



planning, monitoring & evaluation

Department:
Planning, Monitoring and Evaluation
REPUBLIC OF SOUTH AFRICA

Report on the Assessment of Government Evaluations

Evaluation Title:	Economic Evaluation on the Incremental Investment in Forensic Services
Evaluation Number:	602
Start Date:	Monday, 17 November 2014
Completion Date:	Friday, 17 February 2017
Period of Evaluation:	2 years and 3 months
Submitted:	Friday, 08 December 2017 by Cara Hartley
Approved:	Friday, 08 December 2017 by Cara Hartley

Evaluation Details

Evaluation Title:	Economic Evaluation on the Incremental Investment in Forensic Services
Evaluation Number:	602
Evaluation Completion Date:	Friday, 17 February 2017
Created:	Friday, 10 November 2017 by Mike Leslie
Submitted:	Friday, 08 December 2017 by Cara Hartley
Approved:	Friday, 08 December 2017 by Cara Hartley
Period of Evaluation:	2 years and 3 months
Known Cost:	R 2 449 248,00
Known Cost Type:	Referenced
Initiated By:	Department of Planning, Monitoring and Evaluation & South African Police Service
Initiated By Internal:	Yes
Service Providers:	DNA Economics
Undertaken By Internal:	No

Assessors

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

Assessment Documents

Document Name:	Document Type:	Added By:	Added On:
SAPS Draft TOR 28 Feb 2014 Final JB (2).pdf	Terms of Reference (ToR) for the evaluation	Cara Hartley	Tuesday, 05 December 2017
SAPS Inception Report v3.0 250315 DP.docx	An inception report	Cara Hartley	Tuesday, 05 December 2017
SAPS documentary review 050615 STC literature review.docx	Any other relevant documentation pertaining to the evaluation process	Cara Hartley	Tuesday, 05 December 2017
DPME256 Fieldwork report V01 11082015 STC.docx	Progress reports	Cara Hartley	Tuesday, 05 December 2017
SAPS Minutes of Inception Meeting 14 11 24.docx	Minutes or notes of meetings with stakeholders	Cara Hartley	Tuesday, 05 December 2017
Minutes of Forensic Evaluation Validation workshop 16.09.23.docx	Minutes or notes of meetings with stakeholders	Cara Hartley	Tuesday, 05 December 2017
DPME 256 IIP Draft Evaluation Report AJ 20160808.docx	Any other relevant documentation pertaining to the evaluation process	Cara Hartley	Tuesday, 05 December 2017
Evaluation of IIFS Full Report Final.docx	Evaluation report	Cara Hartley	Tuesday, 05 December 2017
Evaluation of IIFS Statistical Appendices 17 02 17.docx	Evaluation report	Cara Hartley	Tuesday, 05 December 2017
Evaluation Framework V05 DP 22062015.docx	Assessment tools	Cara Hartley	Tuesday, 05 December 2017

DPME SAPS Interview guide DP v02 25052015.docx	Assessment tools	Cara Hartley	Tuesday, 05 December 2017
Evaluation of IIFS Summary Report Final.docx	Evaluation report	Cara Hartley	Tuesday, 05 December 2017
Assessment Report at Moderation.pdf	Assessment Report at Moderation	Mike Leslie	Tuesday, 05 December 2017

Quality Assessment Summary

This evaluation of the Incremental Investment in Forensic Services (IIFS) is rated at 3.54 out of 5 overall, indicating an evaluation of good quality. It is particularly strong in terms of relevant stakeholders' engagement throughout the process, the quality of the methodology and report, and its relevance and value to the policy and programme context.

The evaluation planning and design took place in consultation with a range of stakeholders, especially those in the South Africa Police Service (SAPS) who are closest to IIFS. With the Department of Planning, Monitoring and Evaluation (DPME) involved from the outset to guide the process and share evaluation expertise, the evaluation was fairly well conceptualised at this early stage in terms of evaluation type (economic evaluation with a strong implementation evaluation component), methodology, and intended users. The service provider was deemed exceptionally well suited to the evaluation and the inception phase was used to good effect to achieve a common understanding of how the evaluation would be implemented. One element that was not well planned for in this evaluation is resourcing, in terms of time and budget. Instead of the originally planned 9 months, the evaluation took over two years to complete, with several arguably unforeseeable challenges driving significant delays. This placed pressure on the evaluation budget as well. It was also noted in retrospect that a more detailed discussion by all stakeholders of the nature and type of quantitative data that the service provider was going to need from SAPS, including the level of granularity and reliability required, would have set things up even better for a smooth evaluation process. Despite this, in light of the many strengths of the way the evaluation was planned and designed, a score of 3.49 is assigned to the Planning and Design phase.

Despite significant delays and data challenges, stakeholders worked together to implement the evaluation to a reasonably good standard, with the evaluation Implementation phase rated at 3.48. Strengths of this phase included the quality of the literature review; the revisions to the methodology to address data limitations and gaps; and broad coverage of the important role players in the IIFS through multiple types of secondary and primary data. The steering committee had notable strengths, including active engagement with the evaluation's interim deliverables and flexibility to devise solutions to challenges, but also notable challenges, mostly attributable to SAPS internal challenges which caused considerable turnover of the SAPS committee members (among others, the chairperson changed twice). Also linked to internal tensions in SAPS, there is some evidence of attempts to prevent certain negative findings from being reported (although this was rebuffed), and some unwillingness to share data in a timely manner.

