Beyond the trade test: Using the COMET Model to build occupational competence

HELEN BROWN (hbrown@merseta.org.za) merSETA, PhD fellow: Institute for Post-School Studies, Faculty of Education, University of the Western Cape, Cape Town, South Africa
ORCID link https://orcid.org/my-orcid?orcid=0009-0002-3318-1277

JOY PAPIER (jpapier@uwc.ac.za) National Research Foundation Chair: TVET, Institute for Post-School Studies, Faculty of Education, University of the Western Cape, Cape Town, South Africa
ORCID link https://orcid.org/0000-0003-2079-9430

ABSTRACT

The South African trade test is a mandatory end-point assessment that certifies an apprentice to practise as a qualified artisan after a specified period of training. Whereas the manufacturing sector has relied on the traditional trade test to provide assurance of an artisan candidate's required level of competence, recent competence-development studies based on the COMET Model of competence development and measurement have challenged the ability of traditional task-based trade tests to prepare candidates adequately for integrated work processes. Studies in other contexts have shown the potential for COMET-based assessments not only to serve as a means of measuring competence, but also to develop it. This article reports on research that investigated how, through the application of COMET assessments, occupational competencies were developed beyond those measured by the traditional apprentice trade test. A mixed-methods, quasi-experimental approach produced strong evidence that COMET-inspired authentic assessments enhanced learners' levels of competence and developed vocational identity among candidates who undertook such preparation for a trade test.

KEYWORDS
Competence development; competence measurement; trade test; COMET; authentic assessment; apprenticeship

Introduction and context

Global industry and innovation

An industrial economy is associated with export promotion, increased trade openings, economic liberalisation and an improved business climate (Kniivilä, 2007), changes that are referred to as 'megatrends' (Achtenhagen & Winther, 2014:281). This world of work is constantly transforming through industrial development that has more recently been characterised by automation, connectivity and technological innovation. Industrial competence in a global context constitutes the engine of economic growth and employment of most nations (Haraguchi, Cheng & Smeets, 2017).

Technical and vocational education and training (TVET) is expected to respond to this global context, and although the occupational trade test is regarded as one of the key instruments of quality assurance in workforce development, it falls short of the requirements for a competent workforce, as reported in a recent South African case study (Hauschildt, 2016). Earlier studies by Grosse-Beck (1998) criticised the content and method of a trade test as having insufficient focus on company work processes due to their

- separation of skills and knowledge into written and practical parts of an examination;
- primary orientation towards theoretical knowledge;
- use of multiple-choice questionnaires; and
- failure to take into account the work process value chain.

Diagnostic assessments of competence, on the other hand, such as those reported on in this article, have recently started to provide a means of measuring levels of competence with sufficient validity and reliability in order to serve as an empirical basis for the planning, evaluation and measurement of competence development (Jenewein, 2017; Peterman, 2018; Rauner, 2017).

Occupational training system in South Africa

Skills development has been a key feature of South African policy over the past two-and-a-half decades (DHET, 2019). Some of the earliest legislation passed by the first democratically elected parliament focused on the complete reorganisation of education, training, and the apprenticeship system that were rooted in a racialised apartheid history (Gamble, 2021; Wedekind, 2013). More recently, the trade-examination system has been undergoing reform under the auspices of the Quality Council for Trades and Occupations (QCTO), but legacy trade tests continue to be administered through trade test centres (TTCs) across the nine South African provinces. A common trade test certificate for all qualifying artisan candidates has been issued by the QCTO since 2015 (QCTO, 2016).

The trade test is defined as

> a final integrated summative assessment for an artisan qualification in a listed trade, at an accredited trade test centre, by an assessor registered by the National Artisan Moderation Body (NAMB) (DHET, 2015).

This instrument relies on candidates having already achieved a domain-specific National Qualifications Framework (NQF) Level 3 qualification and a minimum of 80 weeks or a maximum of 208 weeks of workplace experience in all aspects of the curriculum before they apply to take a trade test. The trade test is conducted by administering trade-specific practical tasks in a controlled environment, at the end of which the candidate must be pronounced either competent or not yet competent for certification.

Prior to taking the trade test, candidates are encouraged by their training providers to complete a preparation course which is not standardised, but the provider of the course generally checks that the apprentice's logbook has been comprehensively completed and fills any gaps identified in the training for each trade test task.

The assessment approach for a trade test is based on the candidate being declared 'competent' or 'not yet competent' in each of seven tasks (in the case of electrician artisans) in order to be awarded the trade certificate issued by the QCTO. If candidates are found to be not yet competent in three or fewer tasks, they may carry credits towards another attempt on tasks not yet mastered; but if they do not achieve 'competent' status in four or more tasks, then all seven tasks must be tested again after a period of at least six weeks.

The resulting feedback is provided after all the required tasks have been completed; and if there is sufficient time, the examiner will explain the errors that contributed to a result of 'not yet competent'. After candidates have been declared competent in all the specified tasks, they are certificated as being able to practise and be remunerated as qualified artisans.

Notwithstanding the significance of the traditional trade test in the certification of artisans, the trade test as an instrument for assessing occupational competence has not been well researched in the South African context. Neither has the trade test been fully investigated as an assessment instrument that could possibly be used to develop domain-specific competence rather than only measuring it. A large-scale competence diagnostics study was undertaken in five engineering-related artisan occupations between 2013 and 2016, with low levels of competence being found among more than 1 200 candidates who formed part of that study (Jacobs, 2015). Data analysis suggested that the artisan trade test as a competence measure had been underestimated as a mechanism with which to promote the development of competence during the learning phases of an apprenticeship (Hauschildt, 2016) – the test having been conducted only as a final assessment for certification purposes. Therefore, the research reported on in this article was intended to investigate whether, and how, the structure and content of the trade test, in addition to being an end-point assessment, might influence

competence development *during* the learning phases of an apprenticeship. The methodology by which this was gauged was through the so-called COMET Model that is explained more fully in the sections that follow.

Towards a comprehensive understanding of competence

Assessment in TVET

Curtis (2010:6) describes assessment in TVET as a component of an 'ecosystem of skills', including their development and deployment by agents who operate in the social and industrial context for which the assessments are developed, certified and deployed. Assessment is therefore part of a larger structure of teaching and learning for a purpose, normally set out in a qualification and its policy. The way in which assessments – and examinations in particular – influence teaching and learning is commonly described as the 'washback' or 'backwash' effect (Pan, 2009:257–263). 'Washback' indicates 'an intended or unintended (accidental) direction and function of curriculum change on aspects of teaching and learning by means of a change of public examinations' (Cheng, 2005:8). On the one hand, positive washback integrates meaningful and innovative learning activities in teachers' educational methodologies, with the result that educators will devote more attention to students' intentions, interests and choices, and students are motivated to work harder. On the other hand, negative washback occurs when teachers teach only for the purposes of the test, narrowing the curriculum accordingly (Pan, 2009:261). The paradigm shift from the assessment 'of' learning to assessment 'for' learning has also brought diversity to educational practices, especially in the propagation of creativity and critical thinking among learners (Pattalitan, 2016).

