

**Remaining relevant: Reflections on critical incidents in psychological test development
in South Africa**

Unisa inaugural lecture by Prof René van Eeden of the Department of Psychology

Presented on 1 November 2018

Abstract

In this presentation, critical incidents in the South African context that affected test development policies and practices are considered. Anecdotes and personal perspectives illustrate the response to these incidents and I highlight the implications thereof for the test development domain. The presentation is structured in terms of three topics, namely infrastructure and ideology, the challenge of regulation and training as a shared responsibility, and methodological developments and conceptual concerns. The emphasis is on relevance in the local context – an issue that continues to shape this domain.

Professor Moche, the Vice-principal for Teaching, Learning, Community Engagement and Student Support, Professor Phillips, the Executive Dean of the College of Human Sciences, Prof Mojapelo-Batka, the Chair of Department for Psychology and also the respondent tonight, colleagues, family and friends – good evening.

Introduction

Charlie Brown is a model neurotic. He is prone to depression and anxiety and paralyzing fits of over-analysis. ... Snoopy is a typical extrovert. Flamboyant, daring, and outgoing to a fault, he tries to join in every activity and conversation. He (perhaps fictitiously) flies gallant missions against the Red Baron and then brags about his exploits.... Defined by a single word (crabby), Lucy revels in her disagreeableness. Typical portrayals of Lucy feature her bossing around her friends, dominating her little brother, [and] mocking Charlie Brown's self-consciousness.

In this extract, the Big Five model of personality is used to categorise the characters from the Peanuts comic strip (Kaufman, 2010, pp. 1-2). There are numerous examples of such an

analysis of these fictitious characters; these reflect our need for a better understanding of our own behaviours.

Efforts to assess human attributes, describe individual differences and predict behaviour have a long and at times esoteric history – think astrology or palmistry. Assessment as a discipline was first proposed at the end of the seventeenth century (Foxcroft, Roodt, & Abrahams, 2013a) and it continued to develop as a scientific method. Events in the United States and Europe in the early twentieth century firmly established objective measurement, or testing, as part of the discipline of psychology. Such assessment remained popular until World War II, after which resistance against the practice developed in the United States and some European countries (JvR Psychometrics, 2018a). Sentiment again shifted, though, and for the past fifty years practitioners have regarded psychological assessment measures as indispensable in ensuring informed decision-making in the case of individuals, groups and organisations. The topic of psychological assessment in multicultural contexts has dominated this period, both internationally and in South Africa. One should note that the term psychological assessment refers to a multifaceted process in which psychological tests, the focus of this presentation, is but one source of information.

Test development and test use in the South African context have been and continue to be influenced by the country's socio-political history. The question of testing has consequently been one of the most controversial areas in the local history of psychology. Various authors discuss how an environment characterised by unequal resources and opportunities, that were based on racial categorisation, negatively influenced assessment related research and practice (e.g. Foxcroft et al., 2013a; Laher & Cockcroft, 2013). As early as 1916, local adaptations of international tests of cognitive ability were undertaken (JvR Psychometrics, 2018b). Practices in the ensuing years are reflected in the well-documented controversy around the multicultural use (and often misuse) of assessment measures (Foxcroft et al., 2013a). Socio-political changes resulted in a shift towards more inclusive test-development policies and practices during the 1980s under the auspices of the Human Sciences Research Council (HSRC). The HSRC was at that stage the main psychological test provider in South Africa (JvR Psychometrics, 2018b).

My presentation focuses on changes in the local psychological test development domain over the past three decades, that is since the mid-1980s. Anecdotes and personal perspectives

illustrate the response to critical incidents in this domain during this period. I discuss the implications for policies and practices in test development. This is not an exhaustive discussion of the local history of testing. However, it does show how the need for relevance shaped this domain, enabling it increasingly to make positive contributions towards a more inclusive society. In a passionate and proactive manner, role players have been exerting themselves to keep this domain relevant.

The presentation is structured in terms of three topics. First, *infrastructure and ideology* deals with the decentralisation of test development when the state-supported HSRC relinquished its role as test publisher. This responsibility shifted to multiple role players, implying a change in test-development infrastructure. Changes in the collection of available tests have been slow, but policies and practices have been shifting towards greater inclusivity.

In the second place, *the challenge of regulation and training as a shared responsibility* refers to the fact that statutory control became more dynamic owing to the increased involvement by different stakeholders. Professional bodies and societies have been active in the development of local policies and the implementation of international guidelines. Academic institutions, furthermore, are responding to concerns regarding sufficient and relevant training opportunities.