The Reporting phase was particularly strong in this evaluation, scoring 3.66 out of 5. The final evaluation report is considered by all to be of good quality, benefiting from good analysis of data from multiple sources, validation of findings by individual units in SAPS through workshops, inputs from a peer reviewer, and several rounds of in-depth inputs from the steering committee, thereby generating findings which served as a good basis for answering the evaluation questions and drawing strong, convincing conclusions. While recognising the documented limitations of the data and scope, the report can therefore be seen as a quite reliable reflection of what has happened in terms of Incremental Investment in Forensic Services in South Africa up to 2015. Recommendations benefited from consultation with key stakeholders and were recently validated as acceptable during the process of developing an improvement plan, where the group reportedly disagreed with only one of the 23 recommendations. One weakness of the evaluation report (but not the evaluation process) is that ethical considerations are not described in sufficient detail.

In terms of Follow-up, Use and Learning, the evaluation scored 3.40. This somewhat lower (although more than adequate) rating was impacted by the fact that the evaluation was completed so far outside of the intended timeframe and budget. Furthermore, there was little opportunity for reflection among stakeholders about lessons learned through the experience - this is why the overarching consideration Capacity Development is also rated as inadequate, at 2.60. There are nevertheless indications of greater understanding of and appreciation for evaluations in SAPS.

Overall, this evaluation is deemed by all interviewees as being of great conceptual value, bringing greater understanding of what has happened; and symbolic value, sending a strong message about the value of forensics and incremental investment in it, as well as the importance of addressing the challenges to its efficiency and effectiveness. There appears to be agreement with the conclusions and support for the recommendations. Interviewees are confident that it will shape policy and practice; in fact it has already reportedly informed the development of the National Forensic Strategy.

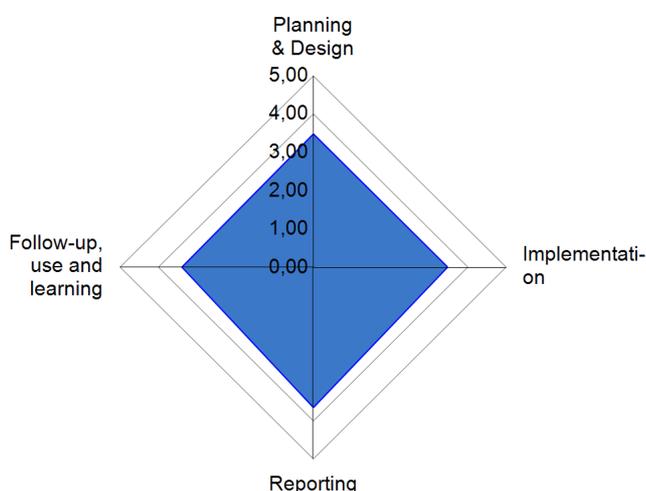
Quality Assessment Scores

Phase of Evaluation	Score
Planning & Design	3,49
Implementation	3,48
Reporting	3,66
Follow-up, use and learning	3,40

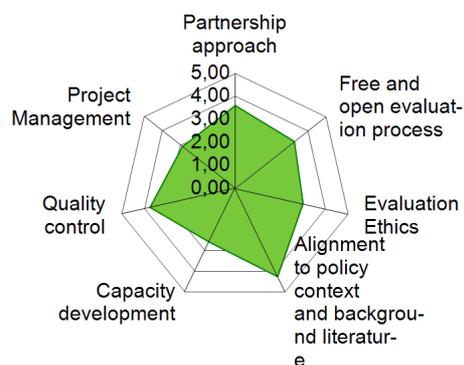
Total	3,54
-------	------

Overarching Consideration	Score
Partnership approach	3,60
Free and open evaluation process	3,28
Evaluation Ethics	3,00
Alignment to policy context and background literature	4,27
Capacity development	2,60
Quality control	3,76
Project Management	2,90
Total	3,54

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	3,71
Planning & Design	Appropriateness of the evaluation design and methodology	3,36
Planning & Design	Project management (Planning phase)	3,00
Implementation	Evaluation ethics and independence	3,00
Implementation	Participation and M&E skills development	3,57
Implementation	Methodological integrity	3,79
Implementation	Project management (Implementation phase)	3,00
Reporting	Completeness of the evaluation report	3,50
Reporting	Accessibility of content	4,00
Reporting	Robustness of findings	3,65
Reporting	Strength of conclusions	4,00
Reporting	Suitability of recommendations	4,00
Reporting	Acknowledgement of ethical considerations	2,57
Follow-up, use and learning	Resource utilisation	2,00

Phase of Evaluation	Area of Evaluation	Score
Follow-up, use and learning	Evaluation use	3,75
Total	Total	3,54

Planning & Design

Quality of the TOR

Standard: The evaluation was guided by a well-structured and complete TOR or a well-structured and complete internal evaluation proposal (e.g. Background, Purpose, Evaluation Questions, Design & Methodology, Deliverables & Timeframes, Resource requirements, Intended Audience & Utilisation, etc).

Comment and Analysis: The TOR for this evaluation was well-structured and covered all of the following: Background, evaluation questions, design, methodology, deliverables, time frames, resource requirements (in terms of the required competencies and skills required of the service provider), intended audience and utilisation.

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 first stated purpose of the evaluation was to determine whether the benefits (outcomes) of the annual incremental investment into the SAPS Forensic Services (IIFS) outweigh the costs (inputs), or not. This type of purpose is best served by an Economic Evaluation. The second stated purpose was to provide useful evidence on the implementation of IIFS and how its effectiveness can be optimised. This type of purpose is best served by an Implementation Evaluation. It was thus appropriate that the required type of evaluation was "Economic and Implementation". The scope in terms of time period aligned with the inception of the IIFS up to the last financial year for which data would be available at the commencement of the evaluation, which was also appropriate.