Assessments are usually expected to produce comparable outcomes, with consistent standards being set over time. However, there are factors that impede both the validity and the reliability of assessment practices in workplace settings: for instance, the inconsistent nature of people; relying on assessors to make judgements without bias; changing contexts or conditions; and evidence of achievement arising spontaneously or incidentally. Public interest in the reliability of educational assessment as well as the complex nature of errors in assessment due to systemic factors or personal circumstances often present challenges for assessment (Gardner, 2013:72–92).

Competence development

Many of the central ideas that shaped the understanding of the development of competence emerged with increasing emphasis from the 1970s onwards – for instance, the theory of complete action. The circle of complete action was the first to challenge the Taylorist approach to the organisation of work in favour of recognising work as a process beyond functional actions of closed and repetitive tasks and requiring the essential element of communication. Boreham (2002) defined work process knowledge as the competence needed for modern workplaces that is characterised by increased functional flexibility, the use of

information and communications technology (ICT), the integration of previously separated production functions, and an emphasis on knowledge creation within normal work activity. Fischer and Boreham (2004) conducted empirical research on the concept of work process knowledge across 22 industrial sectors. Their research yielded three main defining characteristics. First, work process knowledge constitutes an understanding of a complete work process. Second, work process knowledge is used directly in the work process and is an instrumental part of work activity. Third, and finally, work process knowledge is constructed in the workplace itself by synthesising experiential and codified knowledge.

In seeking to understand the development of competence, Rauner, Hauschildt and Heinemann (2013:164) proposed a comprehensive analysis of competence in a four-stage model, where the highest level of competence is defined as 'holistic shaping competence' or

> the level of competence where occupational tasks are considered in their full complexity with due regard to the diverse operational and social conditions in which they are performed, and to divergent requirements in terms of work process and its outcome (Rauner et al., 2013:164).

Rauner (2017) and Rauner et al. (2013) held that holistic shaping competence, if measured as an outcome in TVET, could become a catalyst for finding new ways of teaching, learning and assessment that support industrial competitiveness, growth and innovation in an economy.

The definition of competence as 'context-specific cognitive dispositions that are acquired by learning and [are] needed to successfully cope with certain situations or tasks in specific domains' (Weinert, 2001) became the guiding competence construct for developing models so as to provide a basis for developing measuring instruments and interpreting their results (Hartig, Klieme & Leutner, 2008:10). At the time of the emerging PISA project (Programme for International Student Assessment), only a limited number of competence models existed to provide a basis for comparative measurement.

Competence models for assessment and learning

According to Nickolaus and Seeber (2013), there are three approaches to modelling vocational competence in industrial technical fields (Mulder, 2017:844–845):

1. approaches preferred by companies that use stated levels of competence in self-assessment or external assessment instruments (performance management and recruitment);
2. approaches which use pragmatic reasons for concentrating on professional competence in the narrow sense and modelling professional competence based on item response theories; and
3. holistic approaches that integrate professional, economic, social and creative aspects of professional competence.

Martens and Rost (cited in Deitmer, Hauschildt, Rauner & Zelloth, 2012:160), argue that

> [t]he measurement of occupational competence presupposes a theoretical and standards-based competence model, [and] accordingly, competence models have the following functions: firstly, to operationalise the fundamental criteria that have to be met in the context of problem-solving in the workplace; and[,] secondly, to provide sufficiently concrete guidelines for the formulation of test assignments.

The role of the competence model, according to Rauner (2017), is to connect the guiding principles and objectives of vocational education and the construction of tests and learning tasks. Three empirical studies using multidimensional models of competence have emerged over the past 10 years, illustrating this trajectory of critical enquiry:

1. Winther and Achtenhagen (2009) proposed a model of vocational competence with the achievement of vocational competence as the central goal. They defined levels of competence as conceptual, procedural and interpretative, all of which are governed by dimensions of complexity in modelling, cognition and content categories.
2. An alternative model of vocational competence was proposed by Klotz, Winther and Festner (2015), where a multidimensional model was developed to test 877 industry apprentices in a cross-sectional database using item response theory-based scaling. The resulting four-stage psychometric model represents a systematisation of the development of vocational competence; it is characterised by the degree of occupational specificity and different forms of cognitive processing.
3. The third model, the COMET Model of competence diagnostics, underpins the research reported on in this article. Therefore, this conceptual model is elaborated on in more detail in the next section.

COMET: A conceptual model for competence diagnostics

The COMET Model, with related test instruments and procedures, has been implemented in Germany, China, South Africa, Norway, Switzerland, Poland and Spain. Its implementation has resulted in a number of publications aimed at supporting TVET systems research. In some of the original conceptualisations of the model, the COMET acronym has been used variably to refer to 'competence measurement in education and training' and also to 'competence-based occupational methodology for effective training', depending on the focus of the application. Notwithstanding slight variations in the wording of the acronym, the overriding understanding of COMET is that it is a model for measuring and developing the competence outcomes of occupational qualifications (Rauner et al., 2013), qualifications that may also include higher-level professions. Most studies describe COMET as a diagnostic instrument that is used to assess or measure competence on a large scale. It possesses an implementation logic similar to that of the well-known PISA but is designed for whole occupations.

In seeking both to develop occupational competence and measure it, the model distinguishes between three dimensions of competence, namely the *requirement* dimension (incremental levels of professional competence based on skills that are associated with professional work tasks); the *content* dimension (teaching and learning in a specific subject as a basis for the development of test assignments); and the *action* dimension (a scientific foundation with which to measure 'complete professional action' … in favour of shaping complete professional action) (Rauner et al., 2013:41–53). The three dimensions are aimed at testing the specific requirements of a learning area across all occupations in the form of competence levels, while at the same time providing a guide to selecting specific content for the construction of test tasks. Building on the concept of work process knowledge and the theory of complete action, Rauner et al. (2013) argued that

> when the steps of a complete professional activity are related to the criteria for holistic solution of professional tasks, the concept of complete professional action is transformed into the category of complete (holistic) problem-solving, which is fundamental for the design of vocational training processes and the modelling of professional competence (cited in Deitmer et al., 2012:163).

