




A micro-level outcomes evaluation of a skills capacity intervention within the South African public service: Towards an impact evaluation

**Authors:**

Petronella Jonck¹ 
 Riaan de Coning² 
 Paul S. Radikonyana¹ 

Affiliations:

¹Research and Innovation,
 National School of
 Government, South Africa

²University of Stellenbosch,
 Business School, South Africa

Corresponding author:

Riaan de Coning,
 riaan.deconing@thensg.gov.za

Dates:

Received: 22 Sept. 2017
 Accepted: 16 May 2018
 Published: 18 July 2018

How to cite this article:

Jonck, P., De Coning, R., &
 Radikonyana, P.S. (2018).
 A micro-level outcomes
 evaluation of a skills capacity
 intervention within the
 South African public service:
 Towards an impact evaluation.
*SA Journal of Human Resource
 Management/SA Tydskrif vir
 Menslikehulpbronsbestuur*,
 16(0), a1000. [https://doi.
 org/10.4102/sajhrm.
 v16i0.1000](https://doi.org/10.4102/sajhrm.v16i0.1000)

Copyright:

© 2018. The Authors.
 Licensee: AOSIS. This work
 is licensed under the
 Creative Commons
 Attribution License.

Orientation: Interest in measuring the impact of skills development interventions has increased in recent years.

Research purpose: This article reports on an outcomes evaluation under the ambit of an impact assessment with reference to a research methodology workshop.

Motivation of the study: A paucity of studies could be found measuring the workshop outcomes, especially within the public service as it pertains to training interventions.

Research approach/design and method: A pretest–post-test research design was implemented. A paired-sample *t*-test was used to measure the knowledge increase while controlling for the influence of previous training by means of an analysis of variance and multiple regression analysis.

Main findings: Results indicated that the increase in research methodology knowledge was statistically significant. Previous training influenced the model only by 0.8%, which was not statistically significant.

Practical/managerial implications: It is recommended that the suggested framework and methodology be utilised in future research as well as in monitoring and evaluation endeavours covering various training interventions.

Contribution/value add: The study provides evidence of the impact generated by a training intervention, within the South African Public Service. Thus, addressing a research gap in the corpus of knowledge.

Introduction

South Africa cannot reach its developmental goals by relying on opinion-based policy practices. In this regard, the Department of Planning, Monitoring and Evaluation (DPME) (2014) notes that most policies originate in the planning meetings of political parties. Thus, the probability that the resulting policies would be evidence-based is unlikely. Evidence-based policy development can be defined as an approach that supports the provision of a knowledge base by ensuring that research evidence is the cornerstone of policy development and implementation. As a result, opinion-based policy-making and ad hoc decision-making methods are challenged (Davies, 2004). Research is, *inter alia*, deemed to be the missing link in providing high-quality, evidence-based policy interventions (Zwar, Weller, McCaughan & Traynor, 2006).

Additionally, research plays a supportive role in the achievement of a skilled and capable developmental state. To accomplish the shared vision of realising a developmental state, the National Development Plan (NDP) states that ‘a well-functioning research capacity is vital in sustaining growth and improving productivity’ (National Development Plan 2030 (NDP), 2012:131). The NDP likewise stipulates that research conducted by government departments, and other organs of state, has a crucial role to play in improving South Africa’s global competitiveness (NDP, 2012:293). There is also a broader argument to be made for the development of research capacity in the South African public service. The current age is characterised by globalisation and major knowledge-based economies. Thus, an investment in knowledge and skills development will ensure the progress of the country’s labour force and consequently the country’s ability to compete in the world economy (Goujon, Lutz & Wazir, 2011). Research plays a pivotal role in the knowledge economy in that it lays the foundation for the production and dissemination of knowledge (Leahey & Moody, 2014).

Furthermore, the NDP (2012:364) states that inadequate public service performance could be attributed to skills deficiencies and unsuitable staff appointments. A lack of an adequately

Read online:

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

skilled staff component in the public service has, therefore, been a cause for concern. Consequently, the interest in measuring the impact of skills development interventions has increased (Pillay, Juan & Twalo, 2012). Abrahams (2015) notes that within the context of the public service, monitoring was underscored until the New Public Management (NPM) approach emphasised accountability. Thereafter, a shift occurred towards including evaluation as a key performance management tool. In this respect, impact evaluations are deemed crucial, because these evaluations provide information about the impact produced by an intervention and can be undertaken in a programme, a policy or a capacity-building intervention (Rogers, 2014). Furthermore, an impact evaluation can be undertaken either for formative (i.e. the improvement or reorientation of a programme or policy) or summative purposes (i.e. to inform decision-making regarding the continuation, or discontinuation, of a programme or policy), as pointed out by Rogers (2014). It therefore suffices to state that 'an impact evaluation encompasses any evaluation that systematically and empirically investigates the impact produced by an intervention' (Rogers, 2012:1). The goal of an impact evaluation can be to promote a particular type of intervention as best practice in a specific field or development (Weyrauch & Langou, 2011).

