

# Special edition of African Evaluation Journal on the national evaluation system

## Authors:

Ian Goldman<sup>1</sup>  
Babette Rabie<sup>2</sup>  
Mark Abrahams<sup>3</sup>

## Affiliations:

<sup>1</sup>Department of Planning,  
Monitoring and Evaluation,  
South Africa

<sup>2</sup>School of Public Leadership,  
Stellenbosch University,  
South Africa

<sup>3</sup>Division for Lifelong  
Learning, University of the  
Western Cape, South Africa

## How to cite this article:

Goldman, I., Rabie, B. &  
Abrahams, M., 2015, 'Special  
edition of African Evaluation  
Journal on the national  
evaluation system', *African  
Evaluation Journal* 3(1), Art.  
#166, 4 pages. [http://dx.doi.  
org/10.4102/aej.v3i1.166](http://dx.doi.org/10.4102/aej.v3i1.166)

## Copyright:

© 2015. The Authors.  
Licensee: AOSIS  
OpenJournals. This work is  
licensed under the Creative  
Commons Attribution  
License.

## Read online:



Scan this QR  
code with your  
smart phone or  
mobile device  
to read online.

## Overview of the edition

This edition is a significant marker – recognising the development and implementation of a national evaluation system in South Africa. The system was established in 2011 and 50 evaluations have been completed, are underway or are starting, representing around \$6 – \$7 billion of government expenditure, as well as many elements of a comprehensive system at national and provincial levels. South Africa thus joins a few countries in the world that have national systems: Mexico and Colombia, which are well known and from whom we have learned, Canada, Chile, Peru and our peers who we are working with closely in Africa, Uganda and Benin.

At the same time, a key text has been written on evaluation management in South Africa and Africa, which was launched in May 2015 and which we review in this edition (we refer to this as the Book), which gives a broader historical perspective, as well as more theoretical background to evaluation and practice on the African continent.

This system has not emerged in a vacuum. The article by Abrahams outlines some of the history of the evolution of evaluation in South Africa from the early 1990s. He relates three trends that emerge from this – evaluation as an emerging profession, how evaluation has emerged as part of the governance role in the country and also how this has driven the development of evaluation as an industry, with an emerging set of companies and universities providing evaluation services.

Monitoring and evaluation (M&E) has been studied internationally since the 1960s, covered by Charlene Mouton *et al.*, and by Goldman *et al.* in the Book, and driven since the 1980s by the emergence of the new public management approach with its focus on measurement. Mackay (2008) quotes in the Book:

The performance orientation of the public management is here to stay. It is essential for successful government. Societies are now too complex to be managed only by rules for input and process and a public-spirited culture. The performance movement has formalized planning, reporting, and control across many governments. This has improved the information available to managers and policy makers. (p. 554)

Whilst this is true in South Africa as well, what has emerged is a strongly compliance-driven culture in government, with the Auditor General and Treasury as the strong drivers of performance reporting. The article by Chitepo and Umlaw gives a picture of how this looked in 2012, when 89% of departments had an M&E unit, but 54% of departments reported that problems were not treated as an opportunity to learn. It suggests that 'the focus of M&E is generally monitoring of outputs at operational level rather than enabling departments to probe the effectiveness of strategy and policy in terms of the outcomes and impact resulting for the public'.

The article by Paine and Sadan provides a very important contextual piece on the attitude of senior managers to evidence. This work was conducted in 2011 and involved interviewing 54 senior managers from director-general to director levels. This provides an important and nuanced picture of these attitudes and how to consider the use of evidence in the complex policy terrain of government. There was almost unanimous agreement that evidence-based policy making (EBPM) should represent a move from opinion as the basis for policy to a more rigorous use of the available body of evidence, replacing the use of 'power derived from position .... (to) a discourse of reason'. There were two discernible groups of thinking that underpinned attitudes on what kind and use of evidence was most desirable for EBPM:

- Predictive, scientific and objectively verifiable: independent experts derive unambiguous facts through replicable, valid and generally scientific methods providing objectively verifiable proof (emphasised by 15).
- Formative, emergent, probabilistic and contested: an iterative search for better explanations and understanding of how to achieve politically derived values in which the choice of facts and sources is influenced by existing ideas, ideology, mindset, values and interests and is

subject to specific and changing contextual factors (emphasised by 32).

