International Journal for Research in Vocational Education and Training (IJRVET)
Vol.6, Issue 2, August 2019, 132-152
DOI: 10.13152/IJRVET.6.2.2

Large-scale studies of holistic professional competence in vocational education and training (VET). The case of Norway

Leif Chr. Lahn*, Hæge Nore

OsloMet–Oslo Metropolitan University, Department of Vocational Teacher Education, Kunnskapsveien 55, 2007 Kjeller.

Received: 28.04.2018; Accepted: 08.02.2019; Published: 29.08.2019

Abstract

Context: In this paper, we review and discuss the piloting in Norway of a German methodology for competence diagnostics in vocational education and training: the Competence Development and Assessment in TVET (COMET) project. Our overarching theme is determining to what extent such large-scale assessment systems are valid for international comparisons in this sector.

Method: We present the theoretical underpinnings of the COMET model and position our discussion within the broader context of the concept of "professional competence" (berufliche Kompetenz) and methodological guidelines for its measurement. Terminology from psychometrics on "measurement equivalence" is described and serves as a template for identifying challenges in using the Norwegian data for comparative purposes. Our pilot included students and apprentices in health care, industrial mechanics and electricians and was designed as a three-year follow-up study from the second year of upper secondary school through two years of apprenticeship. Each year, a test on professional competence and a context survey were administered. Similar studies have been conducted in Germany, China and South Africa.

Results: In line with the results from these countries, the Norwegian participants had low scores, particularly the electricians. However, the diagnostic instrument was sensitive to the development of professional competence, and progress on the assessment was influenced

*Corresponding author: l.c.lahn@iped.uio.no

ISSN: 2197-8646
http://www.ijrvet.net

by the quality of the learning support in the companies, as reported by apprentices in the context survey.

Conclusions: The COMET platform may be a viable prototype for the development of diagnostic tools, which may support the monitoring of quality factors at different levels and inspire local improvement projects in schools, companies and training offices. Such an objective would be in line with the latest summaries of the COMET project, in which its contribution to a model for international large-scale assessment is toned down and replaced by a stronger emphasis on its potential for measuring competence development, evaluating contextual factors and generating data for didactic innovations.

Keywords: Vocational education and training, large-scale assessment, professional competence, international measurement comparability, VET

1 Introduction

The pioneering work on international large-scale studies in education was initiated 60 years ago in both a European and US context. Since then, these assessments have expanded in scope, focusing on a broader range of populations and skill domains (Kirsch, Lennon, van Davier, Gonzalez, & Yamamoto, 2013). New countries have been recruited, representing a large mix of economies, cultures and geography (Rutkowski & Rutkowski, 2018). Three overall motives for such efforts could be distinguished (Smith & Comyn, 2003; Kuger & Klieme 2016): (1) a concern for national competitiveness in a global knowledge-based economy, (2) actions to innovate the educational sector, and (3) curriculum development and teaching strategies to improve schools. Baumert (2009) adds that the universalization and standardization of schooling are global trends which stir an interest in international comparisons of educational systems. Thus, it is not surprising that initiatives were taken a decade ago to develop large-scale assessment instruments that should enable across-country comparisons of learning outcomes of different systems for vocational education and training (VET).

There are several reasons why the measurement of competences in the field of VET is more challenging than in compulsory education (Baethge & Arends, 2009). Whereas international large-scale assessments have a limited focus on pupils' performance in mathematics and science (PISA, TIMMS) or generic competencies (IALS, PIAAC) , similar efforts for vocational skills need to take into account a large variety of vocational domains and important differences in the educational configurations between countries. In addition, the development of measurement tools for VET competences cannot draw on long research traditions and accumulated knowledge about vocational expertise. Hence, the design of such projects should be subjected to feasibility testing and explorative inquiries. The elusive character of concepts such as "vocational competence," "generic competence" and "literacy" does intro-

duce added challenges of measurement and formal assessment (Mulder, 2017). In this article a major theme is the international and transcultural comparability of such instruments when they are designed to measure holistic competences in VET.