Despite the importance of the aforementioned, few examples of successfully implemented evaluation studies could be found (Abrahams, 2015), especially in terms of training interventions. O'Malley, Perdue and Petracca (2013) note that many training interventions do not consistently provide evidence that links specific training efforts to desired outcomes, despite a commitment to training. Colquitt and Simmering (1998) identified the three keystone examined training outcomes subsuming declarative knowledge, task performance and post-training self-efficacy, based on a 20-year longitudinal meta-analysis of 106 training interventions. Research indicated that cognitive skills development evolves from an initial knowledge compilation (viz. gaining knowledge via instruction) to procedural knowledge (viz. task performance) and advances towards self-efficacy, which refers to internalised perceived performance capabilities. Hence, determining the knowledge outcome is a necessary precursor which influences task performance and self-efficacy that results from task performance (Yi & Davis, 2003).

In view of the discussion thus far, the aim of this research was to conduct an outcomes evaluation on a research methodology skills capacity workshop within the context of the public service.

Construct definition

An evaluation can be defined as a cross-sectional or periodic application that is aimed at providing credible evidence to guide decision-making. An evaluation may assess relevance, efficiency, effectiveness, impact and sustainability (Department of Planning, Monitoring and Evaluation (DPME), 2007). From the literature studied, six dominant

methods of evaluating skills development interventions have been identified, including methods which encompass efficiency indicators, self-reported behavioural change, on-the-job follow-ups, proxy indicators, policy evaluation and knowledge testing. The subject literature on evaluation studies indicates that knowledge and skills testing provides the best example of objectively evaluating skills development interventions (Pillay et al., 2012:28). To this effect, literature search reveals that pretest and post-test designs are widely used in behavioural research, primarily to measure knowledge gained from participating in a training intervention (Dimitrov & Rumrill, 2003). By comparing participants' post-test scores with their pretest scores, it is possible to determine whether the training or skills development programmes were successful in increasing participants' knowledge of the training content (Dimitrov & Rumrill, 2003; Pillay et al., 2012). It is worth noting that in pre- and post-testing, the researcher does not take into consideration whether increased knowledge will result in behaviour change (Pillay et al., 2012).

Theoretical underpinning

Mouton (2010) notes that, although programme evaluation was introduced into South Africa by international funding organisations, it was not until this practice was accepted and amalgamated in public service policy documents and frameworks that a culture of evaluation emerged. As such, using the NDP's concept of a developmental and capable state, the Department of Planning, Monitoring and Evaluation (DPME), as the organ of state responsible for the planning, monitoring and evaluation (Abrahams, 2015), endorses performance monitoring and evaluation as a key management intervention that should enhance public service capacity and increase the impact of service delivery interventions (DPME, 2014). A key initiative of the aforementioned department has been to introduce the outcomes approach which emphasises linking inputs and activities to outputs and outcomes (Phillips, 2012). An example of a programme evaluation is the implementation evaluation conducted on the business processes services incentive programme in the Department of Trade and Industry using a cost-competitiveness analysis approach by Mashalaba, Wyatt, Mathe and Singh (2015).

The DPME (2007) approved a monitoring and evaluation framework which consists of five key elements. The first key element is the inputs that represent the resources utilised, including fiscal resources and equipment. The second key element is the activities that encompass the process or actions that make use of a plethora of inputs to produce the desired outputs and, ultimately, outcomes. The third key element is the outputs or the final products that represent the goods and services produced. The fourth key element is the outcomes. These are the medium-term results for specific beneficiaries which are the consequence of achieving specific outputs. Outcomes should ideally relate clearly to the strategic goals and objectives of institutions as indicated in their annual reports. Lastly, impact can be seen as the result of achieving specific outcomes.

In light of the above, a theory-based theoretical underpinning was utilised in this study. White (2009) noted that theory-based evaluation, which refers to examining the assumptions underlying the causal chain from inputs to outcomes and possibly impact, is a well-established approach. Bank (2012) defined a theory-based evaluation as an approach to evaluation that underscores a specific manner of structuring and undertakes the analysis based on a theory of change, also referred to as a 'programme logic' or 'logic model'. The theory of change typically commences with a sequence of events and results (i.e. outputs, immediate outcomes, intermediate outcomes and ultimate outcomes) that are expected to occur owing to the intervention (Bank, 2012). Table 1 provides an indication of the logic model utilised to evaluate the training intervention in this study.

O'Malley et al. (2013) note that programmes generally report on training outputs, such as number of participants trained, lending support to the logical framework formulated in Table 1. These output indicators enable stakeholders to aggregate data across a variety of training interventions, such as workshops, lectures or e-learning programmes. Nevertheless, output indicators cannot be utilised in the evaluation of whether or not the training encounter improves knowledge or practice. In addition, although the aforementioned logical framework is focussed on the micro-level of complexity as training is provided to an individual, it should be borne in mind that training has various complexity levels. A system can be defined as a structured entity consisting of components sufficiently interrelated and interdependent, thereby forming a whole. As a component of a system, training can influence the system at various levels of complexity, in accordance with the hierarchy of systems. For example, a vertical hierarchy of systems can in theory include micro-, meso-, macro-, national and supranational levels of complexity (Ureda & Yates, 2005). Thus, the influence of training can extend to various levels of complexity, commencing with the micro-level (Frei, 2011).

Multiple frameworks have been developed to evaluate the complex phenomenon referred to as training. The most frequently utilised framework is the Kirkpatrick Model which identifies four levels at which training can be evaluated, namely reaction, learning, behaviour and results (O'Malley et al., 2013). The last three categories correspond to three levels of complexity. Firstly, learning occurs at the micro-level or at the individual level. Secondly, behaviour ensues at the meso-level or within the organisation, and

thirdly, results arise at the macro-level within the broader community. As such, Table 2, as adapted from O'Malley et al. (2013:6), provides an indication of training evaluation outcomes identified in a systematic review that emphasises training outcomes as well as the levels of complexity.