In the learning context, the *action* orientation seeks to integrate theoretical knowledge and practical abilities through a reality-based, problem-related learning task rather than through closed and repetitive tasks (Argyris & Schön, 1997). In the context of professional work, the learning assignment and test tasks are designed to provide the space for both rational action and creative-dialogue type of action, which are fundamentally significant in all occupations.

The *content* dimension of the COMET Model relies on occupational fields in order to construct learning tasks and test assignments. Professional validity is a criterion for determining the content of tasks for the respective fields and therefore requires professional groups or expert reference groups to agree on the job description as a reflection of what true mastery looks like (Rauner, 2017:88). The content of learning and test tasks is structured so as to develop learners from novice to expert, with the curriculum content based on a systematic approach of defining relatively simple learning tasks first and then building complexity as the learner passes through progressive learning stages.

The *requirement* dimension builds on the action and content dimensions by defining four levels of competence, namely nominal, functional, processual and holistic shaping. These levels are based on the four-level proficiency model of Bybee (1997), which aimed to improve instructional practices to enhance student learning. Empirical research developed this concept further into six levels, which were applied in the Organisation for Economic Cooperation and Development's (OECD) project to measure competence in science during PISA 2006 (Bybee, McCrae & Laurie, 2009). The four levels of competence are described as follows:

1. *Nominal competence* reflects superficial conceptual knowledge of the field and individuals at this level can therefore not be considered competent. Indeed, learners at this level are considered a 'risk group' (Rauner et al., 2013).

2. *Functional competence* refers to basic technical knowledge learnt in isolation. It is the elementary subject knowledge and skills that have not yet been integrated and assimilated. The skill of integrating knowledge in order to solve process-related problems in an occupational task is therefore still very limited (Rauner et al., 2013).

3. *Processual competence* relates to the ability to interpret occupational tasks in relation to work processes and workplace situations. Aspects such as economic viability, customer focus and the expression of technical concepts in a clear and organised way through verbal accounts and technical drawings are evident in the solutions proposed for an occupational task (Rauner et al., 2013).

4. *Holistic shaping competence* is a level of competence where due regard is given to the diverse operational and social conditions in which an occupational task is performed, resulting in solutions that are uniquely different and valuable to the workplace organisation. This level of competence incorporates the possible influences of developments and innovations in technology in an occupational specification (Rauner, 2017).

These four competence levels are assessed using a Likert rating scale of 40 items mapped against eight criteria, with the total score indicating the level of competence achieved. The test instruments rely on a practical solution being arrived at to a dynamic workplace scenario. An extract from a scenario used in the study is briefly stated as follows:

> The training department in Company ABC requires an automated motor starter system to simulate a conveyor system that is used in plant operations for the transfer component parts from one station to another. The simulation is required for training purposes which runs 5 days a week (Monday to Friday) from 07H30 to 16H30. There is a 3-phase supply in the building and all the components required are available at the training store.

By way of comparison, the traditional trade test would set six to eight discrete practical tasks where candidates are declared either 'competent' or 'not yet competent' without reference to a particular work process. COMET test instruments are supported by an additional context questionnaire and a questionnaire on the learners' test motivation. And it is not only the learners' competence levels that are assessed through COMET: TVET lecturers and industry trainers are prepared in advance so that adjustments can be made to lesson plans so as to include COMET learning tasks beyond the basic formative assessment requirements. This ensures that teaching is adapted using the action, content and requirement dimensions to enhance learning before the COMET test assignments (Brown, 2015). In the process, abstract criteria are converted into measurable observations that enable data collection to be performed systematically. Each criterion is converted into an evaluation tool that guides the consistent application of the measurement for each of the COMET sub-competences.

The central constructs of the COMET model as described above not only provided a framework for moving empirically beyond accepted assessment practices in the traditional

trade test, but also laid the foundation for the research methodology that was employed in the research described in this article.

Research methodology

Research design

The study reported on here employed a mixed-methods, sequential explanatory design (Cameron, 2009) which connected quantitative and qualitative data collection. The logic of the research design was to measure the influence of the COMET-inspired methodology and assessment on two groups of artisan candidates, labelled A and B, as explained below.

Group A/Path A (the control group) comprised the artisan trade test candidates who were undergoing the standard preparation for taking the traditional trade test. Group A did not undergo COMET test preparation. The Group B/Path B (the experimental group) artisan candidates were introduced to the COMET model methodology and assessments as part of their preparation for taking the regulatory traditional trade test. All the candidates who passed the traditional trade test (i.e. could be certificated as qualified artisans) were subsequently invited to participate in the alternative COMET-inspired 'trade test' that followed. Essentially, then, the difference between the control and the experimental groups was that the former group was not exposed to the COMET test preparation while preparing for the regulatory traditional trade test, whereas the latter group did enjoy such exposure. Both Group A/Path A and Group B/Path B candidates took the traditional trade test that would indicate an exit level of 'competent/not yet competent'. Only those (in Groups A and B) who passed or were declared competent then took the COMET-inspired trade test and had their competencies measured in terms of the model.

The intention of this research design was to try to ascertain whether the learning and assessment methodology of the COMET-inspired tasks, undertaken by half the candidates (Group B) in preparation for the traditional trade test, would influence the competence outcomes when measured by the COMET-inspired alternative trade test. The competence outcomes (measured by the COMET-inspired trade test) of the experimental group would then be compared with those of the control group (Group A) which underwent only the regular learning preparation for the traditional trade test and for whom the COMET-inspired trade test would be an end-point assessment only.

Scope of the study and sample selection

The electrician and millwright trades were targeted for this study because these trades annually conduct the largest number of trade tests. In addition, the millwright trade includes the full electrician trade curriculum and therefore would ensure that more subject-matter experts would be available to participate in the study.

The four trade test centres with the highest number of electrician and millwright trade tests over a 20-month period were identified for participation in this study. Furthermore,

consideration was also given to the diversity of locations (both urban and rural) across provinces. For this reason, two centres were located in Gauteng, one in KwaZulu-Natal and one in the Eastern Cape.

Each trade test centre was asked to identify a minimum of 10 candidates who were approaching their trade test date (*n* = 40). Ten was seen as a reasonable number in relation to the cost and time involved on the part of staff and other resources needed to complete the fieldwork. For various reasons, of the group of 40 candidates, 14 were not successful in the traditional trade test and were therefore excluded from the groups who went on to take the subsequent COMET-inspired trade test. The final research sample across Groups A and B after nine months was therefore 26 electrical artisan candidates, who undertook the COMET-based trade test across the four selected trade test centres.