We will present and discuss the study, "Measuring competence development in vocational education and training" (MECVET), which was a pilot in Norway of a German effort to develop and test a methodology for professional competence diagnostics ("Kompetenzdiagnostik") in VET, COMET ("Competence Development and Assessment in TVET", Rauner, Heinemann, Maurer, Haasler, Erdwien, & Martens, 2013b). Although the primary objective of the COMET model was to establish a theoretical and methodological platform for an assessment system that should support the pedagogical work of VET teachers (Rauner et al., 2013b), a secondary intention aimed at constructing a tool for international comparisons that would provide data for policymaking in the VET sector. This article leans on the latter by revolving around the comparability of test scores and the adaptability of assessment systems to different institutional and cultural contexts. How did the model and its instrumentation meet such requirements, and to what extent were they implemented in practice? We first position our text in the literature on a competence-based assessment, in which models for a holistic diagnostics of professional competence stands out as an interesting subarea, though problematic if the intention is to generate comparable data across countries. Next, we review the basic principles of the COMET instruments and their implementation in the MECVET study, bringing to the fore issues of measurement equivalence. Our methodological reflections run as a thread through the presentation of test scores and other data on professional competence from our panels of students and apprentices in the vocations of industrial mechanics, electricians and health-care workers – and through our concluding discussion. Norway should be an interesting case for various reasons. It is a small country with a VET system that encompasses approximately 180 vocations and a sequential dual model with two years in school followed by two years in apprenticeships, which differs from the German parallel school/work track. The number of apprentices in each trade is relatively small compared with other European countries. Formal assessment is toned down in Norwegian VET, with the exception of the final trade examination (after apprenticeship, Lahn & Nore, 2018). These factors generate methodological, logistic and cultural challenges when piloting large-scale assessment studies in a Norwegian context.

Holistic approaches to the assessment and comparability of professional competence Competence-based education has become a favorite model for innovations in VET, provoking intense debates about reductionist versus expansive interpretations of such reforms (Mulder, 2007). Surprisingly, the debate about the assessment of professional competence tends to repeat the critical comments made by Hager, Gonczi and Athanasou (1994) more than 20 years ago. They highlighted the weaknesses of two dominant approaches – the behaviorist and the cognitivist, where the main characteristic of the former was a priority attri-

buted to task analysis (functions) and observation of performance, which could easily turn out to be a fragmented and inflexible system of assessment lacking dimensions that take into account the transferability of proficiencies. At the other end, cognitive psychology has a long tradition of advocating the importance of general mental abilities in professional expertise, such as universal mechanisms of problem-solving (Ericsson, 2009) or intuitive thinking (Dreyfus, 2004). A recent version of this perspective that intends to meet the challenges of a turbulent working life is the focus on individual's self-organizing potential or generalized meta-competences (Haasler & Erpenbeck, 2008), adaptive expertise (Carbonell, Könings, Siegers, & van Merriënboer, 2016), 21st century skills (Pellegrino, Hilton, & National Research Council, 2012) and core work skills (Brewer, 2013).

Hager et al. (1994) introduced an integrative or holistic model that seeks to avoid the challenges of the proliferation of tasks (and abilities) by centering the assessment on key activities or elements of the profession – a position shared with others flagging concepts such as "key competences" (Rychen & Salganik, 2003). Like other versions of holistic approaches (Le Deist & Winterton, 2005; Mulder, 2007; Beckett, 2008), they extend the classification of competences to include action competences, social competences – and contingencies connected with varying contexts of vocational performance and individual strategies. This literature seems to converge on the following features when addressing the assessment of professional competences:

- The holistic logic is often contrasted with a reductive that focuses on one dimension or key task at a time and not a totality of key tasks (Lum, 2013). Hager (2017) argues that the capacity for "putting it all together" is crucial.

- Professional competence is inferred from work performance, and dependent on what Beckett (2008) refers to as a "judgment-in-context." Thus, properly qualified assessors should be experts in the relevant professional domain (Tigelaar & van der Vleuten, 2014).