Pursuant to the foregoing discussion, methodological challenges have been reported as especially problematic. These challenges include, *inter alia*, the distal (decentralised) nature of outcomes and impact, and the fact that it may not be possible to generalise the findings to the population (O'Malley et al., 2013). Thus, the problem in measuring training impact is the fact that a micro-level intervention is implemented, but a macro-level impact is expected. This is especially challenging in the case of public service training institutions where a micro-level intervention is effected (e.g. an individual is trained), but impact is expected by departments that require training at a macro-level (e.g. impact is seen as resolving service delivery problems experienced by the procuring department). In addition, not all training interventions are aimed at generating macro-level impact. For example, training interventions for programme 1, namely Corporate Services, which is standard in government departments, may not result in improved service delivery to communities. In spite of challenges in the measurement of impact, it is essential to evaluate the effectiveness of training to ensure that limited fiscal resources, manpower and hours devoted to attending training yield a return on investment (O'Malley et al., 2013).

TABLE 2: Training evaluation outcomes based on a systematic review of relevant published literature.

Level of complexity	Indicator	Number of published articles
Individual (micro-level)	Knowledge, attitude or skills	21
	Individual performance (knowledge test, or according to the annual performance agreement)	30
	Service delivery outcomes	12
Organisational (meso-level)	Performance improvement (based on IMPAT scores, quarterly organisational performance report, annual performance plan or presentation made to the portfolio committee)	16
	Systems improvement	4
	Service delivery outcomes	10
Community (macro-level)	Client satisfaction or service delivery satisfaction	11

IMPAT, Management Performance Assessment Tool.

Source: Adapted from O'Malley, G., Perdue, T., & Petracca, F. (2013). A framework for outcome-level evaluation of in-service training of health care workers. *Human Resources for Health*, 11(50), 1–12. <https://doi.org/10.1186/1478-4491-11-50>. p.6

TABLE 1: Training evaluation logical framework.

Inputs	Activities	Outputs	Outcomes	Impacts
The training intervention was free of charge. However, the department that requested the training paid for catering. Additionally, training material was prepared and copies printed by the training institution (thus the costs associated with the intervention were shared).	The training intervention itself. The activities encompassed lecturing, class discussions and practical exercises on a resource CD.	The number of respondents who underwent training. Thus, 32 respondents participated and completed the training intervention.	Participants with knowledge of research methodology. An increase in research methodology knowledge would be the outcome of the capacity-building intervention. Outcome for procuring department – individual performance outcomes (i.e. the completion of postgraduate studies, thus HR reports on bursary results).	Formative impact (viz. improvement, reorientation or alternatively sustaining the training intervention). Impact for the procuring department would be decreased fruitless expenditure ascribed to employees' failure to complete postgraduate studies for which the department provided bursaries.

HR, human resources; CD, compact disc.

Based on the logical framework presented in Table 1, this article reports on an evaluation of a training intervention to determine a hypothetical knowledge increase. It should be noted that although this article reports on outcomes, the trend is towards an impact evaluation. The rationale for the foregoing contention is based on the work of Weyrauch and Langou (2011:12) who note that impact can be measured at various levels of complexity, of which the first aims to influence a particular project, programme or policy. This refers to a tangible public intervention with a particular objective, defined recipient population, budget and set of activities with clearly defined benefits. It is important to note that what is referred to as an impact evaluation at the micro-level can be targeted at either changing a part of the programme or alternatively sustaining it (Behrman, 2010:1476). Furthermore, the National Evaluation Plan (2012) reports on a project determining the learning outcomes of a Grade R educational intervention as an example of an impact evaluation. In addition, Babbie and Mouton (2011:340) note that the term 'impact assessment studies' refers to the degree to which a programme has produced the desired outcomes. The abovementioned authors elaborate by distinguishing between four types of evaluation studies, namely the evaluation of need, process, outcome and efficiency. The evaluation of outcome, under the ambit of an impact assessment, could entail a knowledge increase, and a behavioural and/or an attitudinal change (Babbie & Mouton, 2011). Because of the fact that this study controlled for the influence of previous training, the view of Samuels et al. (2015), who found that an educational intervention had limited impact on later educational outcomes, is pertinent.

Research method

A pretest–post-test repeated measure research design was incorporated in the study to determine the effect of the training intervention on research methodology knowledge. Babbie and Mouton (2011) note that the logic of an impact assessment is based on the supposition that an intervention has certain effects. As such, the standard evaluation approach to investigate this is a pretest–post-test design. Gertler, Martinez, Premand, Rawlings and Vermeersch (2011:13) further note that retrospective evaluations assess the programme impact after implementation, generating comparisons *ex post facto*. This study could be classified as an *ex post facto* research design as respondents were related to the different variables prior to data collection. Thus, participants were not assigned to experimental and/or control groups (Jonck, 2014). As a result of this, three limitations of a pretest–post-test research design can be identified, namely the absence of a control group against which comparisons can be made, the teaching effect and the unobserved moderating variables intrinsic to the facilitator (Wagner, Kawulich & Garner, 2012).