The third topic, *methodological developments and conceptual concerns*, illustrates the advances made in the field, but also the complexity of efforts to contribute constructively. Despite the use of sophisticated techniques to ensure equivalence, the influence of socio-cultural factors on test performance continues to pose challenges to assessment in diverse contexts. Stakeholders acknowledge the value of testing in the assessment process. However, a shift from traditional approaches has been proposed and recommendations have been made for indigenous tests that are more suitable to the local context.

Infrastructure and ideology

During the 1980's and 1990's, the resources available at the HSRC for the development of psychological assessment measures and for engaging in related research were indeed impressive. Various research teams specialised in different types of tests. Highly experienced researchers served as mentors to juniors who were in many cases recruited through a bursary

programme. In addition to a comprehensive research library, a test library that also served as an archive of the history of testing in South Africa was in operation. HSRC publishers and printers were responsible for the production of test material and related publications, while research teams had access to statisticians and programmers. These specialists were instrumental in the development of an array of statistical techniques and software related to the field of assessment.

The HSRC undertook large scale test standardisation projects in educational settings where the infrastructure included support by the educational departments. A historical disparity in access to psychological services between the privileged sector of the South African school population and the non-privileged majority has been documented (e.g. Maree, 2013; Setshedi, 2008). Psychological assessment services were nevertheless an integral part of the functions of each of the four racially segregated education departments. These departments provided support for related research in the schools under their jurisdiction. Psychologists from the educational aid centres carried out the testing together with a large team of HSRC researchers. For example, in the case of the standardisation of the Senior South African Individual Scale – Revised (SSAIS-R), a sample of 6 000 English- and Afrikaans-speaking pupils aged 7 to 16 years were tested over a three year period (Van Eeden, 1991). From 1986 to 1988 researchers would take part in test administration either locally or on extended fieldwork trips to other provinces. There was obviously no shortage of volunteers for these trips, especially if Durban or Cape Town was the destination. Such a project involved a complex set of logistics – and this was but one of the tests being standardised at that stage. The partial loss of this infrastructure had implications for the renorming of existing tests and the standardisation of additional measuring instruments for the school-age population: substantial samples are needed for this population to ensure age representation. Obsolete norms are consequently at present a serious concern in the case of this age group. This limits the degree of confidence with which practitioners can accept a score on an ability test as a true reflection of ability.

Regular testing in industry preceded the establishment of national institutes such as the HSRC (JvR Psychometrics, 2018b). Applicant pools at organisations continued to be a source of data in research requiring adult samples. The standardisation of the Wechsler Adult Intelligence Scale – III (WAIS-III) in the mid-1990s illustrates the involvement and support of various stakeholders from private practices, higher education and industry (Claassen,

Krynauw, Patterson, & wa ga Mathe, 2001). A symposium on related topics informed the research strategy, while the test was administered by individuals who did so on a voluntary basis. The sample comprised 1 300 respondents, and 138 individuals were acknowledged for their contribution during the testing for this project. Standardisation of an individual intelligence test requires substantial resources. The successful use of external support in the Wechsler project boded well for research on other types of tests – especially after the decentralisation of test development that was soon to follow.

Despite a shift in test related research policies, the available assessment measures continued to be perceived as a legacy of apartheid. Foster, Nicholas, and Dawes (1993) described this as keeping a system in place. Resistance cumulated in the moratorium on testing in school settings in 1995 – especially in group format (Maree, 2013) – and in the temporary ban on testing for organisational selection in 1998 (JvR Psychometrics, 2018b). The HSRC shifted its research focus to redress inequity and relinquished its role as test developer and distributor in 2003. Local branches of international test agencies as well as local test agencies had at that stage already been operational for a number of years and thus took centre stage in the test development arena. Involvement by universities and industry per se became more important. In addition, individual researchers continue to contribute to the development of local tests. Furthermore, postgraduate students often select a topic on the feasibility of tests in a variety of contexts. Valuable information is collected in this manner.

Test publishers have been faced with the dilemma of accommodating research demands within a business model. Many agencies act as distributors for a variety of international tests. The local adaptation of signature tests is a priority and resources are made available for these projects. However, this cannot be carried out for all tests in their catalogues. The tests are furthermore often only suitable for an organisational context; studies by Foxcroft and others indicate that users in other contexts have continued to rely on the existing range of locally developed or adapted tests (Foxcroft, Paterson, Le Roux, & Herbst, 2004; Van Eeden, Van Deventer, & Erasmus, 2016). There is insufficient capacity to update these tests and to establish norms suitable in diverse contexts.