- Vocational proficiency is a complex activity interacting with contextual affordances and restraints, asking for a multi-method and processual approach to assessment (van der Vleuten, Sluijsmans, & Joosten-ten Brikke, 2017).

- Generative aspects of professional competence are often highlighted in the holistic approaches, thereby reflecting the urgency of including capabilities that align with transformations in working life (Deitmer, Hauschildt, Rauner, & Zelloth, 2013). In the German literature, these characteristics are central to the definition of "profes-

sional action competence" ("berufliche Handlungskompetenz", Seeber & Nickolaus, 2010).

These scholars have been concerned with the use of holistic assessment, primarily for formative purposes and local curriculum development, which tend to be reserved toward the appropriateness of such a framework for large-scale international comparisons (Lum, 2013). One notable exception is the VET-LSA feasibility study initiated by the University of Göttingen (Baethge & Arends, 2009) – also referred to as "the PISA-VET" (Baethge, Achtenhagen, Arends, Babic, Baethge-Kinsky, & Weber, 2006). It is presented as based on a holistic understanding of competence, although such claims have been challenged by Dietzen (2017). The latter recognizes the multidimensional logic of the VET-LSA, but criticizes its underlying model of professional action competence (Weinert, 2001) for separately examining cognitive, motivational and social components, and not in an integrative manner. The VET-LSA was piloted in four vocational areas (car mechatronics, electricians in craft/industry, business/administration and social/health care) in eight European countries including Norway, where one of the authors of this article participated. In line with the principles of holistic assessment described above, the researchers constructed near-authentic task descriptions (occupational profiles) and qualification requirements based on the US standard classification of professions O*Net (https://www.onetonline.org/). These were validated by a team of experts from the four vocations. A summary of the Norwegian findings indicates that the participants, particularly representatives of the electricians, did not relate to the proposed profiles (Olsen, 2009). For various reasons, this international comparative exploration was discontinued in 2009, but in Germany ideas from VET-LSA were implemented in the ASCOT initiative (www.ascot-vet.net). The latter narrowed the scope by focusing on specific vocations such as business/administration and health care, and was reported to generate promising results in terms of psychometric models and computer-based assessment solutions (Klotz & Winther, 2017).

Partly as a result of the mixed experiences with VET-LSA, the researchers behind the MECVET study were attracted by another model for the diagnostics of professional development underlying the COMET project, an initiative coordinated by the Institute für Berufsbildung at the University of Bremen (Rauner et al., 2013b). In the international literature, this approach worked out by Professor Felix Rauner is described as strongly influenced by the German concept of "Beruf" when underlining the importance of holistic shaping competences ("ganzheitliche Gestaltungskompetenz") as qualities of personal development and professional identities (Cairns & Malloch, 2017; Winterton, 2012). The COMET model, to which we return in the next section, is firmly based on these principles.

As noted above, the instrumentation of COMET and its implementation in the MECVET study will be discussed using the following categories from a common typology of measu-

rement equivalence or comparability (van de Vijver & He, 2015). Theoretical comparability is threatened when the understanding of key constructs such as "competence" varies in a systematically biased way between participating countries. Sampling comparability is closely linked to the previous category since the conceptual framework has a decisive impact on the selection of content (tasks and items), subjects, etc. Instrument comparability should be attended to since teachers and students may differ in, for example, their familiarity with specific test formats. Linguistic comparability includes an adequate translation of items, scoring rubrics, etc. Administrative comparability is pursued by standardizing the test situation, etc. Rater comparability is not often addressed in the psychometric literature, mainly because the rating is automatic or semi-automatic, but must be seriously considered when scoring procedures involve the use of professional judgment in assessing student achievements on realistic complex tasks (Baartman, Bastiens, Kirchner, & van der Vleuten, 2007). Outcome comparability is included in order to address differences in consequential effects for the testees, such as a low or high stakes assessment. This typology will serve as a template for mapping the instances (in Table 1) when our case of large-scale assessment of vocational competences "breaks down" (Winograd & Flores, 1986), that is when the data do not fit the comparability standards and become the object of inquiry.