The only possible concern that can be raised is that an Economic Evaluation usually requires the measurement of outcomes, but this was not clearly provided for in the statement of the scope of themes and components to be covered. In this sense, the evaluation was well conceptualised, but not ideally so.

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 included a table listing the potential evaluation users and how they would use the evaluation (action they may take as a result of the evaluation). From how they would use the evaluation, their information needs can to an extent be inferred. This is considered adequate in terms of this standard; it would have been good to give more detail on the type of information that would best facilitate their respective intended uses of the evaluation.

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:	According to the key stakeholder from SAPS who was interviewed, the process of developing the TOR was extensive, including multiple consultations between DPME and senior officials responsible for forensics, including the divisional commissioner for forensics, with various divisions in SAPS. Through this, stakeholders in DPME and SAPS took the first steps in clarifying the intervention logic, and they came to agreement both about what the evaluation purpose would be, and also what purposes it would not serve. With the extensive involvement of these key stakeholders closest to the programme that was evaluated, the process is therefore deemed to have adequately involved the key stakeholders although some broader stakeholders (National Prosecuting Authority (NPA); Treasury) only became involved later on.
Rating:	3: Key stakeholders were 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 evaluation was inadequately resourced, especially in terms of time. Only 9 months were allocated for an evaluation with substantial qualitative as well as quantitative components in a government entity for whom this was the first ever evaluation. The methodology (see below) depended on the use of data, and engagements with respondents, that were assumed to be readily available, but this was not the case, because of (1) fragmented and unclear management practices around the resource and expenditure data that was essential to the type of evaluation asked for; and (2) the extreme sensitivity of the information that needed to be analysed, which required a long process to come to agreement about granting of satisfactory levels of clearance to the analysts, and permission for the relevant divisions to share it. (3) It also takes time for the necessary communication among different divisions in SAPS to ensure the granting of permission to participate in research interviews (according to the service provider, more so than in other government entities) - and extensive interviews were planned. These issues can all reasonably be expected to have been identified at the conceptualisation or inception stage, and a more realistic time frame allocated. Likewise, the budget impacts of these time-consuming processes were not anticipated nor budgeted for.
Rating:	2: The evaluation was resourced with tight timeframes and budget which were challenging from the outset
Standard:	The team conducting the evaluation was adequately resourced in terms of staffing and skills sets
Comment and Analysis:	The evaluation required a team that was familiar with government evaluation, quantitative as well as qualitative analysis, the criminal justice system, and forensics in particular. All sources indicate that the team's expertise thoroughly covered all of these areas. One team member also combined about 15 years' forensics experience in South Africa with over 10 years' experience researching and teaching on various aspects of forensics at an American university, where he is currently a professor. The SAPS interviewee and DPME interviewee both commented on the team's quality of expertise and high level of engagement with the evaluation. The team was also large enough to conduct the research and to analyse large datasets within reasonable time frames, i.e. staffing constraints were not mentioned as a factor driving the delays in this project.
Rating:	5: The staffing and skills sets required for the evaluation were ideal for the evaluation purpose, sector and incorporated high quality international expertise

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:	<p>As is typical for National Evaluation Plan evaluations, the TOR stated that a theory of change would need to be developed on early in the programme and tested in the evaluation. The inception report then provided detail on when and how the theory of change would be refined, workshopped and used to guide the evaluation, and that a final theory of change would be recommended as part of the project. Thus the evaluation was planned with the intention that the intervention logic would be central to the direction of inquiry.</p> <p>However, the evaluation reportedly had only a draft, high level theory of change when the evaluation started, and the Inception Report did not go so far as to present a visual representation of it. In practice, the service provider found that the client had not yet clearly defined and delineated the intervention (IIFS); as a result, significant further work was done to clarify it with stakeholders in the early stages of the evaluation.</p>
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 evaluation questions asked for a determination of the benefits of IIFS, its cost-effectiveness, what is working and not working regarding its implementation, the implications of institutional arrangements on its performance, the implications of the evaluation for the IIFS design, and how IIFS can be made more effective. These questions require an in-depth understanding of what IIFS has consisted of; how much was invested in forensics, when and how; how it has been implemented throughout SAPS and in relation to other entities; and its outcomes / effects / benefits for the criminal justice system. Accordingly, a comprehensive methodology was planned which included interviews with officials in all main SAPS divisions and other entities, detailed analysis of expenditure data, analysis of several types of routine reports (e.g. staffing reports, supply chain management reports) and analysis of criminal justice system performance reports. A cost-benefit analysis was planned to address the evaluation questions in relation to benefits and cost-effectiveness.</p> <p>Thus, the planned methodology was well-suited to the questions being asked. Still, the inception report did not yet reflect a clear recognition of the limitations of the data available, especially on expenditure and human resourcing. Stakeholders only came to grips with these limitations later. As pointed out earlier, the planned methodology made assumptions about the data that may have been dispelled if an evaluability assessment had been undertaken.</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>The planned methodology involved planned sampling in terms of the selection of respondents to be interviewed. The list of respondents was discussed and refined in the inception phase, and the agreed sample was comprehensive. It covered 38 national respondents, including SAPS senior leadership, officials in the most relevant divisions, and other relevant entities (the Civilian Secretariat, NPA, Public Works, State Information Technology Agency (SITA), the Department of Justice and Constitutional Development (DOJ&CD) and the Integrated Justice System board). Furthermore it included a total of 48 respondents in four provinces. The four provinces were well selected to reflect the variety of urban/rural and socio-cultural contexts within which forensics is used in South Africa. The fieldwork report also shows that careful planning went into sequencing the interviews with the different types of respondents, to serve different purposes. The SAPS interviewee expressed satisfaction but felt that a more representative selection of detectives could nevertheless have been obtained.</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>The inception phase was prolonged because the SAPS senior leadership requested in the first few months to change the chairperson of, and the composition of the evaluation steering committee. This was reflective of internal issues at SAPS at the time. The DPME and other SAPS officials on the steering committee worked to ensure that the new members and chair were thoroughly briefed on the evaluation and that the steering committee was quorate before moving on to finalise the inception report. In this way, although it delayed the inception phase from the planned 1 month up to 6 months, it was ensured that the steering committee members were on the same page about the evaluation and how it would be implemented, by the time the inception plan was approved. This included common agreement on the interviews to be conducted nationally and in four provinces.</p> <p>A more detailed discussion by all the stakeholders involved of not just the qualitative data collection, but also the nature and type of quantitative data that the service provider was going to need from SAPS, and what the implications of that was beyond just security clearance (e.g. level of granularity required), would have set things up even better for a smooth evaluation project.</p>
Rating:	3: The inception phase was used to develop a common agreement on how the evaluation would be implemented