Quantitative data-collection process

Learning and assessment instrument development and validation
In line with the COMET Model methodology, a group of occupational subject-matter experts was formed in order to conceptualise a number of possible test tasks (for learning and assessment purposes) according to the eight COMET competence criteria. The tasks were then evaluated according to how often these tasks are performed in authentic work situations, the significance of the professional task to the occupation, the level of difficulty, and the significance of the task to personal professional development. Each subject expert used the same questionnaire to evaluate the tasks that had been developed in the group.

On the basis of the evaluation exercise, eight professional work-relevant tasks were selected for the study. The selected tasks were then presented to external professional practitioners for their comment on any industry-specific peculiarities in the description of the skilled work that might need to be amended. As a last step in the validation process, the test tasks were each rated out of 10 for their professional authenticity, representation of competence, and curricular validity.

The final set of tasks was subjected to a piloting process in which candidates completed the assessments and were rated by trained subject-matter experts with a view to establishing the potential of each professional task to describe competence in all of the eight criteria of the competence model. The piloting process validated four tasks for the COMET-inspired trade test. The expectation was that the problems posed by each task should be solved in the most professional way possible. The degree of complexity had to allow for the assessment of contextual understanding matched with the required level of practical skill. The grading of the test outcomes was ability-based; this made it possible to differentiate between test-takers according to the levels of the solutions they offered – whether they were (in terms of the model) functional, procedural or holistic in nature (Heinemann, Maurer & Rauner, 2010).

Implementation of the assessment instrument

Prior to the formal COMET-inspired trade test, candidates in Group B/Path B (the experimental group) were exposed to three COMET learning tasks over a period of six months, whereas candidates in Group A/Path A (the control group) completed only the traditional trade test preparation.

The formal assessment approach was then made up of two parts: the first part was dedicated to two written tests on validated COMET tasks, each separated by one to two months. The second part was dedicated to the practical implementation of a COMET-inspired trade test. The practical test was expanded into three segments over five days. Day 1 was dedicated to the documented conceptualisation and planning of a solution to the test task. This was done under the supervision of an examiner. Days 2 to 4 focused on the practical implementation of the plan or task and its quality control, supported by the candidate's documentation and explanations of any deviation(s) from the original plan. On Day 5 of the assessment, the candidates presented their solution, which was supported by an expert discussion with the examiners to justify the final result. This segment culminated in an agreed rating by two examiners so that the examination result could be fed back to the candidate artisan. The control group completed the five-day COMET-inspired practical project only as an alternative trade test assessment.

Qualitative data collection

Qualitative questionnaires were administered to both the assessors and the candidates in order to provide additional context for the quantitative data collected through the task-based assessments. Questions aimed at the assessors or examiners related to both their expert views on the content and method of the traditional trade test and their observations of the candidates' commitment to the COMET-inspired task, including factors that might have influenced the candidates' examination performance. Questionnaires aimed at the candidates related to their views on the task's level of difficulty, their interest in the task, the usefulness of the task, the effort applied to complete the task, and the usefulness of the task to their occupation.

Findings

For purposes of this article, only a few of the research findings are highlighted here to illustrate the potential of the COMET Model for going beyond traditional artisan candidate trade testing, and to demonstrate that occupational competence might be enhanced by applying a future-oriented COMET approach instead.

Quantitative findings

Finding 1: The COMET Model rendered fine-grained levels of competence
A substantial part of the study was based on quantitative data generated by applying the COMET diagnostic model in an alternative trade test in order to measure the occupational competence of a group of trade test candidates. In addition, the intention was to ascertain the

competence development of candidates who had experienced the COMET learning and assessment methodology in preparation for the trade test, compared with those candidates who took only the practical COMET-inspired trade test as the end-point assessment.

By applying the diagnostic analytics of the COMET Model to candidates who completed the validated test tasks, we found that only eight of the 17 candidates (fewer than 50%) reached a functional level of competence as described by the dimensions of the COMET Model. The 17 candidates represent the sample after test-task validation. The remaining nine candidates achieved a nominal level of competence, which, by international standards, is considered a risk level for occupational competence. These low levels of competence confirm the continuing challenge of dealing with learning and teaching deficits among apprentices in the electrical occupation, which had also been demonstrated in earlier studies (Hauschildt, 2016; Jacobs, 2015).

Further analysis of the results revealed the extent to which all the competence criteria were expressed in the COMET test solutions of candidates in Group A/Path A and those in Group B/Path B, as illustrated in the overall radar graphs below. As shown in Figure 1, the eight competence criteria used to evaluate the candidates' solutions were (Rauner, Heinemann, Hauschildt & Piening 2012:16–17):

K1 – Clarity: Candidates must document and present the results of professional tasks in such a way that both customers and work superiors are able to understand and review the proposed solutions. A core element of communication in the work context is the ability to express one's thoughts in a clear and organised way by giving clear accounts, drawings and sketches.

K2 – Functionality: This refers to the technical competence of instruments or context-independent, subject-specific knowledge and skills. Candidates are expected to provide evidence of the functionality of a solution; such evidence will determine all further requirements that enable work tasks to be solved.

K3 – Use or utility value: Professional activities, workflow, work processes and work assignments must ultimately be usable and oriented towards a customer, whose main concern is the utility of the result. The criterion of use therefore points to the usability of a solution in the entire context of work: a usable solution must be immediately applicable, less likely to fail, and take into account the need for easy maintenance and repair. It should preferably also be sustainable and capable of enhancement.

K4 – Cost-effectiveness: Candidates are expected to consider the context-specific economic viability of a solution, that is, how economically a specific task can be carried out. This should include considering diverse types of costs and influences, including long-term costs, with a view to performing a sound cost–benefit analysis.

K5 – Work process: This criterion refers to the way in which the test task will relate to the preceding and the following operations in the process chain. Candidates will be expected to

take into account the linkages with the preceding and following processes in the chain, not just their specific task.

K6 – Social responsibility: Candidates will be expected to include aspects of work safety and the prevention of accidents in addition to the potential impact of a specific solution on the social environment. It should take into account health protection and the often divergent interests of principals, customers and society.

K7 – Environmental responsibility: Here, the candidates should consider whether environmentally friendly materials are used and whether eco-friendly work organisation is employed in arriving at the solution of the work task. And have they considered energy-saving strategies and the possibility of recycling?