2 The COMET conceptual framework

The COMET project is exemplarily well documented in reports, articles, monographs and anthologies. Although the large majority of texts are in German (Rauner, Heinemann, Maurer, & Zhao, 2014; Fischer, Rauner, & Zhao, 2015), there are several reviews and discussions in English (Rauner, Heinemann, & Hauschildt, 2013a; Rauner et al., 2013b). For readers wanting to learn more about the details of the model and the instruments, we refer to this literature. The COMET project was not presented by its designers as belonging to the family of international large-scale assessment studies. Still, since the first pretests in 2006 these instruments have been piloted in a number of countries (Germany, China, South Africa, Norway, Switzerland, Vietnam, Spain and Poland), and broadened in scope by adding vocations such as industrial mechanics, car mechatronics and elderly care nursing to the original focus on craft and industrial electronic technicians. Thus, the model is claimed to have the potential for "international comparative competence diagnostics" (Fischer et al., 2015; Zhao, Zhang, & Rauner, 2016). Given such an ambition, we will review how the COMET instruments meet standard requirements for the comparability of results and indicate what the methodological challenges were when piloting them in a Norwegian context. However, in the following, we first go more thoroughly into the key elements of this "theory-based competence model" (Rauner et al., 2013a, p. 8).

Rauner et al. (2013b) distinguish between performance assessment and professional competence diagnostics. In the first case, VET teachers and trainers would be able to evaluate the professional action competences of apprentices on specific tasks. The examinations or tests should have curricular validity without any pretense of being generalizable to other contexts. In contrast, competence diagnostics generates comparative scores in order to provide solid input to reviews of learning opportunities across vocations and countries. The measurement focus is then displaced from action competence to cognitive (domain-specific) performance dispositions (ibid., pp. 5). In line with other versions of holistic assessment, Rauner (ibid.) maintains that a model of competence is needed in order to make inferences from observed performance. The COMET framework is founded on three dimensions.

The content dimension draws on expertise research, notably the concept "developmental tasks" (Havinghurst, 1974) and the stage model of Dreyfus (2004) and Benner (1982). Thus, this understanding of the content dimension enables an assessment of specific competence levels and developmental progression in specific vocations to be compared with a trans-occupational test concept (Rauner et al., 2013b, pp. 46). In face of the criticism that has been aimed against the Dreyfus model for advocating a restricted linear learning trajectory, an underestimation of training impacts and the diversity of knowledge domains (Lahn, 2010), it may not be as universal as claimed by Rauner et al. (2013b). At the same time, these theoretical ideas may provide the heuristics for comparative inquiries that will enable an empirical validation of the Dreyfus developmental model.

As one of its main building blocks, the action dimension has the concept of "complete professional action" ("vollständingen beruflichen Handlung"), with its roots in German work psychology (Hacker, 1973; Oesterreich & Volpert, 1986; Monnier, 2015) and cultural historical activity theory (Leont`ev, 1981). It frames human activity as a generative process when work performance includes goal anticipation and planning processes on the one hand, and the evaluation of action and its outcomes on the other. The idea is that effective learning will take place when workers know the professional task as a whole and can take action to reshape work if necessary – the essence of such concepts as "action competence" ("Handlungskompetenz") and "holistic shaping competence" ("ganzheitliche Gestaltungkompetenz"). In the vocabulary of the COMET researchers, these concepts are linked to "work process knowledge" ("Arbeitprozesswissen"), which emphasizes the way that a practical understanding or judgment of work processes at different levels and time scales strengthen the generative capabilities of skilled workers. Again, one may question the generality of these theoretical assumptions, and critics contend that they fit with industrial- and a limited number of service jobs, but do not catch the dependencies in advanced professional work (Mieg, 2001).