I want to conclude this section on infrastructure and ideology with a quote from Laher and Cockcroft (2014, p. 7): “... although not without its critics, the HSRC was one of the most

productive agencies for change in psychological assessment in South Africa and provided the groundwork for where the field stands today.”

The challenge of regulation and training as a shared responsibility

Statutory control of psychological assessment measures restricts the use of these measures to registered professionals in psychology. The Health Professions Act (No. 56 of 1974) defines the use of these tests as a psychological act, implying a need for appropriate registration (Foxcroft, Rood, & Abrahams, 2013b). The Professional Board for Psychology is presently responsible for the classification of a measuring instrument as a psychological test; subsequently, the test is subject to legal restrictions. There have, however, been calls for an independent test accreditation body. Stakeholder groups have been formed to address this and related issues.

Additional statutory measures apply in educational and organisational sectors respectively. In the educational context, Article 5.(2) of the South African Schools Act (SASA 84 of 1996) prohibits testing as part of the admission process in public schools. Routine group testing is also no longer common practice. These changes reflect policies on inclusive education. Multi-disciplinary professional teams continue to serve the school environment but the availability of specialist assessment is often hampered by a lack of resources. In addition, the contribution of the educational psychologist can no longer be primarily focussed on diagnostic assessment but should include the empowerment of educators to provide developmental opportunities based on assessment results (Foxcroft et al., 2013a). According to Kanjee (2013), large scale assessment tailored for educational reform could empower policy makers as well as teachers.

Testing in industry has been affected by the Employment Equity Act (No. 55 of 1998, section 8) which states that testing is prohibited unless the test or assessment being used is reliable, valid, fair and unbiased for all employees. By the mid-1990s testing was being viewed with suspicion and trade unions were opposed to testing. A paper presented in 1995 by Minister Blade Nzimande entitled *To test or not to test?* was a turning point. Copies were circulated to researchers at the HSRC to create greater awareness of the impending changes and the impact on the research we were involved in. Stakeholders accepted the challenge and became involved in working towards inclusive and responsible practices. People Assessment in

Industry (PAI), an interest group of the Society for Industrial and Organisational Psychology of South Africa (SIOPSA), was established. This group, together with others, gave input to government on the wording of the Employment Equity Act. The Association of Test Publishers of South Africa (ATP) was established in 2005 (Kriek, 2013). This association successfully reversed certain legal restrictions (e.g. JvR Psychometrics, 2018b), resulting in a wider range of measuring instruments that could be used in industry and allowing for administration of some instruments by non-registered individuals. The Psychological Society of South Africa (PsySA) was also, and continues to be, an active agent of change.

The various bodies and societies encourage the use of a vast number of international as well as locally developed guidelines and standards (Davies, Foxcroft, Griessel, & Tredoux, 2013; Foxcroft et al., 2013b). Documents by the International Test Commission (ITC) include guidelines for adapting tests, for test use, for computer-based and internet-delivered testing and even a test-taker's guide to technology-based testing. The International Organisation for Standardisation (ISO) published a standard for assessment in organisational settings, while assessment in multicultural populations is discussed within a document by the American Psychological Association (APA). Locally developed documents on the validation of assessment measures and assessment in the workplace are also available from SIOPSA. Efforts are currently underway to tailor international guidelines for the African context. Guidelines regarding computer-based and internet delivered testing deal with issues of copyright, security, confidentiality, authenticity and ways of providing for differences in the level of digital literacy. This type of testing complicates the issue of who is allowed to administer tests. Despite the highly regulated test environment in South Africa, the practical implementation thereof remains a challenge.

There is also increasing concern regarding training and the implications for capacity. The Professional Board for Psychology accredits professional training programmes while the Psychometrics Committee of the Board provides assessment related training guidelines. Capacity building in terms of test development and related research skills is recognised in the Board's requirements for the training of psychometrists as well as the training of research psychologists. Initiatives by academic institutions, however, highlighted a disconcerting trend. Despite the research and development agendas of the various test agencies, the decentralisation of test development is gradually leading to a shortage in researchers who have the expertise to meet the country's needs for tests that benefit the development of a

diverse population. In 2014, the Stellenbosch Department of Industrial Psychology held a workshop with educators from various universities to discuss assessment related training (Malan, 2014). Amongst others, a need for comparability in curricula and the promotion of psychometric literacy among professionals were discussed.