K8 – Creativity: The creativity of a solution is an important indicator of professional problem-solving, but a creative or unusual solution has to be interpreted and operationalised in an occupation-specific way, showing sensitivity to the problem(s) to be solved; and it can also be expected to make a meaningful contribution to the attainment of a goal.
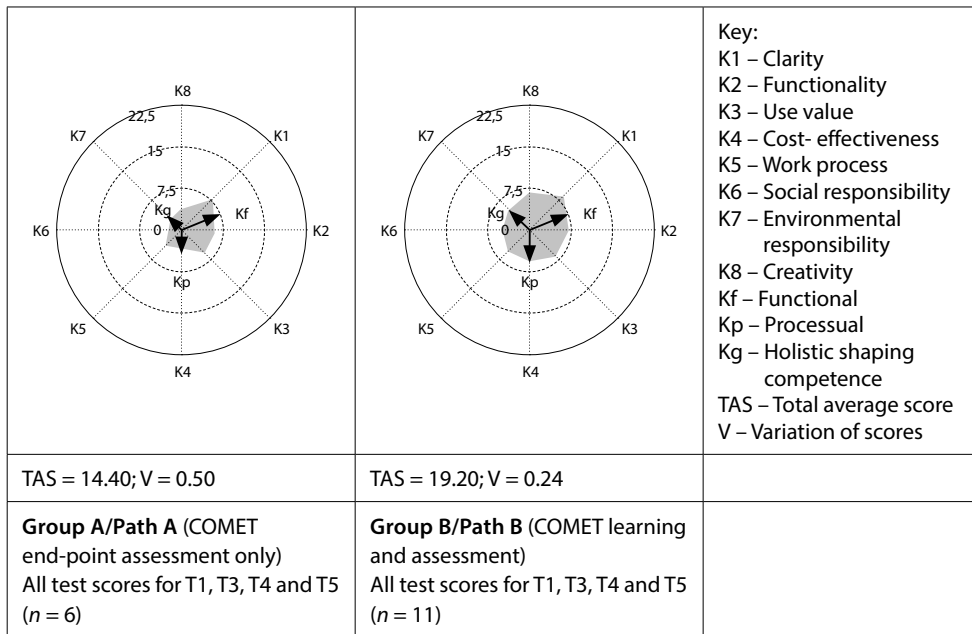

		Key: K1 – Clarity K2 – Functionality K3 – Use value K4 – Cost- effectiveness K5 – Work process K6 – Social responsibility K7 – Environmental responsibility K8 – Creativity Kf – Functional Kp – Processual Kg – Holistic shaping competence TAS – Total average score V – Variation of scores
TAS = 14.40; V = 0.50	TAS = 19.20; V = 0.24	
Group A/Path A (COMET end-point assessment only) All test scores for T1, T3, T4 and T5 (*n* = 6)	**Group B/Path B** (COMET learning and assessment) All test scores for T1, T3, T4 and T5 (*n* = 11)	

FIGURE 1: Comparative performance of Groups/Paths A and B, excluding non-valid test tasks 2, 6, 7 and 8[1]

---

1   Each test task should be able to demonstrate the application of all eight criteria. This is normally assessed through the 'V' (variation score). Through inspection of the radar graph, certain criteria could not be scored with evidence, which meant excluding these four test tasks.

All the candidates in Group A and Group B passed the traditional trade test between one and six weeks before the COMET-inspired trade test was taken. The comparisons in Figure 1 offer insights into the outcomes of the COMET-based test procedure: Group A achieved a lower total average score (TAS = 14.4), without all eight competence criteria being equally developed (V = 0.50), particularly the competence criteria representing environmental responsibility (K7) and creativity (K8). The higher variation of scores (V) around the mean (represented by 0) presented in Group A confirmed that the integration of knowledge – for instance, the practical solution considered with regard to the cost-effectiveness of the solution, the work process or the social and environmental responsibilities – inherent in the task had not been adequately demonstrated in the COMET-inspired practical task. The graphs confirmed a higher total average score (TAS = 19.2) for Group B, with a lower variation score (V = 0.24) across the criteria. The average performance of Group B (experimental group) was therefore able to earn more points for their solution across all eight COMET criteria.

Finding 2: Criterion-referenced evaluation through the COMET Model pinpointed learning needs

The usual practice in a trade test is to rate candidates against a defined rubric for a particular set of tasks and then declare them either 'competent' or 'not yet competent' in each task. In the traditional trade test, in the case of six practical tasks the requirement would be to be declared competent in each of the six tasks before being certified as a qualified artisan, that is, there is no grading of each individual task – the overall result is simply competent or not yet competent. In the traditional trade test, functionality is a primary consideration, which is exemplified in the following questions: Did the installation work? Was the earth leakage mechanically strong? Was the overload calculated? Was each eye separated by washers? Were there more than six 'non-critical' or small things identified by the assessor that would place the result of the concluded task in the 'not-yet-competent' category?

In the COMET rating process, on the other hand, such assessment questions would form only one-eighth of the evaluation carried out by the rater in the rating procedure. When rating a COMET solution to a specified task, the whole solution is rated against the eight criteria, that is, five questions are asked per criterion and they are scored using a four-point Likert-type scale from 0 to 3, defined as follows: 0 = criteria not met at all, 1 = criteria mostly not met, 2 = criteria mostly met, or 3 = criteria fully met.[2] This process enables a visual presentation of the distribution of total average scores (TAS) around the mean, as presented in Figure 1.

Even though the apprentices in the study were all at the end of four years of learning and were practising their trade, the diversity of results at the end of their apprenticeships was evident in the high variation of scores around the mean, depicted in the box plot in Figure 2. This presents a substantial challenge for any training provider because mentoring,

---

2   English translation of the original German rating terminology.

coaching and training have to be performed in a manner in which both stronger and weaker candidates benefit, to the extent that no one is in a 'risk category' of competence after passing the traditional trade test.
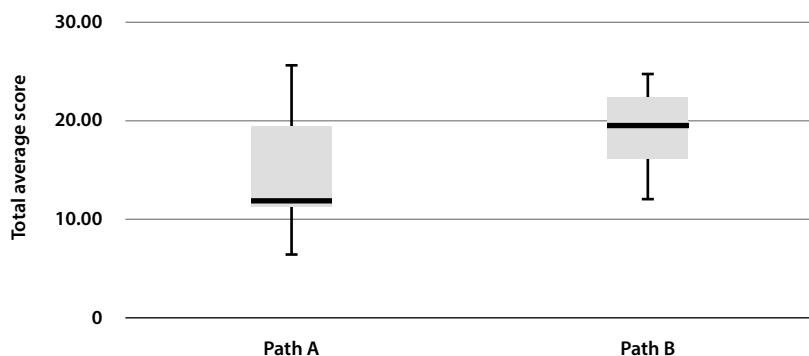

FIGURE 2: Box plot indicating the distribution of scores around the mean

The box plot illustrated in Figure 2 describes the distribution of scores around the mean, with no outliers indicated (Group A/Path A n = 13; Group B/Path B n = 13). The total scores for each group were normally distributed, as assessed by the Shapiro–Wilk's test (p >0.05). Group B/Path B (M [mean] = 18.92, SD [standard deviation] = 4.25) achieved a higher mean average than Group A/Path A (M = 14.38, SD = 6.17), with Group A/Path A scoring more widely than Group B/Path B, thereby confirming a statistically significant difference of p = 0.39.