Implementation

Evaluation ethics and independence

Standard: Where data was gathered in contexts where ethical sensitivity is high, informed consent, assurances of confidentiality and appropriate clearance were achieved; e.g. through an ethics review board, in evaluation involving minors, institutions where access usually requires ethical or bureaucratic clearance

Comment and Analysis: Ethical sensitivity was high in this evaluation, because individual case records in the Crime Administration System (CAS) database were analysed. Sensitive data about the functioning and capacity of SAPS was also handled, which could hypothetically be used for ill. Therefore, the need to facilitate security clearance for members of the evaluation team was recognised early on and SAPS committed to making office space available for data that could not be taken off-site. According to the interviewees for this quality assessment, it was later discussed that the security clearance process could take up to 6 months and therefore in the meantime the team also signed confidentiality agreements. The DPME interviewee indicated that there was discussion with SAPS about the further assurance of confidentiality provided by certain clauses in the Service Level Agreement (SLA) that the service provider had signed. The necessary clearances and office space were made available (although not before considerable delays). According to the interviewee from SAPS, these measures satisfied the steering committee's concerns about confidentiality.

The service provider indicated in the interview for this assessment that the interview respondents were assured of confidentiality with a standard script read to them at the start of the interview, and, unlike in most evaluations of this nature, it was agreed that interviews would not be audio recorded, to further protect the respondents. They were anonymised in the final report.

Confidentiality considerations also fed into the decision to keep the online surveys anonymous (see discussion of the surveys below).

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

Standard: Where external, the evaluation team was able to work without significant interference and given access to existing data and information sources

Comment and Analysis: The evaluation faced challenges with regards to the evaluation team's ability to work without interference and being given access to data. All those interviewed for this evaluation acknowledged that there were attempts to prevent certain negative findings from being reported, specifically the widespread allegations of procurement irregularities. It is noted that DPME supported the service provider in insisting that the negative findings, which were corroborated across multiple data sources, remain in the report and that recommendations be made with regard to them. Because of this, all interviewed stakeholders indicated that the final evaluation report was ultimately not biased by these attempts to interfere.

The service provider also expressed the view that repeated delays in granting permission and providing the requested information were not just a factor of the necessary processes in a hierarchical organisation, but also of some officials' unwillingness to cooperate. (It later became known that a forensic audit had recently been conducted involving some of the same officials who were now being asked to share data or give interviews for the evaluation; this naturally raised suspicion about the purpose of the evaluation.) The persistence of the steering committee and service provider is however believed to have finally yielded most of the necessary data, albeit after considerable delay.

Rating: 2: There was some evidence that the evaluation team was not supported to get access to existing data and information sources

Participation and M&E skills development

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

Comment and Analysis: A steering committee oversaw the evaluation (see list of members on page 4 of the report). It consisted of the DPME, NPA and various relevant divisions in SAPS. The steering committee was reported by interviewees to have been satisfactorily involved, albeit with some "push and pull" among some SAPS members to ensure active participation. Although not listed as members, the DPME interviewee indicated that Treasury (relevant because of its role in making substantial incremental investment into forensic services) and SITA (relevant because of its role in criminal system data management, and procurement) joined the steering committee towards the end and then gave valuable comments. However, the regular, active involvement of the NPA and the various SAPS divisions was the most essential consideration, since they are closest to the intervention.

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: A DPME evaluation officer participated in some evaluation activities with evaluation team. This was guided by a structured capacity building plan. She reportedly did learn and benefit from this. Unfortunately her time was limited, so that she did not participate as intensively as intended throughout the project.

The DPME interviewee said that the above-mentioned DPME evaluation officer, as well as the SAPS members of the steering committee, also benefited from the "demonstration effect", of seeing first-hand how evaluations work. This included, according to the service provider interviewee, notable growth in the steering committee members' understanding of the relevance of theories of change. The SAPS interviewee also described growth in the SAPS's members thinking around activities, outputs and outcomes, and the importance of keeping intended outcomes in mind while being busy with the daily activities of forensic services. This capacity development was not structured however.