In 2015, the Unisa Department of Psychology published research pertaining to the decrease in the number of university departments that offer training for registration as a psychometrist (Van Eeden, et al., 2016) and the fact that opportunities for training seem to be increasingly limited to the organisational sector. The training presented by Unisa imparted an overwhelming sense of how broad the need for psychological assessment is. Internship placements included pro bono work with children in rural Kwa-Zulu Natal, career counselling at universities and universities of technology, assessment at drug rehabilitation centres, and of course large scale testing for personnel selection and placement. It is clear that the need to equip students with skills to deal with diversity and to cope in low-resourced settings is non-negotiable, regardless of the level of registration. The range of service delivery in the mental health sector has been and is a concern, with large sectors of the community being under resourced. This is also true for assessment related services, and research on the availability of pro bono networks has been proposed.

Methodological developments and conceptual concerns

To ensure equality in terms of educational and work related opportunities, there was a need to develop tests that were not constructed along racial lines and to establish norms that were not racially based (Claassen, 1998). The term equivalence as used in the context of multicultural testing has to be broadly defined to include various forms of conceptual and statistical equivalence. In simple terms, “[statistical] equivalence ... refers to a lack of bias that is due to group-related effects [that is] a lack of group-related bias” (De Kock, Kanjee, & Foxcroft, 2013, p. 90). Initial studies therefore considered the bias in existing tests (e.g. Owen, 1989) and this focus was followed by comparative studies on tests developed for more than one population group (e.g. Claassen, 1990; Prinsloo, 1992; Van Eeden & Visser, 1992). Sophisticated models and techniques proposed by authors such as Van de Vijver and Poortinga (e.g. 2004) and Hui and Triandis (e.g. 1985) were diligently explored as options to identify the level of equivalence in diverse contexts. The use of terms such as differential

item functioning, congruence in factor structures and predictive equivalence became a permanent feature of test related research.

It is, however, the methodological design and the interpretation of results that allow a researcher to comment on conceptual equivalence. That is, are we measuring the same thing regardless of context? Socio-cultural influences on test performance consist of a number of variables that are closely related and therefore difficult to separate. Heterogeneity in terms of these variables results in between-group as well as within-group differences. This makes it difficult to identify a stable “majority culture” and thus identify meaningful reference points for the comparison of an individual’s score on a test. Comparison to a national sample implies a strong possibility of under- or over-estimation of an individual’s performance based on subgroup status. This applies to all types of tests, but the problem is especially salient in the case of ability tests. The complexity of test standardisation in multicultural populations has been explored in depth in the ongoing debate related to the standardisation of the Wechsler scales of intelligence in South Africa (Shuttleworth-Edwards, 2016; 2017; Sunderaraman, Zahodne, & Manly, 2016; Taylor, 2016). Researchers used equal representation of racial groups (Claassen, Krynanuw, Patterson, & wa ga Mathe, 2001) as well as proportional representation (Taylor, 2016) to compile adult normative samples. However, neither approach adequately addresses the need for optimum functioning of a test in a multicultural context. The influence of socio-cultural variables, especially language proficiency and education, continues to complicate issues.

There are numerous examples of studies illustrating the predicament. Socio-economic status was for the first time formally introduced as a norming variable in the early 1990s when group and individual intelligence tests for the school-aged population were released (e.g., Van Eeden, 1991). It is debatable if such a distinction could still be meaningfully applied. Although ethnicity per se is not a meaningful norming category, the potential influence of culture was acknowledged in studies on the Sixteen Factor Personality Questionnaire (16PF). This stimulated a debate regarding the interpretation of differences in performance by different population groups. On the one hand, the nature of these differences was regarded as discriminatory by authors such as Abrahams and Mauer (1999), whereas advocates for the continued use of the test argued in favour of ongoing research about potential socio-cultural explanations for these differences, for example Prinsloo and Ebersöhn (2002). In support of the latter, it was established that there are cultural differences in response patterns, with

moderate response styles (i.e., a preference for middle response options) being found for collectivist cultures. Personal experience with the test showed that the situation was complicated: cultural response patterns seemed to be more or less pronounced depending on the level of education.

Language might well become a barrier in the assessment process so that, where possible, test instructions have been translated into the official South African languages. Examples include the Paper and Pencil Games (PPG) used during the foundational phase (Claassen, 1996) and the Learning Potential Computerised Adaptive Test (LPCAT) utilised for high school children and adults (De Beer, 2013). Test translation is, however, potentially problematic due to a lack of appropriate concepts required for equivalence in the target language. It is, moreover, difficult to translate idiomatic expressions, as we found in attempting a Venda translation of a personality questionnaire (Van Eeden & Mantsha, 2007). One should also consider the argument that assessing individuals in an African language might disadvantage them where the language of education is English (Van den Berg, 1996).