The third dimension, the requirement dimension, specifies the levels of "professional problem-solving" ("Berufliche Problemlösung", Rauner et al., 2013b) and is based on the four-stage literacy concept formulated by Bybee (1997) and incorporated in different versi-

ons of PISA. (1) The level of "nominal competence" ("Nominelle Kompetenz"), indicating that the apprentices have superficial conceptual knowledge and are not able to solve the task adequately in practice. (2) At the "functional competence" (Funktionale Kompetenz") level, the testees will have instrumental abilities in a domain, but be weak at connecting different elements in practice. (3) The level of "processual competence" (Prozessuale Kompetenz") is the one in which the solutions to the task include an appreciation of work processes and situational constraints, and (4) "holistic shaping competence" ("Ganzheitliche Gestaltungkompetenz"), which entails complex vocational problem-solving that encompasses a generative aspect. The COMET model does not include the nominal level (1) since it does not represent vocational competence. The last three levels are subdivided into eight content criteria (functionality, presentation, efficiency, sustainability, work-process orientation, environmental responsibility, social responsibility, creativity); the model therefore provides the basis for differentiated competence profiles.

3 COMET instrumentation and its implementation in the MECVET design

If we go from the basic theoretical concepts to the COMET test development process, the first stage is the construction of test tasks by experts in the vocation taking the model above as conceptual platform. An instrument in this case is the expert workers' workshops, which integrate both research-based and experiential input into a shared understanding of core tasks and learning trajectories in the vocation (for a full description of these workshops, see Rauner et al., 2013b; Bremer, Rauner, & Röben, 2001). The COMET design will prescribe the development of open test and learning tasks that enable all the eight content criteria to come into play and prompt a variety of solutions. The learners are sensitized to the open test format by participating in the solution of open learning tasks several months before testing. The testees work on two tasks for 120 minutes each. An underlying assumption is that these tasks represent the core vocational competences of a skilled worker (after trade examination).

The rating of test solutions was subject to stringent procedures. (1) The eight criteria were subdivided into 40 items. (2) For each task, a "solution space" was constructed defining the range of possible solutions. (3) Extended rater training in seminars intended to generate a consensual standard for scoring, strengthen inter-rater reliability and construct validity, and (4) two raters are independently rating each test participant's solution, hence making it possible to calculate the degree of correspondence using the Finn-correlation (Rauner et al., 2013b). In line with their understanding of vocational learning as an appropriation of a specific identity, the COMET researchers have devised a questionnaire on students'/apprentices' test motivation, vocational identity, occupational and organizational commitment and learning environments in school and companies.

In short, the piloting of the COMET model outside Germany (China, South Africa and Norway) was mainly done by translating test tasks, criteria, solutions spaces and the procedures for test development into the language of the respective countries. Moreover, the selection of vocations and the temporal patterning of measurements (cross-sectional/longitudinal/combined) were partly replicated. In general, these facets were adapted and validated by specific methods and procedures – especially the use of workshops with experts from the different vocations. The PROJECT study integrated the following elements:

Test development procedure - The different steps progressed in a manner common in test development projects: Review/translation of COMET test tasks, etc., validation of test tasks in workshops, the pretesting of test tasks and revisions, the test administration of test tasks and surveys, rater training and rating, the analysis of results and test context.

Sampling and temporal organization of the study (see Fig. 1) - Three vocations, electricians, industrial mechanics and health-care workers were chosen to approximately match those of the COMET project in Germany, and to fulfill our ambition of having a contrasting case in another sector