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 evaluation benefited from a 65-page review report covering the South African policy context and recent developments; legislative review; international comparative review; and analysis of key IIFS expenditure and performance reports. It concluded by stating, on the basis of the review, what key topics that should be investigated (further) in the evaluation; and main categories of expenditure to consider in the analysis. It also put forward a more detailed list of questions in relation to the key evaluation areas of inquiry, based on what had been learned about forensics and the IIFS in the literature review. The service provider stated that some of the work done in this report on describing SAPS's own organisational structure was helpful to the steering committee. The final report also notes key lessons from the literature for the theory of change. In this way the literature review appears to have been insightful in terms of the analytical framework and provided a 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:	<p>The methods employed in the process of the evaluation diverged somewhat from what was planned:</p> <ul style="list-style-type: none"> - The main change was in the methodology employed towards the assessment of costs and benefits. The intention had been to do a full cost-benefit analysis, but as the final report notes, "the lack of detailed and granular expenditure, performance, and administrative data made this impossible... Since the level of detail required was not available, the evaluators adopted a "soft" cost benefit analysis. The methodology combines information from the FSL admin database, CAS and CRIM databases to trace the outcomes of cases from the laboratories to the court over a ten-year period." Although inconsistent with what was planned, the decision to respond to the unavailability of data by changing the methodology in such a way that it nevertheless yielded insight about the benefits of increased investment on the effectiveness of the CJS, should be commended. The report notes that this may be the first time such an analysis has been done in South Africa. - Furthermore, it was decided after the inception stage had been concluded, to extend the methodology by conducting electronic surveys: one of crime scene investigators; and one of prosecutors (purposive sample selected with assistance of the NPA). Again this is a deviation that added to the value of the evaluation, according to the service provider this was added "to bring out the voices of the users of the equipment and consumables purchased through IIFS". (A third survey, of forensic analysts, was also proposed but did not materialise in the end because there was no electronic platform through which they could participate.) - Instead of just interviews, the decision was made to conduct focus groups with experienced detectives in the four provinces. This was done according to the fieldwork report to augment the interviews with senior managers in detective services, and group settings were considered more likely to set them at ease to speak freely. - The interviews were broadly conducted as planned. <p>Because the deviations mostly improved the scope and data available to the evaluation, the evaluation is rated 4 (good) on this standard.</p>
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:	The service provider in the interview stated that the evaluation team did three pilots for the forensics interview guides, conducting interviews with supervisor staff who had previously held the positions of the officials who were going to be interviewed. After each pilot, the team reportedly did an iteration of comments on the instrument, until after the third pilot it was finalised. The pilot results were not used in the final evaluation. This is an instance of good practice that is rare in qualitative data collection and seems to have benefited the quality of the instruments. The service provider said that all the interview instruments were piloted in this way, except those that were developed for a single person (where it was hard to find someone on whom to test the instrument.) The surveys were developed in consultation with SAPS staff, but they were not piloted. Since the surveys were also important data points, it would have been good also to pilot them. Since not all instruments were piloted but the interview instruments were piloted well, a rating of 3 ("adequate") is assigned.
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 discussed earlier, a wide range of the key stakeholders were involved through interviews and surveys. The interviewed stakeholders had no concerns about any key stakeholder groups left out of the data collection. The online surveys meant that more key stakeholders were able to participate than had been planned. The only small critique was the SAPS interviewee who felt the selection of detectives for interviews could have been more representative. On this basis the coverage of key stakeholders is rated as "good".
Rating:	4: Data was collected from the intended key stakeholder groupings in line with the envisioned range and type of stakeholders (approx. 80-89% of intended)