To further complicate matters, a number of studies found the effects of language proficiency and education to be interrelated (e.g. Claassen et al., 2001). Stratification in terms of level and quality of education clearly indicated the crucial role of this variable in performance on the Wechsler Intelligence Scale for Children - Fourth Edition (WISC-IV) (Shuttleworth-Edwards, Gaylard, & Radloff, 2013; Shuttleworth-Edwards, Van der Merwe, Van Tonder, & Radloff, 2013). These studies focussed on the use of the test for neuropsychological purposes. The authors recommended that the individual's performance should be compared to that of a group that was very similar in terms of important demographic variables such as the level of education. This solution is warranted in the case of specialised contexts and if it is applied by professionals with a high degree of expertise – it does not resolve the dilemma of assessment in the broader population. And, although situation-specific norms are not uncommon in an organisational context (Patterson & Uys, 2005), the use of demographic criteria to form separate norm groups could be perceived as a repetition of earlier discriminatory practices.

In addition to the need for inclusive norms, there is the dilemma of outdated norms. The Flynn effect refers to an increase in general IQ scores of a population over time. The average gain for Western, industrialised countries is three IQ points per decade; Te Nijenhuis, Murphy and Van Eeden (2011) also identified gains for South African populations.

Therefore, if test norms are not regularly updated, the measured performance of an individual will become artificially inflated. Outdated norms, especially in the case of ability tests, have become one of the major concerns in the assessment context. I have earlier illustrated the scope of the study aimed at standardising the SSAIS-R for children. This type of project is no longer practical and test developers often have to rely on convenience samples.

Innovative developments in computerised testing during the 1980s and 1990s included the use of computerised adaptive testing or tailored testing. During the testing session, items are interactively selected to match the individual's level of ability, implying a unique set of items for each individual (De Beer, 2013). This form of assessment has successfully been used in the LPCAT, a non-verbal learning potential measure. De Beer and her co-workers also explored the value of tailored testing for assessment of constructs such as interests; the development of a local interest questionnaire, based on this model, is nearing completion. According to Laher and Cockcroft (2014), the measurement of learning potential is regarded as one of the most promising solutions to the challenge of equitable testing, especially in low-resourced settings. The principle of dynamic assessment used in these tests, allows for the measurement of the current level of performance as well as the projected future level of performance – the learning potential. Training is incorporated in the assessment to provide for potential differences between individuals in terms of educational and socio-economic opportunities.

Most of the research referred to has been based on international tests adapted and standardised for local use, or on locally developed tests grounded on existing models of testing. The ideal in a multicultural context would be to include only experiences common to all the different groups. The importance of content familiarity has been recognised, as noted in African-inspired item content (e.g. Bekwa, 2016) and the use of material that is readily available in a low-resourced setting (Laher & Cockcroft, 2017). The latter includes sand-play and the use of clay and beads as part of the assessment process. The South African Personality Inventory (SAPI) (Hill, Nel, van de Vijver, Meiring, Valchev, Adams, et al., 2013, p. 2) affords an example of an indigenous test. The aim of an indigenous or emic approach is to “develop an insider's perspective of psychological phenomena in a culture”. “[C]ulturally and linguistically adequate” personality descriptors were identified for each of the 11 language groups in South Africa. Analyses of these descriptive terms resulted in a test that measures nine personality constructs including universal factors as well as social-

relational factors specific to more collectivist cultures. This is, however, a lengthy and complicated solution and there seems to be insufficient capacity for a broad application of this methodology.

Regardless of the methods used, the interpretation of results should always be contextually meaningful (Shuttleworth-Edwards, 2014) and integrated with information from all relevant sources. Contextual approaches to assessment in multicultural contexts tend to be more qualitative. For example, the practitioner could use qualitative indicators together with quantitative scores in assessing school readiness so as to obtain a more comprehensive picture of the child (e.g. Theron, 2013). Narrative techniques have also been used in career counselling (e.g. Maree, 2013). Knowledge of relevant cultural factors remains a prerequisite when using these approaches. Continuous research in applied contexts keeps the practitioner updated on the utility of different tests in these contexts. My own research and that of my postgraduate students illustrate the broad areas of application. These studies include the use of leadership and personality questionnaires as part of an exploration of the dynamics within a management team. In a related study, psychometric data are being used amongst others to describe the construct of spiritual leadership. The effectiveness of mindfulness practices can be partly determined by means of objective measurement, and questionnaires are also being used in a study on the imposter phenomenon amongst academics at a higher education institution. Considering health and well-being at a community level, a cognitive stimulation programme applicable in low-resourced settings has been developed for infants. The results on a related developmental test provided partial support for the effectiveness of this programme.