Research hypotheses

The primary research hypothesis states that: 'A research methodology capacity-building intervention does have a

statistically significant influence on participants' research methodology knowledge'. The secondary research hypothesis specifies that: 'Prior research methodology training does have a statistically significant influence on research methodology knowledge after the training intervention'. The primary hypothesis was verified by Gertler et al. (2011:7) who note that in the context of an impact assessment the research question would hypothetically be: 'What is the impact or causal effect of a programme on an outcome of interest?'

Research process

The skills development facilitator of a national department requested the training intervention after conducting a training needs analysis, at which time a need for research methodology capacity-building was registered. From consultation with the facilitator, it would appear that the request arose because participants lacked the research capacity to complete their higher education postgraduate studies, and this influenced bursary requirements and fruitless expenditure. Thus, participants volunteered to undergo training. The two-day training intervention consisted of a quantitative and qualitative section. As a pretest–post-test method was implemented, assessment took place prior to and after the two-day training programme. Standard ethical guidelines according to Wagner et al. (2012) were adhered to throughout this research. Respondents were informed of the nature and scope of the study. Participation was completely voluntary and respondents were not compelled to participate. Respondents completed the questionnaire anonymously, and the information received remained confidential. Finally, no physical or psychological harm occurred as a result of respondents' participation. In fact, the research study was used as an example in the training intervention.

Description of the intervention

The training intervention occurred at the micro-level, which is indicative of the notion that training was provided to the individual, taking into consideration that individual training would inevitably result in a hypothetical knowledge increase (i.e. research knowledge) and ultimately an organisational outcome (i.e. the development of evidence-based policies). The research methodology workshop was hosted at a public service training facility. Facilitators utilised teaching aids, such as PowerPoint presentations, flip charts and electronic devices. The mode of delivery encompassed traditional lecturing, class discussions and practical exercises on a resource compact disc (CD) that was distributed to participants. The syllabus consisted of two sections, namely a qualitative and quantitative section. The qualitative section focussed on the fundamentals of research, including differentiating between personal information-seeking and academic research. By way of introduction, the three research methodologies (viz. quantitative, qualitative and mixed methods) were briefly mentioned before attention was focussed exclusively on qualitative research. The aspects covered by this topic included the nature of qualitative research, various designs (e.g. phenomenological, thematic

analysis and grounded theory), development and implementation of data-gathering instruments and, finally, data analysis. The section on quantitative research emphasised various themes. These included the following:

- The research process, quantitative research designs (viz. experimental, quasi-experimental and non-experimental designs) and sampling
- Developing and/or designing an appropriate questionnaire (viz. types of survey questions, and scaling techniques which might include using either a Likert or a semantic differential scale)
- Coding and capturing data
- Cleaning the data set
- Basic descriptive statistical analysis (viz. measures of central tendencies and dispersion).

Research participants

The study made use of a non-random sampling technique, namely convenience sampling (i.e. participants who attended training were involved in the study). The final sample consisted of a total of 32 public service employees which comprised 16 (48.5%) male and 17 (51.5%) female respondents. Note that one respondent did not complete the post-test and was eliminated from further analysis. As far as the age distribution of the sample was concerned, 6.1% of the sample was in the 25 years and younger age category. Moreover, 15.2% of the respondents were in the 26–35-year age category, followed by the majority of the sample (42.4%) who were between 36 and 45 years of age. Thirty-three per cent of the respondents were in the age group of 46 to 55 years and one (3%) respondent was in the 56–65-year age category.

In terms of the highest academic qualification, one respondent (3%) had a Grade 12 coupled with a certificate, two (6.1%) respondents had diplomas, seven (21.2%) respondents held bachelor's degrees, five (15.2%) respondents had postgraduate certificates and/or diplomas, and eight (24.2%) held honours degrees, while nearly one-third ($n = 10$; 30.3%) of the sample held master's degrees.

Respondents were requested to indicate their previous work experience. Results indicated that 9.1% of the sample had experience in the legislative sector, 54.5% of the sample had originally been employed in national government, 18.2% had work experience in provincial departments, 3% (representing one respondent) had work experience at local government level, two (6.1%) of the sample worked in the private sector and, lastly, two (6.1%) respondents indicated that their previous work experience could not be categorised in the abovementioned response categories.

The previous training which participants had received ranged from information seeking (3.4%), Sabinet training (3.4%), research methodology skills development (13.8%) or a combination of these (27.9%), to mention a few. However, 44.8% of the sample indicated that they had not had any previous research-related training.

Measuring instrument

Primary data were collected by implementing a questionnaire consisting of two sections, namely a biographical section and a section containing questions relating to the content of the workshop. The section that explored respondents' biographical information included questions regarding gender, age, highest academic qualification and previous research methodology training. The latter is important as it can influence the level of research methodology knowledge and can be considered a moderating variable.

Section B of the questionnaire consisted of 28 questions that probed respondents' knowledge of both qualitative and quantitative research methodology. Typical questions included, for example, 'Sample size is pivotal in a quantitative research design', 'Quasi-experimental research designs are typically found in qualitative research' and 'Qualitative research involves various stages of coding which include initial coding and open-ended coding'. Participants were requested to select the most appropriate option from a four-point Likert scale, with options ranging from 'Strongly agree' (1) 'Agree' (2) 'Disagree' (3) and 'Strongly disagree' (4). A Likert scale was used because the underlying assumption of the statistical tests performed is that the residual scores should be normally distributed and not categorical (Pallant, 2011). As the questionnaire was specifically developed for the study, the reliability and validity of the scale had to be investigated (De Souza, Alexandre & Guirardello, 2017).