Standard:	The methodology included engaging beneficiaries appropriately as a key source of data and information
Comment and Analysis:	The ultimate beneficiaries of high quality forensic services are the those affected by crime in South Africa. However, at a more immediate level, an objective of IIFS was "to increase the availability of high quality and timely forensic evidence for detectives and prosecutors". If one considers these the beneficiaries, then it is relevant that (1) detectives were interviewed, and their behaviour in terms of using forensic evidence in their work was also analysed quantitatively. (2) 41 prosecutors were surveyed as well. This can be described as "meaningful" engagement of these beneficiaries as a primary source of data and information, even though the report notes that the samples were not randomly drawn and therefore not statistically representative of the broader programme. The interviewed stakeholders were all satisfied with the level of inclusion of these "intermediate" beneficiaries as a key source of data and information.
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>The steering committee and the service provider worked together adequately to support the evaluation objectives. However, the significant delays are to a large extent attributable to repeated changes in the SAPS members of the steering committee (not just the change that was described above in relation to the inception phase, but also two more changes of the steering committee chair). These changes were driven by turnover and internal issues within SAPS, and in the interviews the officials in DPME and SAPS described the work they did to meet with and bring new members up to speed as best as possible, but there was a period when certain SAPS members were reportedly uncooperative (e.g. staying away from meetings), and moreover the multiple changes influenced the ability of the steering committee to work efficiently towards the achievement of the evaluation objectives.</p> <p>On the other hand, the evaluation benefited from the willingness and constructive efforts of the service provider and most steering committee members while they served as such. As described above, there were instances of creative problem solving around the unavailability of data and a flexibility that allowed for enhancements in the data collection.</p> <p>The rating of 3 is assigned to reflect both the significant problems that the steering committee posed for the evaluation, as well as the notable constructive efforts nevertheless made towards the achievement of the evaluation objectives.</p>
Rating:	3: The steering committee, technical working group and service provider worked together adequately to facilitate achievement of the objectives of the evaluation
Standard:	Support provided by the evaluation secretariat (e.g. the administrators responsible for the evaluation) facilitated achievement of the objectives of the evaluation (eg turnaround times, addressing problems, preparation for meetings etc)
Comment and Analysis:	The secretariat function was played by DPME. It was described by the interviewees as adequate under the circumstances.
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:	<p>Interviewed stakeholders expressed satisfaction with the quality of the draft report and deemed it sufficient quality to go out to stakeholders. The draft report is detailed and clear and does not appear "rough" (e.g. a quick scan does not reveal any layout, formatting or typing errors). The core findings and much of the textual discussion, graphs and tables remained unchanged in the final draft. Overall it appears to be a good draft report. Some sections of the final report were restructured, phrasing was revised, and there were substantial changes in the conclusions (suggesting further refinement of what the findings mean for the programme), and some additional findings from data received from SAPS at a late stage - but arguably these were enhancements and do not reflect negatively on the draft report.</p> <p>Perhaps contributing to the quality of the draft report is the fact that the service provider had "clarification sessions" with each of the relevant SAPS divisions prior to producing the draft report. According to the interviewed member of the service provider team, this served almost as a "pre-validation" process - it no doubt contributed to the evaluation team's clarity on the issues already in the first official draft.</p>
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:	<p>The final evaluation report is well-structured and complete in almost all respects. It includes an executive summary, context, purpose, questions, methodology, findings (including analysis), conclusions and recommendations. The evaluation scope is stated although it is not a standalone section. The methodology section is not well presented as it is very cursory; it does not describe the different data collection methods in sufficient detail to introduce the (good) evidence base on which the evaluation rests. The surveys are not mentioned at all until they are later introduced when their findings are presented. Focus groups were not mentioned in the report at all, leaving it unclear how the data collection from them was used.</p>
Rating:	3: The final evaluation report is complete, follows a clear structure and addresses at minimum: 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:	<p>The final evaluation report is user-friendly and easy to navigate. It is clearly written with only negligible typing and formatting errors. The interviewed DPME and SAPS officials expressed pride and satisfaction with the final product and report that those to whom the results have subsequently been presented, also expressed a high opinion of it.</p>
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:	The evaluation is strong in its incorporation of relevant tables, figures, boxes and quotes. It presents different types of data in an integrated manner that holds the reader's attention. Most graphs include legends, but where they do not, the accompanying text clarifies the different elements on the graph sufficiently. The regression used to test whether IIFS had a statistically significant effect on justice outcomes, describes its methodology and its main limitations and is clear about what results are statistically significant or not. Even though statistical significance is indicated with "yes" or "no" next to the results, the p-values are not stated in the presentation of the results (they could perhaps have been added in an Appendix); however, since this regression and its results were not ultimately central to the findings and the omission of p-values appears more a matter of style than inappropriate presentation, this is not taken as a significant issue with regards to this standard. Overall this report makes effective use of a variety of different visual representations of data.
Rating:	4: Figures, tables and conventions are well used for a variety of types of data presentations and supporting explanations make them accessible to readers

Robustness of findings

Standard:	Data analysis appears to have been executed to an adequate standard
Comment and Analysis:	The data analysis appears to have been executed well for all datasets. This is quite a feat considering the large and varied amount of data that was analysed. The regression appears to have been adequately done as well.
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:	The presented evidence is adequately explained, assumptions pointed out, and the report tends to make it clear what the evidence means for the findings in relation to the relevant area of inquiry. Data sources are well integrated, with findings emerging from the use of multiple sources of data where they were available. Where relevant, the report highlights data sources that contradict each other or clarify each other. Some synthesising or concluding sections could have helped to build a stronger narrative through the report, but there are introductory paragraphs and paragraphs reminding the reader of what has already been said, which help to build the narrative to an extent.
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: The report uses the multiple data sources to good effect to discuss possible reasons for findings (e.g. possible reasons for crime scene investigators' low morale are explored using data from interviews, surveys, and HR data on retention). The report also takes care, when presenting data, to note important points and caveats that have bearing on how the data should be interpreted, and sometimes explains why the reader should caution against a specific interpretation. In at least one place, the authors also explicitly mention an argument that key stakeholders made in response to the draft results, and discuss the validity of their arguments. The evaluation also points out which assumptions have and have not held according to the evaluation findings. Thus although there is no explicit discussion of broader alternative schools of thought or ideological approaches to the subject matter, it appears that the possibility of alternative interpretations is considered adequately in relation to the findings presented in the report. Reinforcing the impression of sufficient recognition of possible alternative explanations, the interviewees did not mention any major differences of opinion or alternative points of view that call the evaluation findings into question.

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: There are no obvious areas where analysis appears unclear, data appears inappropriately interpreted, or pieces of evidence assigned inordinate weight in the discussion. There are no significant apparent flaws. Furthermore, the report documents the fact that key stakeholders' inputs were obtained on the theory of change and reports.

Although not stated in the full report, the interviewees also noted that there was a peer reviewer who provided valuable comments, and that the steering committee members went through the deliverables in detail, including the data collection instruments and final report. The Forensics staff from SAPS, NPA and DPME were all cited as having provided valuable inputs on the drafts.

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: The evaluation notes limitations around some of the available case datasets; limitations of attribution (contribution to outcomes rather than attribution of impact); and (briefly) limitations of the survey samples. More could perhaps have been done to guide the reader in how to interpret the survey results, if not as entirely representative of the populations the samples were drawn from (e.g. the report notes that those without computer access could not participate in the electronic surveys; is this a systematic bias perhaps concentrated in specific provinces or particularly under-resourced units?).