This difference can also be seen in the different approaches to evaluation worldwide. Many in the 'formative' group noted the high levels of complexity and instability in policy contexts in South Africa, which require rapid cycles of evidence-driven learning. Officials noted that where evidence is used, it is often to defend decisions already made, for example to secure funding from Treasury. In many of the negative examples given, there was no time for consideration of even minimal evidence, even when the decisions have very significant long-term implications or risks. One respondent indicated that:

It is a tough job to get policy research on the agenda and drive it through – there are so many role players and factors. People look for quick ways of pushing it through. It matters more who is pushing, not the evidence. We take the route of least resistance because producing hard evidence is a tough job and nobody thanks you for it.

An interesting issue was at which stages of the policy cycle evidence was used. The majority of officials felt that in the *agenda setting* phase the failure to use evidence effectively, particularly learning from past experience, was a key reason for policy weakness and failure. They also highlighted the absence of effective problem, needs or *options* analysis. Several noted that some policy options are 'taboo' even if all the evidence points to their being the best available option in the circumstances. Paine and Sadan report that officials screen evidence, often based on political assumptions, thus limiting political principals' capacity to make informed choices, and that policymakers did not generally request or assess options before deciding on a policy approach. Implementation requirements and feasibility seldom inform choices about the best and most sustainable policy option. Even where some analysis is done, a clear '*theory of change*' or hypothesis spelling out how it is assumed the intervention will work is seldom made explicit enough to be tested through experience, making evaluation and ongoing evidence-driven learning and policy improvement difficult.

So these two articles provide an important context and lens within which to view the emerging national evaluation system. How much is it able to provide scientific evidence to support prediction, how much to support a formative process where government uses evidence to learn and improve?

In 2012 the Department of Performance (now Planning), Monitoring and Evaluation was a relatively new actor on the national scene, having been created only in 2010, and many of its systems, such as the management performance assessment and the national evaluation system, only emerged in 2011. The chapter in the Book by Goldman *et al.* on evaluation in SA covers this emergence of the range of DPME's systems and an article by Goldman *et al.* in this edition focuses specifically on the emergence of the national evaluation system. The system started with a visit to Mexico,

Colombia and the US in mid-2011, which led to the approval of a National Evaluation Policy Framework by Cabinet in November 2011 and the roll-out through a series of National Evaluation Plans from 2012. The Goldman *et al.* article gives a timeline for this, along with the emerging set of systems, including standards, competences, training, guidelines, communication mechanisms and so on. Specific articles then go through these elements in turn.

The other article by Goldman *et al.* in this edition discusses the development of standards, how this has developed into a quality assessment system and the results that emerge from this. A number of international standards were reviewed and in July 2012 DPME with the South African Monitoring and Evaluation Association (SAMEA) decided that the most useful framework for South Africa was the OECD DAC approach, based on the phases of evaluation. The South African standards were drafted and first published in August 2012. These cover the following phases:

- Overarching considerations.
- Phase 1 – Planning, design and inception.
- Phase 2 – Implementation.
- Phase 3 – Reporting.
- Phase 4 – Follow-up, use and learning.

A tool was developed to assess the different standards. The article outlines the tool and how it was developed, as well as some of the emerging results from applying the tool. The tool was applied to evaluations identified through an audit and around 15% of evaluations fell below the quality threshold of 3 on a five-point Likert scale. All of the evaluations undertaken through the national evaluation system since have scored above 3 and it is encouraging to see provinces with provincial evaluation plans now also using this system. The article points to above satisfactory levels of methodological appropriateness and appropriate data gathering techniques, but a number of specific shortcomings in evaluation practice.

An article was meant to be in this edition on capacity, the emergence of competencies, a set of training courses, other methods of capacity development including learning-by-doing and so on. There are some major pieces of work underway at present that relate to this, so it was decided to delay the article to draw from that experience.

There is then a series of articles around specific early evaluations conducted through the national evaluation system. The article on early childhood development (ECD) by Davids *et al.* describes the first pilot evaluation, which was used to develop many elements of the evaluation system. This was a complex area covering a wide range of factors affecting early childhood development, including educational issues, issues of child nutrition and so on. The evaluation was very influential in recommending the extension of ECD to cover the first 1000 days of life from conception, the need for a more comprehensive set of services and better targeting of poor children, all of which has been endorsed in a new

ECD policy which was approved by Cabinet for gazetting in March 2015. What was interesting also about this evaluation methodologically was rather than primary research, it was a diagnostic evaluation which drew from 112 existing reports, that is, a research synthesis. It also illustrates the complexity of the policy environment and how it linked to parallel interventions, notably an ECD conference organised by the minister, and how the improvement plan drew from the evaluation as well as the conference to create a plan which could be approved politically as well as technically.