Conclusion

Psychological assessment is regarded as central to the profession of psychology. Consequently it is listed as a core competency in the scope of practice of professionals. In response to continued challenges, the various stakeholders responded with “a reaffirmation of the value of assessment, reconceptualisation of the scope of assessment, and innovations in the development of measures” (Foxcroft & Roodt, 2013, p. 291). Dialogue and debate amongst academics, test publishers and professional and legal bodies remain pertinent.

Nonetheless, the lack of resources implies that professionals are often obliged to rely on the available body of tests, many of which are still questionable in terms of cultural and linguistic appropriateness. The continued popularity of these tests (Foxcroft, Paterson, Le Roux, & Herbst, 2004; Van Eeden, Van Deventer, & Erasmus, 2016) may, for many, symbolise little change from past practices. Even when a test is scientifically valid and reliable, the method of assessment might still be perceived as negative by those being tested (Patterson & Uys, 2005). Educating the public will go a long way to disperse the myths surrounding psychological assessment.

Literature refers to the domain as a “vibrant field” (Laher & Cockcroft, 2014, p. 9), characterised by efforts to accommodate diversity in terms of socio-cultural factors including language and education, thus “actively contributing towards equity”. Acculturation in our country has not just been a one-way process and we are in a unique position to contribute to the body of research through the accommodation of indigenous knowledge systems. I have illustrated the pro-active manner in which researchers and other stakeholders have responded to critical incidents in the local history of psychological assessment. I am confident that their dedicated and sincere efforts will ensure the continued relevance of this domain.

Acknowledgement

Tonight is in a sense the culmination of my involvement in the domain of psychological assessment. I want to say thank you to all of you who shared my professional and personal life over this period. My family, my friends, my mentors, my colleagues and my students. I also want to acknowledge those whose place is empty tonight.

References

- Abrahams, F., & Mauer, K. F. (1999). The comparability of the constructs of the 16PF in the South African context. *Journal of Industrial Psychology, 25*, 53–59.
- Bekwa, N.N. (2016). *The development and evaluation of Africanised items for multicultural cognitive assessment*. Unpublished doctoral thesis, University of South Africa.
- Claassen, N.C.W (1990). The comparability of General Scholastic Aptitude Test scores across different population groups. *South African Journal of Psychology, 20*, 80-92.

- Claassen, N.C.W. (1996). *Paper and Pencil Games (PPG): Manual*. Pretoria: Human Sciences Research Council.
- Claassen, N.C.W. (1998). Cross-cultural measurement in the human sciences. In J. Mouton, J. Muller, P. Franks, & T. Sono (Eds.), *Theory and method in South African human sciences research: Advances and innovations* (pp. 209-237). HSRC Press: Pretoria.
- Claassen, N.C.W., Krynauw, A.H., Patterson, H., & wa ga Mathe, M. (2001). *A standardization of the WAIS-III for English-speaking South Africans*. Pretoria: Human Sciences Research Council.
- Davies, C., Foxcroft, C., Griessel, L., & Tredoux, N. (2013). Computer-based and internet-delivered assessment. In C. Foxcroft & G. Roodt (Eds.), *Introduction to psychological assessment in the South African context* (4th ed.) (pp. 213-229). Cape Town: Oxford University Press.
- De Kock, F., Kanjee, A., & Foxcroft, C. (2013). Cross-cultural test adaptation, translation and tests in multiple languages. In C. Foxcroft & G. Roodt (Eds.), *Introduction to psychological assessment in the South African context* (4th ed.) (pp. 83-106). Cape Town: Oxford University Press.
- De Beer, M. (2013). The Learning Potential Computerised Adaptive Test in South Africa. In S. Laher & K. Cockcroft (Eds.), *Psychological assessment in South Africa: Research and applications* (pp. 137-157). Johannesburg: Wits University Press.
- Employment Equity Act, number 55. (1998). *Government Gazette*, 400 (19370). Cape Town, 19 October.
- Foster, D., Nicholas, L., & Dawes, A. (1993). Psychology in South Africa: reply to Raubenheimer. *The Psychologist*, 6, 127-174.
- Foxcroft, C., Paterson, H., Le Roux, N., & Herbst, D. (2004). *Psychological assessment in South Africa: A needs analysis. The test use patterns and needs of psychological assessment practitioners*. Pretoria: Human Sciences Research Council.
- Foxcroft, C., & Roodt, G. (2013). What the future holds for psychological assessment. In C. Foxcroft & G. Roodt (Eds.), *Introduction to psychological assessment in the South African context* (4th ed.) (pp. 287-302). Cape Town: Oxford University Press.
- Foxcroft, C., Roodt, G., & Abrahams, F. (2013a). Psychological assessment: A brief retrospective overview. In C. Foxcroft & G. Roodt (Eds.), *Introduction to psychological assessment in the South African context* (4th ed.) (pp. 9-27). Cape Town: Oxford University Press.