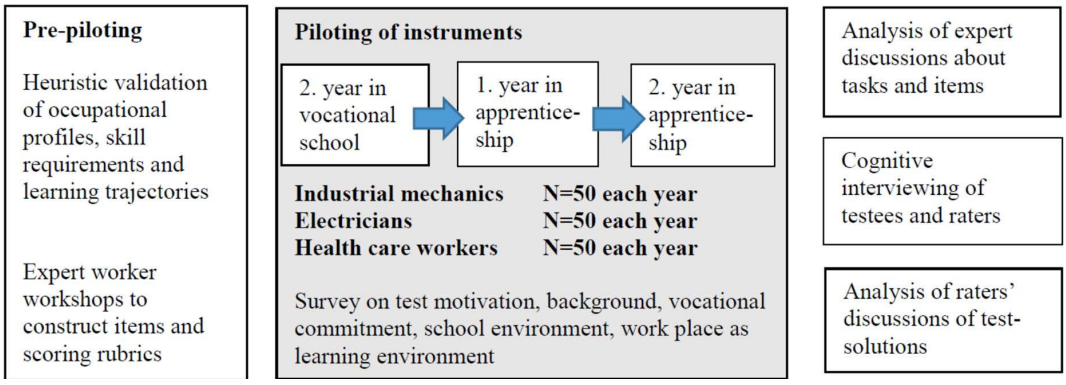

Fig. 1: The overall research design of the MECVET study

Our longitudinal design was intended as a follow-up of a cohort of students into their two years of apprenticeship. We realized early in the process that it would be impossible to get in touch with the same individuals across the three years, since the Norwegian dual sequential VET system involves a break between two years in school and then two years in companies. It also demarcates a point with students have to choose a specific vocation, and also a stage where a substantial number of dropouts of the VET track. For each year, we recruited new testees, although a subsample was measured in all three years. Due to these difficulties, our planned sample of 100 participants in each vocation for each year was reduced to approximately half, as indicated in the figure. We also attempted to obtain a regional balance, but

had to accept an overrepresentation of testees from schools and companies on the west coast of Norway, where there is a more favorable climate for VET than in the eastern part of the country.

Test tasks and web-based survey - For both the vocations of industrial mechanics and health-care workers, three out of four test tasks were translated and one considerably revised to meet typical work tasks and requirements in a Norwegian context. In the case of the electricians, three test tasks were made anew, while one was partly modified due to the professional profiles and (restricted) responsibilities for electricians in Norway. Another divergence from the COMET set-up with testees working on two tasks for four hours was our decision to contend with one task within a two hour limit. Schools and companies were not willing to allocate "nearly a day" for the MECVET project. We administered a web-based survey on the testees' engagement in the assessment, their work conditions and learning environments, as well as their vocational identities based on a German version. Some items were not relevant because of the different VET systems in Germany and Norway, which also required two surveys – one for the year in school and one for the two years of apprenticeship.

Field studies of workshops, test administration and rater training. Both workshops and rater training sessions were videotaped, selectively transcribed and analyzed using software for qualitative analysis (AtlasTi). Whereas teachers organized the testing in Germany, this was done by the MECVET researchers in Norway, which provided us with field data recorded in memos.

Mixed method data analysis - Since the process of constructing the assessment system was compromised along several dimensions, we had to rely less on harmonized templates and procedures, and instead rely on methods that enabled the analysis of the contextual contingencies of the testing (variation logic for inferential work (Maxwell & Chmiel, 2014). This would be an example of an explanatory sequential design (Creswell, Plano Clark, Gutmann, & Hanson, 2003), in which qualitative data should deepen the interpretations of the numerical material processed for descriptive and inferential statistics using IBM SPSS+. The international comparisons were done a posteriori and indirectly based on documentary (reported) data, since our access to data files from the other countries was very limited. For the Norwegian material, log-linear statistics were combined with principles of configurational comparative methods (Ragin, 1987) in order to explore structural and local causal relationships in these data (statistical analyses in the MECVET final report, Lahn, Nore, Johannessen, & Bakken, in print).