The article on Grade R (reception year of schooling) by Samuels *et al.* describes the first impact evaluation done under the NES. This was interesting methodologically in how it managed to collate existing data collected from annual national assessments in over 18 000 schools, with data on the number of Grade R enrolments, to estimate the treatment effect of Grade R. Grade R is a flagship programme for government to address educational inequality and there was some disappointment at the low levels of effects seen in poor schools and in underperforming provinces. However, the evaluation does point to the need to improve quality and not just to roll out increased access. Methodologically the evaluation also indicated the limitations of a quantitative methodology without a formative component to understand why particular effects are occurring and to unpack the theory of change, in this case mitigated by the extensive knowledge and work of the principal investigator. However, as there was no primary data collection this was a very inexpensive exercise.

The article on business process services is an evaluation of a programme of a key partner, the Department of Trade and Industry (the dti), which has put forward a number of implementation evaluations of its programmes. This is a much smaller and more contained intervention than the previous two and so the evaluation was also more straightforward. It was possible to make clear and simple recommendations for improvement, which have been adopted and the scheme relaunched, a good example of instrumental use.

Finally there are two different articles, the first on emerging work around communication and the second analysing the partnership with SAMEA. Some basic communication has been done from an early stage of the evaluation system, including evaluation summary reports consisting of a one-page policy summary, a five-page executive summary and a 25-page main report, an evaluation newsletter, reports going public and evaluation reports being sent to Parliament. However, there is room for much more work to extract the use and impact from the evaluations, work which has started with a communication strategy. The article talks about ongoing interactive communication between practitioners and evaluators during the evaluation process, as well as communication of findings. Amisi raises the work of Rochow saying that 'for communication to lead to action it has to draw and capture the attention of the audience; respond to the needs of that audience; satisfy the need; invite the audience to visualise a situation where the problem has been resolved and compel them to take action by appealing to their

emotions and using symbols/visuals to present ideas'. It also talks of the challenges DPME faces in that since evaluations are a collaboration between DPME and the custodian departments, communication is effectively guided by how much the custodian departments want to communicate.

The last article, by Beney *et al.*, looks at the relationship between SAMEA and DPME. SAMEA was established a few years before DPME and draws together a wide range of practitioners, of whom perhaps 50% – 60% are from government. The two partners have made considerable efforts to work together effectively, including through a joint standing committee. As DPME has got stronger in the evaluation space, some tensions have emerged with fears from some that DPME is dominating. The article explores this relationship using a matrix of organisation identity versus mutuality.

There are four quadrants in the model. Quadrant 1, partnership, represents a partnership where mutuality and separate organisation identity are maximised. Quadrant 2, contracting, represents a situation where specific organisational characteristics and contributions are determined by one organisation, but sought in another based on organisational identity to fulfil predetermined ends and means. Quadrant 3, extension, describes when one organisation calls the shots and the other organisation has little identity and follows the dominant organisation's lead. Quadrant 4, co-optation and gradual absorption, happens when a partner organisation compromises its identity by exchanging its services for the benefit of serving the dominant organisation, either consciously or unconsciously. The conclusion is that the relationship between SAMEA and DPME does appear to be a genuine partnership, although as mentioned some worry about government becoming too dominant.

So this edition provides a good overview of the national evaluation system as at the end of 2014. Some will criticise it as being more descriptive than evaluative at this stage, but this is intentional, to provide a base to which to refer in future articles reflecting on how the system is working in practice, what is working well and what isn't and how the system can improve.

There are some important areas that are evolving as we speak:

- The evolving relationship between SAMEA, DPME and higher education institutions around increasing the professionalism of evaluation and what should be a roadmap for the country.
- Evaluation moving from a voluntary activity to one where the use of evaluations is rated as part of the management performance assessment process, so driving behaviour and moving to a government-wide system.
- Increasing the focus on programme planning, a key area of weakness in government, and how this links to evaluation.
- South Africa developing partnerships with other countries, notably Uganda, Benin, Mexico and Colombia around government evaluation systems.

- A move more into the use of evaluation results and tracking what influence they are having. To that end an evaluation of the evaluation system is planned for 2016 and 2017, to be done in collaboration with the World Bank.

We will hear more about these developments in future articles. This edition is therefore intended to serve as a marker of an important moment in South Africa, where a national evaluation system has been established, and which will guide much work in the future.