- Foxcroft, C., Roodt, G., & Abrahams, F. (2013b). The practice of psychological assessment: Controlling the use of measures, competing values, and ethical practice standards. In C. Foxcroft & G. Roodt (Eds.), *Introduction to psychological assessment in the South African context* (4th ed.) (pp. 109-124). Cape Town: Oxford University Press.
- Hill, C., Nel, J.A., Van de Vijver, F.J.R., Meiring, D., Valchev, V.H., Adams, B.G. et al. (2013). Developing and testing items for the South African Personality Inventory (SAPI). *SA Journal of Industrial Psychology/SA Tydskrif vir Bedryfsielkunde*, 39(1), Art. #1122, 13 pages. <http://dx.doi.org/10.4102/sajip.v39i1.112>
- Hui, C.H., & Triandis, H.C. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-cultural Psychology*, 16, 131-152.
- Kanjee, A. (2013). Large-scale assessment studies in South Africa: Trends in reporting results to schools. In S. Laher & K. Cockcroft (Eds.), *Psychological assessment in South Africa: Research and applications* (pp. 516-534). Johannesburg: Wits University Press.
- Kaufman, J.C. (2010). The Charlie Brown theory of personality. *Psychology Today*, online. <https://www.psychologytoday.com/us/blog/and-all-jazz/201003/the-charlie-brown-theory-personality> (Retrieved 26/03/2018)
- Kriek, H. (2013, August). *Association of Test Publishers of South Africa (ATP)*. Paper presented to the Parliamentary Portfolio Committee for Labour, Cape Town.
- JvR Psychometrics, (Retrieved 27 March 2018a). *The histories project: History of personality assessments in the workplace*. Johannesburg: JvR Africa Group. <http://histories.jvrpsychometrics.co.za/history-of-personality-assessments-in-the-workplace/>
- JvR Psychometrics, (Retrieved 27 March 2018b). *The histories project: History of psychometric testing in South Africa*. Johannesburg: JvR Africa Group. <http://histories.jvrpsychometrics.co.za/history-of-psychometric-testing-in-south-africa/>
- Laher, S., & Cockcroft, K. (2013). Contextualising psychological assessment in South Africa. In S. Laher & K. Cockcroft (Eds.), *Psychological assessment in South Africa: Research and applications* (pp. 1-14). Johannesburg, SA: Wits University Press.
- Laher, S., & Cockcroft, K. (2014). Psychological assessment in post-apartheid South Africa: The way forward. *South African Journal of Psychology*, 44, 303-314. Doi: 1177/0081246314533634
- Laher, S., & Cockcroft, K. (2017). Moving from culturally biased to culturally responsive assessment practices in low-resource, multicultural settings. *Professional Psychology: Research and Practice*, 48(2), 115-121. <http://psycnet.apa.org/doi/10.1037/pro0000102>