When we consider the variation of candidate scores (or the spread of scores around the mean), it can be seen that the higher average score (TAS) leads to a lower spread of scores around the mean. Therefore, if the trade test is able to give adequate attention to all eight criteria of the COMET diagnostic model, the outcome could be higher average scores and less variation around the mean. Any variation of scores around the means of groups of candidates in a year, or of groups of candidates at a location, could also be used to monitor the quality of teaching and learning.

Finding 3: Candidates exposed to COMET methodology showed more holistic task solutions

The COMET-based test procedure confirmed a higher total average score for Group B/Path B candidates compared with Group A/Path A candidates, as illustrated in Table 1.

The results in Table 1 show that the competence outcomes of the traditional task-based trade test favoured task presentation and functionality over a more holistic task solution. There is a small 0.4 average point difference in the functional competence score (K1 and K2). The remaining criteria that represent processual competence and holistic shaping competence have much larger differences in the average score, ranging from 1.8 points to 2.7 points. This

indicates that the traditional trade test method is oriented towards, for example, the standardised functioning of an electrical installation rather than improving the performance of the installation that includes considerations of cost-effectiveness (K4), use value to others in a work process (K5), social and environmental responsibility (K6 and K7), and creativity (K8) with regard to technological advancements.

TABLE 1: Difference in individual COMET Model criteria scores between Group A/Path A candidates and Group B/Path B candidates

COMET MODEL CRITERIA	GROUP A/ PATH A AVERAGE SCORE	GROUP B/ PATH B AVERAGE SCORE	DIFFERENCE
K1: Clarity/presentation	8.1*	8.4	+0.3
K2: Functionality	6.6	7.1	+0.5
K3: Use value	5.8	6.8	+1.0
K4: Cost-effectiveness	3.5	5.3	+1.8
K5: Work process	4.1	5.5	+1.4
K6: Social responsibility	2.2	4.9	+2.7
K7: Environmental responsibility	2.5	5.0	+2.5
K8: Creativity	3.9	6.9	+3.0
Total points for functional competence level 1	7.4	7.8	+0.4
Total points for processual competence level 2	4.1	5.9	+1.8
Total points for holistic shaping competence level 3	2.9	5.6	+2.7

*   *Note that it is not normal COMET practice to say how many total points could have been scored or how many total points are allocated to each criterion (e.g. 10). The diagnostic capability relies on benchmarking scores, that is, scores of Path B become a new benchmark through which data in future studies could be evaluated.*

Finding 4: COMET methodology assessment feedback helped to identify particular disparities

The diversity of the competence outcomes can be evidenced in multiple ways through results measured between test sites and across gender and age groups. For example, the results across the four different test sites indicated diversity between total average scores and variation measures. Those test sites with somewhat lower average scores (shaded in Table 2) are an indicator of teaching and learning deficits that would benefit from improvement strategies.

TABLE 2: Differentiation of competence scores by test site

TEST SITE	NUMBER OF CANDIDATES	TOTAL AVERAGE SCORE (TAS)	VARIATION COEFFICIENT (V)	HIGHEST CANDIDATE (TAS)	LOWEST CANDIDATE (TAS)
Test site 1	3	20.1	0.17	24.8	16.0
Test site 2	12	18.4	0.37	25.0	11.4
Test site 3	6	16.6	0.37	21.5	9.0
Test site 4	5	9.9	0.74	13.0	6.5

Gender comparisons indicated that the total average score of females compared with their male counterparts was 3.2 points lower for Group B/Path B candidates. However, when the scores on individual competence criteria were examined, females excelled in the criteria of clarity of presentation (K1) and social responsibility (K6) to a greater extent than their male counterparts. Even though the sample was small, this information could be valuable to teachers and trainers in indicating specific areas in which students require additional assistance. Here the 'washback effect' of the COMET Model trade test construction and its feedback has the potential to reveal more particular disparities in the development of occupational competence.

Qualitative feedback

Finding 5: The traditional trade test falls short of developing professional competence
The responses from the examiners revealed that the current learning paradigms which influence content and assessment methods do not fully prepare trade test candidates for the dynamic world of work. Among the concerns mentioned were that the traditional trade test preparation allowed for shortcuts to be taken in the curriculum in that *'[t]he task-based content of the trade test does not test all the knowledge components of the curriculum'* (Examiner 3B) and that *'not enough embedded knowledge is covered in most of the tasks'* (Examiner 4B), this latter statement suggesting that the knowledge locked in processes, products, culture, routines, artefacts or structures (Gamble & Blackwell, 2001) is not fully exploited in the assessment.

Data from the examiners suggested that the traditional trade test content and method fall short of developing the professional competence of candidates in line with the expectations set by industry. In a traditional trade test, candidates pass the test based on them being found competent in a set number of closed tasks in an examination, but the tasks themselves are not adequate preparation for the dynamic world of work. In this context, candidates are usually not aware of any skill or knowledge deficit and are not encouraged in the preparation phases *before* the traditional trade test to increase their competence in related work processes such as writing a report, planning, finding innovative technical solutions, environmental and social responsiveness, or cost implications.

Finding 6: COMET Model assessments improve candidates' understanding of work processes
The majority of artisan candidates in the study reported that practical applications would enable them to add value to the work process, as the following extracts illustrate:

> *Because it would improve my skills and knowledge [regarding] how to limit unplanned downtime in the plant through planning and reporting everything* (Candidate 037);

and

> *It will [better] equip me … [for] my trade and make me a better electrician* (Candidate 062).

On the nature of the assessment using the COMET Model, the following comments represent a majority of similar responses:

> *Because it trains a learner to have an open mind … [and] to be able to think about future challenges instead of focusing on the now only* (Candidate 040); and

> *This project helps grow the mind; the way you think changes afterwards* (Candidate 048, Group B/Path B).

The data substantiated candidates' perceptions of the benefits of the COMET-based trade test and also their understanding of the importance of work processes and technology being embedded in occupational tasks.