Reliability refers to the likelihood that a given measure would yield the same results in various iterations, while validity refers to the extent to which a specific measurement provides data that relate to the accepted meaning of a particular concept. In general, reliability is measured by means of Cronbach's alpha coefficient, while validity can be determined by means of face and construct validity (De Souza et al., 2017). Cronbach's alpha coefficient was used to calculate the inter-item consistency (0.225), with an alpha of 0.88, emphasising the reliability of the scale.

In terms of face validity, the measuring instrument was circulated for inputs to five researchers within the public service with numerous years of research experience in the public sector, as well as institutions of higher learning. Factor analysis was used to determine the construct validity of the questionnaire as Lu (2014), for example, indicates that factor analysis can be seen as an efficient tool to ascertain the underlying construct validity of a measurement. Results indicated that the data were factorable, as the Kaiser–Meyer–Olkin (KMO) test for sampling adequacy returned a value of 0.663, and Bartlett's test of sphericity reverted a statistically significant value on the 99th percentile, as indicated by the p -value accompanied by double asterisks ($\chi^2 = 494.527$; $df = 378$; $p = 0.000^{**}$). An exploratory factor analysis with oblique (oblimin) rotation was performed, and it was determined that nine components had an eigenvalue exceeding 1, accounting for 73.367% of the total variance. Nonetheless, an

inspection of the scree plot indicated a clear break after the third factor. To verify the number of factors, a Monte Carlo parallel analysis was performed. Results obtained from the analysis indicated that two components had eigenvalues exceeding the corresponding criterion value for a randomly generated data matrix of the same size (28 variables \times 33). It was therefore decided to retain two components for the purposes of further investigation in accordance with the scree plot and Monte Carlo parallel analysis results. Confirmatory factor analysis was performed with a two-factor rotation, with results displayed in Table 3.

Pursuant to the confirmatory factor analysis illustrated above, two underlying dimensions were identified. Factor 1, which emphasises aspects related to qualitative research, included items such as document analysis (factor loading of 0.723), reporting on qualitative data (with a loading of 0.695) and the aim of qualitative research, which is to gain a deep and insightful understanding of phenomena (factor loading of 0.666). Factor 2 focussed on quantitative research, for example, ethics in quantitative research with a factor loading of 0.670, specific quantitative research designs (e.g. quasi-experimental design) with a factor loading of 0.668 and the symbol for reliability (0.632 factor loading).

Descriptive statistical analysis was conducted to provide a profile of the sample. In addition, measures of central

TABLE 3: Forced two-factor component matrix.

Questionnaire items	Component	
	1	2
Q25 – Qualitative research design	0.723	-0.315
Q17 – Qualitative research design	0.695	-
Q22 – Combination of both methodologies	0.685	-0.351
Q26 – Qualitative research design	0.666	-0.308
Q21 – Combination of both methodologies	0.658	-
Q8 – Quantitative research design	0.647	-
Q19 – Combination of both methodologies	0.631	-
Q1 – General research methodology	0.621	-
Q28 – Qualitative research design	0.614	-
Q14 – Qualitative research design	0.599	-
Q23 – Qualitative research design	0.565	-
Q9 – Quantitative research design	0.513	-
Q27 – General research methodology	0.500	-
Q10 – Quantitative research design	0.435	0.315
Q4 – Quantitative research design	0.428	0.308
Q5 – Qualitative research design	0.357	-
Q2 – Quantitative research design	0.357	-
Q15 – Quantitative research design	0.307	-
Q12 – Quantitative research design	-	-
Q16 – Qualitative research design	-	0.675
Q20 – Quantitative research design	-	0.670
Q11 – Quantitative research design	-	0.668
Q3 – Quantitative research design	0.305	0.632
Q7 – Quantitative research design	-	0.630
Q24 – Quantitative research design	-	0.614
Q13 – Qualitative research design	-	0.462
Q18 – Quantitative research design	0.307	0.426
Q6 – Qualitative research design	-	0.346

Q, question.

tendency were determined to indicate the research methodology knowledge of respondents before and after the training intervention. Inferentially, the primary research hypothesis, which states that: 'A research methodology capacity-building intervention does have a statistically significant influence on participants' research methodology knowledge', was tested using a paired-sample *t*-test. Pursuant to this, an ANOVA (one-way analysis of variance) was performed to determine whether prior research methodology training had a statistically significant influence on research knowledge. Hence, the secondary research hypothesis was tested using an ANOVA. However, to further investigate the relationship, a standard multiple regression analysis was conducted to determine how much of the variance in research methodology knowledge after the training intervention could be explained by prior training (i.e. to control for prior training as counterfactual). Babbie and Mouton (2011:349) maintain that a *t*-test and ANOVA would indicate whether a statistically significant difference between the pretest and post-test results for participants would be yielded by the analysis. A statistically significant difference would, hypothetically, indicate that any differences that are observed could probably be ascribed to true differences and not chance factors.

