention would result in exposure to enriching experiences and positive role models for participating mentees.
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 discussed with all relevant stakeholders. The evaluator presented an initial set of recommendations which were discussed and debated by a representative steering committee inclusive of beneficiaries. While the steering committee was well represented, consultation on the recommendations did not include stakeholders outside the steering committee.
Rating:	3: Recommendations are made in consultation with relevant government officials, stakeholders and sectoral experts

Standard:	Recommendations are useful- they are relevant, specific, feasible, affordable and acceptable
Comment and Analysis:	The report presents twelve recommendations which are well formulated and actionable. Each of the recommendations is accompanied by detailed action plan to guide the commissioning organisation in terms of implementation. The programme manager noted that the recommendations were useful. These final recommendations were acceptable to stakeholders. The evaluator noted that there was initially a bone of contention around the finding that the programme was not properly conceptualised in terms of targeting, however the Office of the Premier did acknowledge that it could have been done better.
Rating:	4: Recommendations are well-formulated for use- they begin to differentiate by user and are relevant to the current policy context, specifically targetted, feasible to implement, affordable and acceptable to key stakeholders

Acknowledgement of ethical considerations

Standard:	The full report documents procedures intended to ensure confidentiality and to secure informed consent where necessary (in some cases this is not needed - e.g. evaluation synthesis - in which case N/A should be recorded)
Comment and Analysis:	The report explicitly discusses the ethical principles that were applied in the data collection in section 1.7.4.3 and attaches the informed consent forms in the appendices. There is no discussion of how the data will be stored in future.
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:	The report itself adheres to ethical standards and presents no risk to participants and institutions. Additionally, the programme manager indicated that while there are no specific risks to participants or institutions, however the report itself needs to be improved in terms of accessibility prior to dissemination.
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 exceeded the planned timeframe of 7 months stated in the inception report and instead was completed over a period of 9 months while remaining within budget. The delays were due to a few factors: - delays due to an extensive feedback loop on all deliverables submitted and incorporation of all stakeholder inputs - data collection resulted in some delays due to challenges with language and non-programme participants arriving to the sessions. The data collection period had to be extended as a result. All time extension of the evaluation were approved by the commissioning organisation.
Rating:	2: The evaluation was completed outside of the planned timeframes and over budget, but with approval of the commissioning organisation

Evaluation use

Standard:	Results of the evaluation have been presented to relevant stakeholders
Comment and Analysis:	The results were presented to the steering committee and no other stakeholders beyond this group. The programme manager noted that this was not part of the ToR.
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:	The project close out session included a reflective aspect where stakeholder discussed what went well and areas of improvement for the evaluation and the programme itself. The reflections included: - improvements to the procurement process and timeframes are required - the internal project team should critically assess participants in the steering committee and ensure they play their role - acknowledgement that the programme could have been executed better
Rating:	3: A reflective process has been undertaken by the steering committee with the service provider to reflect on what could be done to strengthen future evaluations

Standard:	The evaluation study is seen by interviewed stakeholders as having added significant symbolic value to the policy or programme (e.g. raised its profile)
Comment and Analysis:	The evaluation is seen as having added symbolic value for stakeholders that participated in the steering committee and staff working on the programme through validation of the programme theory and emphasising the gender focal points. The symbolic value is, however, not noticeable outside of this group.
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 programme manager noted that the evaluation was of significant conceptual value and shared that a highlight was that "it gave us a frame in terms of how a programme can be implemented in a more systematic manner". Moreover, the experience from the programme and the evaluation has informed the design of the recently launched Tshepo 1 million programme.
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

Unleash Potential Unlimited. 2015. Final report: formative evaluation of the Gauteng Young Women Development Programme

Unleash Potential Unlimited. 2014. Inception report: Gauteng Young Women Development Programme version 3

Gauteng Office of the Premier. 2013. Evaluation of the Gauteng's Young Women Development Programme. Draft Terms of Reference

List of Interviewees

Annette Griesel, Former head of Head of Strategic Planning, Monitoring and Evaluation, Gauteng Office of the Premier (Declined to comment electronically)

Dr Octavia Mkhabela, Managing Director, Unleash Potential Unlimited, Telephonic interview. 8 December 2017

Wowo Mbuli, Programme Director, Gauteng Office of the Premier. Telephonic interview. 22 January 2018