Candidates commented on how the COMET-based approach to trade testing had influenced their learning strategies, a finding strengthened by the examiners reporting that all the candidates had demonstrated a high commitment to the COMET-inspired trade test and were focused on meeting the requirements that would demonstrate a working solution for the project specifications.

Discussion

The findings of this study showed that, in South Africa, the model of competence shaping the trade test is not expressed as a construct that is measurable in any of the guiding policy documents. It can be argued that in fact there is no implicit or explicit model for measuring competence that shapes or defines the trade test. The implications of this are that the reliability and validity of the traditional trade test instruments cannot be scientifically measured. For instance, the effectiveness of the traditional trade test instrument is usually relegated to the feedback of examiners or subject-matter experts about:

- whether tasks had clear instructions;
- whether mark sheets matched task-outcome requirements;

- whether the tasks are valid in the context of the occupation; and
- whether each task can be completed in the time available.

The critical opportunity to demonstrate how the assessment instrument responds to the objective of 'holistic shaping competence' is never dealt with or seized upon.

Despite South African education and policy intentions regarding integrated assessment, the results show that candidates who passed the task-based regulatory trade test in this research study were unable to integrate their knowledge and practical skills when presented with the dynamic COMET Model assessment applied to all 26 candidates in the study. No candidate achieved the processual or holistic shaping competence levels as described in the COMET Model, which illustrates that success in the traditional trade test is not a proxy for competence in the dynamic world of work. Each candidate would, given explicit instructions, probably be able to execute a defined task in the workplace, but the potential for acquiring a 'shaping competence' would not have been included in the learning pathway to the traditional trade test.

Work-oriented and integrated assessments should, according to the COMET Model, look past the action dimension of competence (activities such as analysing information, evaluating alternatives, planning, preparing, implementing and reporting) to include a level of expertise that demonstrates the ability of a candidate to 'think like an artisan', as is expressed in the eight criteria of the COMET Model. This study showed that a dynamic assessment approach which emulates the requirements of the occupation in the modern workplace through authentic work-related projects, is urgently required.

How can the current artisan trade testing system be improved?

Trade test examiners participated actively in this study by developing COMET-inspired test tasks and by rating the solutions delivered by the candidates. This involvement provided valuable feedback on how the current trade test system could be improved. An overall comment expressed was that the policy notion of 'applied competence' needed to be more comprehensively understood, because this was not being achieved through the current trade test. The examiners stated that they were not confident that the traditional trade test is fit for its purpose; nor did they believe that it contains cognitive challenges aligned with the dynamic world of work.

Furthermore, the low levels of holistic shaping competence recorded were indicative of the necessity to reform the current trade testing system. Although the trade test candidates in this study did not achieve processual or holistic shaping competence, they did display high levels of motivation after completing the COMET-inspired trade test and expressed this through positive comments on the value of incorporating real work processes, insights gained into the future of work and technology, and improved learning strategies into their training and testing.

This comparative study demonstrated the potential of COMET-based trade testing that is aligned to the demands of a dynamic world of work. Even though the sample size is too small to make a determination with a high level of confidence, it makes a case for expanding the study to include a much larger sample size.

It can be argued that the COMET Model offers a strategy with which to improve the trade test system in South Africa, as it incorporates a reflective assessment model for evaluating the competence outcomes of trade testing. Such a model for diagnostic analysis would encourage lecturers and trainers to adjust their content and methods of teaching to align them with more fine-grained measures of competence. Without a conceptual competence model, there are very limited points of reference to guide assessment that is suited to dynamic work processes.

Indications for future research

While the electrician occupation was selected for study because it was the most tested occupation in four national accredited assessment centres across three provinces, a broader sample of occupations could yield important comparative insights.

A methodological challenge for the future could be to extend the application of the COMET Model to more practitioners, since such expansion will require a higher level of skill in rigorous quantitative analysis. In addition, there would be the need to construct competence profiles, motivational factor analyses and specialist support for generating and analysing quantitative data. The COMET Model approach employs quantitative data, large sample sizes, statistical tests of significance and comparisons of variables related to competence criteria, motivational factors and so on – research activities that would require extensive capacity-building and technical support if such assessments are to be used. In the light of the potential benefits demonstrated by the approach to date, investment in such capacity-building may be well worth the effort.

Conclusions

The intention of the study reported on here was to explore the potential of the COMET Model methodology not only to diagnose and measure competence, but also to build and improve competence development through the assessment methodology offered by the model. The traditional artisan trade test used in South Africa provided a comparative assessment process, in that its overall 'competent/not-yet-competent' outcome presented a counterpoint to the fine-grained analysis of competence espoused by the COMET Model.

The findings of the study also revealed the relationship between competence development and summative assessment in the artisan trade test and the preparation towards taking it. While more detailed findings could not be elaborated on in this article of limited scope, we were able to illustrate, through selected evidence, the potential of a competence measurement model that is aligned to a transforming world of work vis-à-vis the deficits revealed in the

traditional trade test. The competence profiles of candidates in the study proved that the traditional trade test system does not adequately equip artisan candidates with the domain-specific occupational competencies needed. Fast-paced technological innovation requires a competence model that accommodates technological transformations in the workplace, which necessitates a responsive end-point assessment approach that is supported by scientific measurement and goes beyond the limited trade testing paradigms.

References

Achtenhagen, F & Winther, E. 2014. Workplace-based competence measurement: Developing innovative assessment systems for tomorrow's VET programmes. *Journal of Vocational Education & Training*, 66(3):281–295.

Argyris, C & Schön, DA. 1997. Organizational learning: A theory of action perspective. *Revista Espanola de Investigaciones [Spanish Research Journal] (REIS)*, (77):345–348.

Boreham, N. 2002. Work process knowledge in technological and organizational development. In N Boreham, R Samurçay & M Fischer (Eds). *Work process knowledge*. London: Routledge, 1–14.

Brown, H. 2015. Competence measurement in South Africa: Teachers' reactions to feedback on COMET results. In E Smith, P Gonon & A Foley (Eds). *Architectures for apprenticeship: Achieving economic and social goals*. Ballarat, VIC: Federation University (FedUni), 91–95.

Bybee, RW. 1997. *Achieving scientific literacy: From purposes to practices*. Westport, CT: Heinemann.

Bybee, R, McCrae, B & Laurie, R. 2009. PISA 2006: An assessment of scientific literacy. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 46(8):865–883.

Cameron, R. 2009. A sequential mixed model research design: Design, analytical and display issues. *International Journal of Multiple Research Approaches*, 3(2):140–152.

Cheng, L. 2005. *Changing language teaching through language testing: A washback study* (Vol 21). Cambridge: Cambridge University Press.