Limitations

The following limitations should be taken into consideration when interpreting the results: Firstly, there was an overnight time gap between the implementation of the quantitative and qualitative sections of the course material, during which it may have been possible for participants to acquire additional relevant information from sources other than the training intervention. This could be considered to be a moderating variable that was not taken into consideration during the data analysis. However, further reading should be considered as an outcome of the training intervention, and various books and other resources to this effect are listed in the course material. Secondly, results are based on a small sample which cannot be seen as representative of the population. However, the aim in reporting the results of the study in this article was not to generalise the findings to the larger population, which would have required a more adequate sample size, but to report on findings within the scope of the sample. Despite the fact that the aim of the current research negates the necessity of a representative sample, caution is advised when interpreting the results. Thirdly, very little is known about the motivation of respondents other than the need that was registered by the skills development facilitator who requested the training. The reasons that respondents selected this training intervention are therefore unknown. As a result, the correlation (if any) between learning and level of motivation could also be seen as a moderating variable that was not taken into consideration during data analysis.

Ethical considerations

The authors certify that the underlying analysis is in compliance with standard ethical guidelines.

Findings

Before testing the stated hypotheses, it was important to determine the current and prior research methodology knowledge in a sample of public servants. Hence, measures of central tendency were determined, with results illustrated in Table 4.

From the descriptive results, it was evident that 44.8% of the respondents had had previous training prior to the training intervention. However, despite half of the respondents indicating that they had had previous training, their research methodology knowledge was below 50% (mean = 47.45; median = 50.00; SD = 16.80), as can be seen from Table 4. Although an increase in knowledge occurred after the training intervention, Table 3 indicates that respondents' research methodology knowledge remained below 50% (mean = 54.78; median = 57.50; SD = 10.385). A paired-sample *t*-test was performed to determine whether the knowledge increase that was observed was statistically significant.

In order to test the primary research hypothesis, which was principally to evaluate the influence of the training intervention on participants' knowledge of research methodology, a paired-sample *t*-test was conducted, with the results indicated in Table 5.

As can be seen in Table 5, there was a statistically significant increase on the 99th percentile in respondents' knowledge from the first iteration (mean = 47.45; SD = 16.8; $t = -15.884$; $p \leq 0.000^{**}$) to the second iteration (mean = 54.78; SD = 10.385; $t = -28.750$; $p \leq 0.000^{**}$). The mean increase in knowledge was 7.33, with a 95% confidence interval ranging from 52.412 to 40.497 in the first iteration and 56.526 to 49.037 in the second iteration. The eta squared statistics indicated a large effect (0.96).

TABLE 4: Measures of central tendency for the variables measured.

Variable	Minimum	Maximum	Median	Mean	SD
Research knowledge before the intervention	8	73	50.00	47.45	16.80
Research knowledge after the intervention	28	68	57.50	54.78	10.385

SD, standard deviation.

TABLE 5: Paired-sample *t*-test results for research methodology knowledge.

Time	Paired differences		Mean error	95% confidence interval		<i>t</i>	<i>df</i>	<i>p</i> (2-tailed)
	Mean	SD		Lower	Upper			
Post-test	52.781	10.385	1.836	-56.526	-49.037	-28.750	31	0.000**

* , $p \leq 0.05$; ** , $p \leq 0.01$.

SD, standard deviation; *df*, degrees of freedom.

TABLE 6: One-way analysis of variance results for prior training as independent variable and previous and current research methodology knowledge as dependent variable.

Variable	Sum of squares	<i>df</i>	Mean squares	<i>F</i>	<i>p</i>
Previous research methodology knowledge (pretest)	780.125	1	780.125	2.836	0.103
Current research methodology knowledge (post-test)	133.963	1	133.963	1.252	0.272

* , $p \leq 0.05$; ** , $p \leq 0.01$.

df, degrees of freedom.

To determine whether previous training had a statistically significant influence on respondents' increase in knowledge, as reported in Table 4, an ANOVA was performed (as shown in Table 6).

As can be deduced from Table 6, prior training did not have a statistically significant influence on the previous and/or current research methodology knowledge of respondents. To examine this relationship further, a multiple regression analysis was performed to determine how much of the variance in current research methodology knowledge after the training intervention can be explained by prior training. The results of this analysis are displayed in Table 7.

As can be seen from Table 7, it would appear that the results displayed in Table 5 were verified because prior training did not predict current research knowledge as statistically significantly. More specifically, the model predicted 0.8% of the variance in current research methodology knowledge. It should be noted that the adjusted R^2 value expressed as a percentage was used as a result of the small sample size (Pallant, 2011). The most important aspect to note is that the relationship was negative. Thus, as research methodology knowledge increased, the influence of prior training decreased. Because of the small sample size, it is advised that the results be interpreted with caution. However, the direction of the correlation is not in accordance with the normal assumption, which was that various training courses would cumulatively increase an individual's knowledge base.

The abovementioned supposition is based on a study conducted by Hailikari, Katajauori and Lindblom-Ylänne (2008) which found that students who possessed relevant prior knowledge from previous training were likely to perform better on future related courses. On the contrary, the findings of Samuels et al. (2015) support the presented

TABLE 7: Multiple regression analysis results with research methodology knowledge after the training intervention as dependent variable and prior training as independent variable.