- Malan, J. (2014). From the office of the chair. *Newsletter from the Department of Industrial Psychology, 14*, 1-3.
- Maree, J.G. (2013). Latest developments in career counselling in South Africa: Towards a positive approach. *South African Journal of Psychology, 43*(4), 409-421. DOI: 10.1177/0081246313504691 sap.sagepub.com
- Nzimande, B. (1995, June). *To test or not to test?* Paper presented at the Congress on Psychometrics, Council for Scientific and Industrial Research, Pretoria.
- Owen, K. (1989). *Test and item bias: The suitability of the Junior Aptitude tests as a common test battery for white, Indian and black pupils in Standard 7*. Pretoria: Human Sciences Research Council.
- Patterson, H., & Uys, K. (2005). Critical issues in psychological test use in the South African workplace. *South African Journal of Industrial Psychology, 31*(3), 12-22. DOI: <https://doi.org/10.4102/sajip.v31i3.204>
- Prinsloo, C.H. (1992). *Manual for the Sixteen Personality Factor Questionnaire South African 1992 version (16PF, SA92)*. Pretoria: Human Sciences Research Council.
- Prinsloo, C. H., & Ebersöhn, I. (2002). Fair usage of the 16PF in personality assessment in South Africa: A response to Abrahams and Mauer with special reference to issues of research methodology. *South African Journal of Psychology, 32*(3), 48–57.
- Setshedi, M.J. (2008). *Investigating the use of psychological assessment in South African schools*. MA dissertation, University of the Witwatersrand, Johannesburg.
- Shuttleworth-Edwards, A.B. (2014). Four fatal flaws of brain injury in the cross-cultural context. *Journal of Psychology in Africa, 24*(5), 464-467. <http://dx.doi.org/10.1080/14330237.2014.997022>
- Shuttleworth-Edwards, A.B. (2016). Generally representative is representative of none: commentary on the pitfalls of IQ test standardization in multicultural settings. *The Clinical Neuropsychologist, 30*(7), 975–998. <http://dx.doi.org/10.1080/13854046.2016.1204011>
- Shuttleworth-Edwards, A.B. (2017). Countrywide norms declared obsolete: best practice alert for IQ testing in a multicultural context. *South African Journal of Psychology, 47*(1), 3–6. <http://dx.doi.org/10.1177/0081246316684465>
- Shuttleworth-Edwards, A.B., Gaylard, E.K. & Radloff, S.E. (2013). WAIS-III test performance in the South African context: Extension of a prior cross-cultural normative database. In S. Laher & K. Cockcroft (Eds.), *Psychological assessment in South Africa: Research and applications*(pp. 17–32). Johannesburg: Wits University Press.

- Shuttleworth-Edwards, A.B., Van der Merwe, A.S., Van Tonder, P. & Radloff, S.E. (2013). WISC-IV test performance in the South African context: a collation of cross-cultural norms. In S. Laher & K. Cockcroft (Eds.), *Psychological assessment in South Africa: Research and applications* (pp. 33–47). Johannesburg: Wits Press.
- Sunderaraman, P., Zahodne, L.B., & Manly, J.J. (2016). A commentary on ‘generally representative is representative of none: pitfalls of IQ test standardization in multicultural settings’ by A.B. Shuttleworth-Edwards. *The Clinical Neuropsychologist*, *30*(7), 999–1005. <http://dx.doi.org/10.1080/13854046.2016.1211321>
- Taylor, N. (2016). Generally representative is generally representative: comment on Shuttleworth-Edwards. *The Clinical Neuropsychologist*, *30*(7), 1017–1022. <http://dx.doi.org/10.1080/13854046.2016.1213884>.
- Te Nijenhuis, J., Murphy, R., & Van Eeden, R. (2011). The Flynn effect in South Africa. *Intelligence*, *39*, 456–467. <http://dx.doi.org/10.1016/j.intell.2011.08.003>
- Theron, L.C. (2013). Assessing school readiness using the Junior South African Individual Scales: A pathway to resilience. In S. Laher & K. Cockcroft (Eds.), *Psychological assessment in South Africa: Research and applications* (pp. 60-73). Johannesburg: Wits University Press.
- Van De Vijver, F., & Poortinga, Y.H. (2004). Conceptual and methodological issues in adapting tests. In R.K. Hambleton, P.F. Merenda, & C.D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 39-63). Mahwah: Erlbaum. DOI: 10.4324/9781410611758
- Van Den Berg, A.R. (1996). Intelligence tests. In K. Owen & J.J. Taljaard (Eds.), *Handbook for the use of psychological and scholastic tests of the HSRC* (pp. 157-190). Pretoria: Human Sciences Research Council.
- Van Eeden, R. (1991). *Manual for the Senior South African Individual Scale – Revised (SSAIS-R). Part I: Background and standardization*. Pretoria: Human Sciences Research Council.
- Van Eeden, R., & Mantsha, T. R. (2007). Theoretical and methodological considerations in the translation of the 16PF5 into an African language. *South African Journal of Psychology*, *37*, 62–81.
- Van Eeden, R., Van Deventer, V., & Erasmus, H. (2016). The relevance of psychometrics as a professional resource: Training related issues. *SA Journal of Industrial Psychology/SA Tydskrif vir Bedryfsielkunde*, *42*(1), a1307. <http://dx.doi.org/10.4102/sajip.v42i1.1307>

Van Eeden, R., & Visser, D., (1992). The validity of the Senior South African Scale – Revised (SSAIS-R) for different population groups. *South African Journal of Psychology*, 22, 163-171.