Curtis, DD. 2010. Teaching, learning and assessment in TVET: The case for an ecology of assessment. *SEAVERN Journals*, 2(1):1–24.

Deitmer, L, Hauschildt, U, Rauner, F & Zelloth, H (Eds). 2012. *The architecture of innovative apprenticeship* (Vol 18). Dordrecht: Springer Science & Business Media.

Department of Higher Education and Training (DHET). 2015. Trade test regulations. *Government Gazette* (No 38758 of May 2015). Pretoria: Government Printers.

Department of Higher Education and Training (DHET). 2019. National skills development plan 2030: An educated skilled and capable workforce for South Africa. *Government Gazette* (No 42290 of March 2019). Pretoria: Government Printers.

Fischer, M & Boreham, N. 2004. Work process knowledge: Origins of the concept and current developments. In M Fischer, N Boreham & B Niham (Eds). *European perspectives on learning at work: The acquisition of work process knowledge*. Johannesburg: Emerald Group Publishing, 12–53.

Gamble, J. 2021. The legacy imprint of apprenticeship trajectories under conditions of segregation and apartheid in South Africa. *Journal of Vocational Education & Training*, 73(2):258–277.

Gamble, PR & Blackwell, J. 2001. *Knowledge management: A state of the art guide.* London: Kogan Page .

Gardner, J. 2013. The public understanding of error in educational assessment. *Oxford Review of Education*, 39(1):72–92.

Grosse-Beck, R. 1998. Was hat Innovation im Prüfungswesen mit den Ewiggestrigen bei der PAL gemein [What does innovation in the examination system have in common with the old-timers at PAL television standard]. *Gewerkschaftliche Bildungspolitik* [*Union Education Policy*], 2:3–5.

Haraguchi, N, Cheng, CFC & Smeets, E. 2017. The importance of manufacturing in economic development: Has this changed? *World Development*, 93:293–315.

Hartig, J, Klieme, E & Leutner, D. 2008. *Assessment of competencies in educational contexts.* Gottingen: Hogrefe & Huber Publishers.

Hauschildt, U. 2016. *COMET South Africa: Final report and documentation of test results: Electricians, mechatronics, motor mechanics and welders.* Research Report commissioned by the Manufacturing, Engineering and Related Services Sector Education and Training Authority (merSETA) through the University of Bremen, Germany.

Heinemann, L, Maurer, A & Rauner, F. 2010. Ensuring inter-rater reliability in a large-scale competence measurement project in China. *International Network for Innovative Apprenticeships (INAP).* London: Transaction Publishers.

Jacobs, P. 2015. *The potential of the COMET competence diagnostic model for the assessment and development of occupational competence and commitment in technical vocational education and training.* PhD thesis. University of Bremen, Germany.

Jenewein, K. 2017. *Rezension von Klause Jenewein: Methodenhandbuch. Messen und Entwickeln beruflicher Kompetenzen (COMET)* [Review by Klause Jenewein: *Handbook of methods. Measuring and developing professional skills (COMET)*].

Klotz, VK, Winther, E & Festner, D. 2015. Modelling the development of vocational competence: A psychometric model for economic domains. *Vocations and Learning*, 8(3):247–268.

Kniivilä, M. 2007. Industrial development and economic growth: Implications for poverty reduction and income inequality. *Industrial Development for the 21st Century: Sustainable Development Perspectives*, 1(3):295–333.

Mulder, M. 2017. *Competence-based vocational education and professional education: Bridging the worlds of work and education.* Geneva: Springer.

Nickolaus, R, & Seeber, S. 2013. Berufliche Kompetenzen: Modellierungen und diagnostische Verfahren. *Handbuch Berufspädagogische Diagnostik*, 1:155–180.

Pan, YC. 2009. A review of washback and its pedagogical implications. *Nghiên cứu nước ngoài (VNU) Journal of Foreign Studies*, 25(4):257–263.

Pattalitan, JA. 2016. The implications of learning theories to assessment and instructional scaffolding techniques. *American Journal of Educational Research*, 4(9):695–700.

Peterman, F. 2018. Development and diagnostics of professional competences. (H Hauschildt, trans). *Journal of Educational Science*, 21(1):205–210.

Quality Council for Trades and Occupations (QCTO). 2016. *Assessment policy for qualifications and part qualifications on the occupational qualifications sub-framework (OQSF) – March 2016.* Hatfield: QCTO.

Rauner, F. 2017. *Methodenhandbuch: Messen und Entwickeln beruflicher Kompetenzen (COMET) [Methods manual: Measuring and developing professional skills (COMET)]*. Bielefeld, Germany: Bertelsmann.

Rauner, F, Hauschildt, U & Heinemann, L. 2013. Measuring occupational competences: Concept method and findings of the COMET project. In L Unwin, L Senker & E Fuller (Eds). *Architecture of Innovative Apprenticeship*. Dordrecht: Springer, 159–175.

Rauner, F, Heinemann, L, Hauschildt, U & Piening, D. 2012. Project report COMET-pilot test South Africa, including pilot test vocational identity/occupational commitment, first QRC results, April 2012. University of Bremen TVET Research Group. <https://www.merseta.org.za/wp-content/uploads/2021/04/COMET-RSA-Study-Report-April-2012.pdf>.

Wedekind, V. 2013. Rearranging the furniture? Shifting discourses on skills development and apprenticeship in South Africa. In S Akoojee, P Gonon, U Hauschildt & C Hofmann (Eds). *Apprenticeship in a globalised world: Premises, promises and pitfalls* (Vol 27). Münster: LIT Verlag, 37–46.

Weinert, FE. 2001. Concept of competence: A conceptual clarification. In DS Rychen & LH Salganik (Eds). *Defining and selecting key competencies*. Seattle: Hogrefe & Huber, 45–65.

Winther, E & Achtenhagen, F. 2009. Measurement of vocational competencies – a contribution to an international large-scale-assessment on vocational education and training. *Empirical Research in Vocational Education and Training*, 1(1):85–102.

BIOGRAPHIES

Dr Helen Brown

Helen Brown holds a PhD in education that focused on artisan trade testing in the Engineering TVET space. Her work in quality development of VET systems has been in the areas of large-scale competence diagnostics, professional development of TVET college lecturers, ICT teaching and learning platforms, and skills for the industrialisation of new product innovations.

Prof. Joy Papier

Joy Papier is the South African Research Chair, Post-School Studies: TVET, in the Institute for Post-School Studies (IPSS) at the University of the Western Cape (UWC), Cape Town. She has been actively involved in TVET research, policy development, and capacity building for over 25 years, and has published widely in the field.