Variable	Prior training							
	<i>R</i>	<i>R</i> ²	Adjusted <i>R</i> ²	<i>F</i>	β	<i>t</i>	<i>p</i>	VIF
Research methodology knowledge	0.200	0.040	0.008	1.252	-0.200	-1.119	0.272	1.00

*, $p \leq 0.05$; **, $p \leq 0.01$.

R, *R*-value; *R*², *R*-square; *F*, *f*-value; β , beta; *p*, significance; VIF, variance inflation factor.

finding. The authors of the study mentioned above did an impact evaluation of a Grade R programme, with results indicating that an educational intervention had limited impact on later educational outcomes. Future research could, therefore, investigate a possible mismatch between theoretical and practical knowledge, specifically in terms of research methodology as subject matter as well as possible contextual factors which may influence the results. For example, would the application of research methodology differ in the context of higher education and within the public service as such?

Discussion and conclusion

According to the results illustrated in the preceding section, a 7.33% increase in research methodology knowledge occurred, which was statistically significant on the 99th percentile. A control was done to establish the influence of previous training, and it became evident that previous training was only responsible for 0.8% of the variance. However, the significance of the training intervention should be taken at face value, in that during the training course, information and knowledge was disseminated on a complex topic over a two-day period to a range of participants. Although the majority of the sample was in possession of a higher education qualification, only 44.8% of the respondents indicated that they had received previous training, whereas only one respondent held a Grade 12 qualification coupled with a diploma. Research methodology forms part of a higher education qualification as it is a critical cross-field outcome currently embedded in all higher education curricula. As such, critical cross-field outcomes are generic outcomes, which are the foundation of all teaching and learning and which all higher education students need to achieve (De Jager, 2004, as cited in Jonck, 2014:267). Hence, the assumption would be that most of the respondents would have had at least a basic understanding of the topic under investigation. Based on the results discussed, the primary research hypothesis was accepted, while the secondary research hypothesis was rejected. Furthermore, alternative explanations for the findings could theoretically include (1) micro-level situational factors, for example, intrinsic motivation (i.e. participants were highly motivated and outcomes could be ascribed to the unique characteristics of participants), (2) the training emphasised aspects that were later assessed, which could be assumed as the pretest–post-test design would be related to the content of the courses (i.e. respondents would have been familiar with the items covered in the assessment) and (3) the facilitation style(s) or personality of the facilitator(s) could have played a role. As far as could be established, similar findings have not previously been reported. Samuels et al. (2015) did an impact evaluation of a Grade R programme. Mashalaba et al. (2015) did an implementation evaluation of the business process services

(BPS) incentive programme undertaken by the Department of Trade and Industry. The approach adopted in the aforementioned study did not correspond to the approach in this study as a cost-competitiveness analysis was utilised.

In accordance with the objective of a formative outcomes evaluation under the ambit of an impact assessment, it is recommended that the research methodology training intervention be sustained because the objectives were achieved. Moreover, the framework utilised should be used as a benchmark for best practice in the capacity-building sphere. In terms of the practical significance, this study is only the first step in empirically investigating ways to determine the efficacy of training interventions and could also be used to stimulate debate regarding impact assessments with specific reference to capacity-building initiatives. Thus, it is recommended that the suggested framework and methodology be utilised in future research, as well as monitoring and evaluation endeavours covering various training interventions, in an effort to validate the current findings. Furthermore, future research could include a control group against which comparisons can be made.

Acknowledgements

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Authors' contributions

P.J. conceptualised the article and was responsible for the development of the measuring instrument and the data analysis as well as writing the article. P.S.R. and R.D.C. assisted with data collection as well as compiling the article. All parties were involved with the facilitation of the workshop.

References

- Abrahams, M.A. (2015). A review of the growth of monitoring and evaluation in South Africa: Monitoring and evaluation as a profession, an industry and a governance tool. *African Evaluation Journal*, 3(1), a142. <https://doi.org/10.4102/aej.v3i1.142>
- Babbie, E. & Mouton, J. (2011). *The practice of social research*. Cape Town: Oxford University Press.
- Bank, J. (2012). *Theory-based approaches to evaluation: Concepts and practices*. Retrieved March 02, 2017 from <https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/centre-excellence-evaluation/theory-based-approaches-evaluation-concepts-practices.html>
- Behrman, J.R. (2010). The International Food Policy Research Institute (IFPRI) and the Mexican PROGRESA anti-poverty and human resource investment conditional cash. *World Development*, 38(10), 1473–1485. <https://doi.org/10.1016/j.worlddev.2010.06.007>
- Colquitt, J.A., & Simmering, M.J. (1998). Conscientiousness, goal orientation, and motivation to learn during the learning process: A longitudinal study. *Journal of Applied Psychology*, 83, 654–665. <https://doi.org/10.1037/0021-9010.83.4.654>