In addition to the explicit acknowledge of these three types of limitations up front, where the report presents new pieces of data in the findings section, it does so with explicit acknowledgement of limitations of the data. This includes recognition that the operationalisation of the DNA Act shortly after the period under review in this evaluation, will impact on certain types of resources and therefore render the evaluation findings in this regard outdated. The SAPS interviewee considers this an important limitation of the evaluation findings; it may have been useful to state this limitation upfront, and not only at the end of the relevant section. The report recommends one area for further "interrogation".

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:	The conclusions are presented after the findings and in relation to each evaluation criterion (as well as an additional one, "emerging impact", which was added since the analysis revealed important trends in this regard). The conclusions all point back directly to findings and evidence that were analysed and integrated in the previous sections. In this sense, the conclusions are well-founded and convincing. It is not clear what more could have been done to strengthen the conclusions.
Rating:	5: Conclusions are derived from evidence that has been triangulated and thoroughly analysed, limiting any arguments against the conclusions

Standard:	Conclusions address the original evaluation purpose and questions
Comment and Analysis:	The purpose of the evaluation was to determine whether the benefits (outcomes) of the annual incremental investment into the SAPS Forensic Services (IIFS) outweigh the costs (inputs), or not; and to provide useful evidence on the implementation of IIFS and how its effectiveness can be optimised. The evaluation has certainly succeeded in the latter. The former proved a hard purpose to achieve given data limitations, but the evaluation nevertheless gives a clear idea of what the costs / inputs and benefits / outcomes have been so far. It also highlights specific areas that are more and less efficient. The key evaluation questions are answered in the conclusions, and are aligned to the evaluation criteria. They are answered in a manner that appears comprehensive. The interviewed stakeholders also did not raise any objections to the conclusions that the evaluation came to.
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:	The first conclusion, on the relevance of IIFS, refers specifically to the IIFS intervention logic. The second conclusion, around efficiency, is structured according to the results chain articulated in the theory of change. Thus the intervention logic is referred to in these conclusions, and guides the conclusions adequately. The conclusions do not reflect specifically on the implications of the evaluation for what parts of the intervention logic / theory of change are valid or should be reconsidered, although this can be easily inferred. In all the conclusions deal adequately with the intervention logic.
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 substantially reworked between the draft and final reports, suggesting that they received considerable attention and inputs from stakeholders. This was confirmed in the interviews: According to the evaluation service provider, the evaluation team made draft recommendations on the basis of the findings; these were then presented to key stakeholders at the validation workshop, where each SAPS division validated the findings relevant to it, as well as recommendations. Participants in these workshops included representatives of the IIFS "direct beneficiaries" (detectives and prosecutors) if not representatives of the ultimate beneficiaries (victims of crime). The peer reviewer is a sectoral expert; this person commented on the draft recommendations along with the rest of the draft report. The SAPS interviewee expressed satisfaction that the stakeholders' inputs had resulted in improved recommendations.
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:	The SAPS interviewee indicated that they consider the final recommendations to be quite relevant - particularly in their final form, reflecting the inputs of the relevant stakeholders. To a person less familiar with the subject area, the recommendations appear clearly relevant (directly related to the evaluation findings, focusing on what can be improved); specific (there are 7 overarching recommendations with 23 sub-recommendations in total, each of which specifies what the action is and who should take it, sometimes also noting to what end/purpose); and acceptable (according to the SAPS and DPME interviewees who have been present when the recommendations were presented to broader sets of stakeholders). Specifically, the DPME interviewee reported that the broad collection of SAPS staff who participated in the Improvement Plan workshop (based on the recommendations), accepted all except one of the recommendations. The fact that stakeholders developed an Improvement Plan based on all but one of the recommendations suggests they are also sufficiently feasible and acceptable.
Rating:	4: Recommendations are well-formulated for use- they begin to differentiate by user and are relevant to the current policy context, specifically targetted, feasible to implement, affordable and acceptable to key stakeholders