- Davies, P. (2004). Is evidence-based government possible? Jerry Lee lecture presented at the 4th Annual Campbell Collaboration Colloquium (pp. 1–29). Washington, DC. Retrieved from <http://web.archive.nationalarchives.gov.uk/20091013084422/http://www.nationalschool.gov.uk/policyhub/downloads/JerryLeeLecture1202041.pdf>
- Department of Planning, Monitoring and Evaluation. (2007). *Policy framework for the government-wide monitoring and evaluation system*. Pretoria: The Presidency.
- Department of Planning, Monitoring and Evaluation. (2014). *Overview paper: Evidence-based policy-making and implementation*. Pretoria: The Presidency.
- De Souza, A. C., Alexandré, N. M. C., & Guirardello, E. D. B. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Applications of Epidemiology*, 26(3), 1–10.
- Dimitrov, D. M., & Rumrill, P. D. (2003). Pretest-posttest designs and measurement of change. *Work*, 20, 159–165.
- Frei, R. (2011). A complex systems approach to education in Switzerland. In T. Lenaerts, M. Giaconini, H. Bersini, P. Bourguine, M. Dorigo & R. Doursat (Eds.), *Advances of artificial life ECAL* (p. 242–249). MA, USA: Massachusetts Institute of Technology.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L.B., & Vermeersch, C. M. J. (2011). *Impact evaluation in practice*. Washington, DC: The World Bank.
- Goujon, A., Lutz, W., & Wazir, A. (2011). *Alternative population and education trajectories for Pakistan*. International Institute for Applied Systems Analysis (IIASA), Austria Interim Report IR-11-029. Retrieved November 01, 2017 from <http://www.iiasa.ac.at/research/POP/staff/index.html?sb=31>
- Hailikari, T., Katajavuori, N., & Lindblom-Ylänne, S. (2008). The relevance of prior knowledge in learning and instructional design. *American Journal of Pharmaceutical Education*, 72(5), 113. <https://doi.org/10.5688/aj7205113>
- Jonck, P. (2014). A human capital evaluation of graduates from the Faculty of Management Sciences employability skills in South Africa. *Academic Journal of Interdisciplinary Studies*, 3(6), 265–274. <https://doi.org/10.5901/ajis.2014.v3n6p265>
- Leahey, E., & Moody, J. (2014). Sociological innovation through subfield integration. *Social Currents*, 1(3), 228–256. <https://doi.org/10.1177/2329496514540131>
- Lu, C.H. (2014). Assessing construct validity: The utility of factor analysis. *Statistical Year Book*, 15, 79–94.
- Mashalaba, N., Wyatt, A., Mathe, J., & Singh, R. (2015). Implementation evaluation of the business process services incentive programme. *African Evaluation Journal*, 3(1), 1–12. <https://doi.org/10.4102/aej.v3i1.146>
- Mouton, C. (2010). *The history of programme evaluation in South Africa*. Unpublished master's thesis, University of Stellenbosch, Stellenbosch.
- National Development Plan 2030. South Africa. (2012). Pretoria: Sherino Printers.
- National Evaluation Plan, South Africa. (2012). Retrieved from [http://www.dpme.gov.za/publications/Policy%20Framework/National%20Evaluation%20Policy%20Framework%20\(NEP\).pdf](http://www.dpme.gov.za/publications/Policy%20Framework/National%20Evaluation%20Policy%20Framework%20(NEP).pdf)
- O'Malley, G., Perdue, T., & Petracca, F. (2013). A framework for outcome-level evaluation of in-service training of health care workers. *Human Resources for Health*, 11(50), 1–12. <https://doi.org/10.1186/1478-4491-11-50>
- Pallant, J. (2011). *SPSS survival manual: A step by step guide to data analysis using SPSS* (4th edn.). Sydney, Australia: Allen and Unwin Publishers.
- Phillips, S. (2012, February/March). The Presidency: Outcome-based monitoring and evaluation approach. *PSC News*, p. 13–15.
- Pillay, P., Juan, A., & Twalo, T. (2012). *Impact assessment of national skills development strategy II: Measuring impact assessment of skills development on service delivery in government departments*. Pretoria: Human Sciences Research Council (HSRC) & Development Policy Research Unit, UCT.
- Rogers, P. (2014). *Methodological brief, no 1: Overview of impact evaluation*. Florence, Italy: United Nations Children's Fund (UNICEF).
- Rogers, P.J. (2012). *Introduction to impact evaluation (Impact Evaluation Notes No. 1)*. Washington, DC: The Rockefeller Foundation.
- Samuels, M., Taylor, S., Shepherd, D., Van der Berg, S., Jacob, C., Deliwe, C.N., & Mabogoane, T. (2015). Reflecting on an impact evaluation of the Grade R programme: Method, results and policy responses. *African Evaluation Journal*, 3(1), 1–10. <https://doi.org/10.4102/aej.v3i1.139>
- Ureda, J., & Yates, S. (2005). A systems view of health promotion. *Journal of Health and Human Service Administration*, 28(1), 5–38.
- Wagner, C., Kawulich, B., & Garner, M. (2012). *Doing social research: A global context*. Berkshire: McGraw-Hill Higher Education.
- Weyrauch, V., & Langou, G. D. (2011). Sound expectations: From impact evaluations to policy change. *International Initiative for Impact Evaluation, Working paper 12*. New Delhi: 3ie.
- White, H. (2009). Theory-based impact evaluation: Principles and practice. *International Initiative for Impact Evaluation, Working paper 3*. New Delhi: 3ie.
- Yi, M.Y., & Davis, F.D. (2003). Developing and validating an observational learning model of computer software training and skills acquisition. *Information Systems Research*, 14(2), 146–169. <https://doi.org/10.1287/isre.14.2.146.16016>
- Zwar, N.A., Weller, D.P., McClaughan, L., & Traynor, V.J. (2006). Supporting research in primary care: Are practice-based research networks the missing link? *The Medical Journal of Australia*, 185(2), 110–113.