Acknowledgement of ethical considerations

Standard:	The full report documents procedures intended to ensure confidentiality and to secure informed consent where necessary (in some cases this is not needed - e.g. evaluation synthesis - in which case N/A should be recorded)
Comment and Analysis:	<p>Although the interviewed stakeholders are all entirely satisfied with the procedures taken to ensure confidentiality and informed consent, and although this quality assessment also has rated the evaluation high in that regard, these aspects are only briefly documented in the full report. It states that during the inception phase, the steering committee discussed the sensitive nature of the data to be analysed, and it was agreed that SAPS would facilitate security clearance for the relevant researchers (it is not stated that this was indeed obtained, or whether it was obtained before, during or after the team's work with the data). The examples of research instruments provided in the Appendix also show that potential respondents would be given some background information about the study, which could serve as the basis for informed consent, although the intention to obtain their consent is not made explicit.</p> <p>Although numerous other measures were taken (as described earlier in this quality assessment), they are not documented in the full report or appendices. Given the sensitive nature of the study, it can be argued that the report should have been clearer on the full set of measures taken to ensure confidentiality and informed consent. This would not only give readers clarity about the ethical standards that stakeholders upheld, but also assist future studies in similarly sensitive areas.</p>
Rating:	2: The full report does not acknowledge whether confidentiality was ensured or informed consent secured but there is some evidence that this is the case
Standard:	There are no risks to participants or institutions in disseminating the evaluation report on a public website
Comment and Analysis:	There are no significant risks to publishing the summary report online. Although as noted, the evaluation process dealt adequately with sensitive information and the protection of respondents, DPME and SAPS are in agreement that the full report provides details on sensitive aspects of the functioning of the criminal justice system, which could strengthen the position of those seeking to undermine it. On this basis they deem it not in the public interest to publish the full report, but only a summary.
Rating:	3: There are no risks to participants or institutions in disseminating a summary version of the 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 went far over its original intended time frame of 9 months - from the first inception meeting (17 November 2014) to the submission of the final report (17 February 2017) was approximately 2 years and 3 months. The reasons for this have already been discussed above. A 15% extension had been agreed to at the inception phase in line with extensions to the methodology and scope, but there were no further budget extensions for the unanticipated delays, additional engagements, and significantly more effort and follow-up that was required to secure the agreed interviews and data.
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 have been quite extensively presented in various parts of government. They have widely presented in the many different divisions of SAPS: the SAPS senior management, the directorate: forensic services, detectives, supply chain management, and finance. Furthermore the results have been present to representatives of Treasury, the NPA, Correctional services and others in the criminal justice system. It was also presented to the justice, crime prevention and security cluster; and then to Cabinet in November 2017. The results have not been presented to non-government stakeholders (interest groups, media, think tanks etc), with the exception of the peer reviewer. The interviewed stakeholders did not specify any stakeholders either inside or outside government who they believe should have been more closely involved; they are satisfied with how widely the results have been shared.
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:	Only informal reflective conversations have taken place among individual members of the steering committee and service provider; they did not formally discuss it in a meeting. There has also been some written correspondence between DPME and the service provider about what should be done in future if an evaluation comes across allegations of procurement irregularities, as was the case in this evaluation.
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 SAPS interviewee indicated that the evaluation was of symbolic value to the programme, that it motivated from an external perspective the importance of continuing to invest in forensic services. He pointed out that the importance attached to the evaluation - especially the improvement plan and DPME's expected follow-up on it - has helped ensure that SAPS "takes it seriously" and that "now someone has to do something about [one of the trends highlighted in the report] because it is documented". Furthermore, the fact that the evaluation not only picked up on but also confirmed from an external perspective some challenges that officials at SAPS have been grappling with, was seen as having important symbolic value. The DPME noted that the evaluation has attracted SAPS senior management's attention and that SAPS has subsequently put forward further proposals for evaluations, suggesting this first evaluation experienced has shown them the value of such an exercise.

Rating: 4: The evaluation study is seen by interviewed stakeholders as being of substantial symbolic value to the policy or programme and has noticeably raised its profile amongst stakeholders

Standard: The evaluation study is of conceptual value in understanding what has happened and possibly in shaping future policy and practice

Comment and Analysis: All interviewees are in agreement that the evaluation was extremely valuable to stakeholders in understanding what has happened. The service provider interviewee has the impression that the evaluation started a change management process in the organisation, to improve on the challenges the evaluation identified. Similar views were expressed by the other interviewees. The SAPS interviewee indicated that the process of drafting the National Forensic Strategy benefited from the findings as well as recommendations that emerged from the evaluation, hence the evaluation has already begun to shape future policy. He also noted that the evaluation drew a clearer conceptual connection between daily activities related to forensics, and the ultimate intended impact of this work, which was not so clear or front-of-mind for role players before the evaluation. The DPME interviewee also believed that the evaluation will through the improvement plan help ensure practical improvements.

Rating: 5: The evaluation study is of great conceptual value and all interviewed stakeholders expressed confidence that it would constructively shape policy and practice

References

DNA Economics, 2015. Economic Evaluation of the Incremental Investment in Forensic Services: Fieldwork Report. Version 1 (14 August 2015).

DNA Economics, 2016. Economic Evaluation of the Incremental Investment in Forensic Services: Draft Evaluation Report. Draft 4 (18 August 2016).

DNA Economics, 2015. Economic Evaluation of the Incremental Investment into the SAPS Forensic Services: Literature Review for the Department of Performance Monitoring and Evaluation.

No author, 2014. Inception Meeting: Implementation/Economic Evaluation of the Incremental Investment into SAPS Forensic Services (Minutes of the meeting held 17 November 2014). 24 November 2014.

DNA Economics, 2015. Economic Evaluation of the Incremental Investment into the SAPS Forensic Services: Discussion Guides. Version 2 (15 June 2015).

Curriculum Vitae of Dr Keith Brian Morris (no date).

DNA Economics, 2017. Report on the Economic Evaluation of the Incremental Investment in Forensic Services: Summary Report. 17 February 2017.

DNA Economics, 2015. Economic Evaluation of the Investment into Forensic Programme: Evaluation Report. Version 5 (15 June 2015)

DNA Economics, 2017. Report on the Economic Evaluation of the Incremental Investment in Forensic Services: Statistical Appendices. 17 February 2017.

DPME and SAPS, 2014. Terms of Reference for the Economic Evaluation of the Incremental Investment into the SAPS Forensic Services. Draft (13 March 2014).

DNA Economics, 2014. Inception Report: Economic Evaluation of the Investment into the SAPS Forensic Services: Inception Report for the Department of Performance Monitoring and Evaluation. Version 3.0 (November 2014).

DNA Economics, 2017. Report on the Economic Evaluation of the Incremental Investment in Forensic Services: Full Report. Version 6 (10 February 2017).

List of Interviewees

Mr Jabu Mathe, Director of Evaluation, Department of Planning, Monitoring and Evaluation. Telephone Interview, 23 November 2017.

Brigadier Joe Smith, Section Head: Forensic Database Management. Telephone Interview, 24 November 2017.

Ms Amanda Jitsing, Director of Public Policy, DNA Economics. Telephone Interview, 21 November 2017.