4 Measurement equivalence in principle and practice

This article is basically a methodological reflection on the contingencies and constraints imposed on the transfer of the COMET project from German to Norwegian VET. As noted above, international large-scale assessments rely on standards of measurement equivalence, which may be a challenge in VET given the diversity of vocations and the heterogeneity of educational systems, both within and across countries. Table 1 refers to the typology introduced earlier, and links it to stages in test development:

Table 1: Comparability typology related to COMET design and practices in the MECVET study

Comparability types	Stages in test development	Guidelines for COMET projects	PROJECT implementation (ethnostatistical account)
Theoretical comparability	Research framework	Holistic shaping competence	Norwegian competence-based VET
Sampling comparability	Sampling Test construction	Expert groups Same test tasks Educational level Level of aptitude	Experts groups from industry Many test tasks were reconstructed Educational level differences for specific vocations Variation in competence levels
Instrument comparability		Paper-pencil test without use of external resources	Trade examinations are performance based Assessment with the use of resources is common
Linguistic comparability		Translation and back-translation	Rating criteria were difficult to translate in a meaningful way
Administrative comparability	Pilot implementation	Standardization of instruction, procedures…	Voluntary participation of schools, companies, training offices, apprentices-generated variation
Rater comparability	Rating of scores	Intensive rater training Solution spaces/rubrics	Rater training was resource-demanding Few raters used solution spaces
Outcome comparability		Feedback to schools, teachers and testees	Assessment scores without consequences for schools, teachers and testees

The third column indicates how the COMET design and its implementation in the MECVET study were tuned to such principles. The empirical evidence for the ethnostatistical account (Maddox, Zumbo, Tay-Lim, & Demin, 2015) in the fourth column is based on documentary data and field observations.

Theoretical comparability - The terminology of "work process knowledge" and "holistic shaping competence" is part of the German discourse on "Berufspädagogik," and provides an ideological basis and a recognition that is not present among Norwegian actors in the VET sector. More importantly, the construction of the COMET model and measurement instruments was partly motivated by the German German VET reforms, which facilitated the acceptance of the project in German VET. These interests were not present in the Norwegian case. Here, we would add that in the psychometric discussion of international large-scale assessments, the country-specific rationales for participating in those programs have not been

a highlighted topic, although it may have a strong impact, not only on the use of test results, but also on how the assessment is implemented (Baethge & Arends, 2009). Both the Chinese and South African involvement in COMET can be traced back to their ambition of improving the connection between schools and companies in VET (Rauner, Heinemann, & Maurer, 2011; Rauner, Heinemann, Hauschildt, & Piening, 2012), whereas MECVET was framed as a research project that had an interest in developing new tools for formative assessment. Thus, a comparability in purpose could have been added to the typology of measurement comparability.

Sampling comparability includes decisions about recruitment to expert groups (for test task development), and about test tasks and the characteristics of the testees. Compared with the German composition of expert groups (and the rater groups), the Norwegian had a stronger representation of participants from industry. Additionally, the test tasks in the MECVET project were more totally reconstructed compared with the other countries – notably for the electricians.

A major issue in both the COMET project and the VET-LSA project has been the equivalence of the educational levels of testees across the participating countries. The different national COMET datasets include measurements of similar vocational groups training at various levels (upper secondary/tertiary in the technical sector), while the VET-LSA managed to establish an adequate comparability between the occupational profiles at the end of VET (Baethge & Arends, 2009). Differences in scholastic aptitude between vocations in the participating countries may account for the observed variation. The measurement of this variable is not easily equated and seems to influence the test results, but the COMET studies suggest that this interaction was lower than expected (Rauner et al., 2014). In the Norwegian material, however, the testees' grades at the second year of upper secondary vocational schools correlated significantly with their professional competence on the COMET test.

Instrument comparability - Biases in measurement may be introduced when the test format favors or penalizes the groups to be compared. The COMET test tasks were essay-based paper and pencil tasks to be worked on without any support materials, which go against the norm of assessment in Norwegian VET schools. The MECVET researchers had to accept that the testees could have access to, e.g., resources on the web, which of course made the comparison with the other countries less reliable.

Linguistic comparability may be an issue in the COMET studies, given that the test tasks are case-based, with a fairly elaborate presentation of a complex problem or order from clients. However, purely linguistic translation of these texts was less of a challenge than the exact wording of the criteria and items on the rating schemes, which had a strong influence on how raters weighted different aspects of the testees' solutions. For instance, we used the following translation from German:

Utility	Use value/sustainability
Economy	Effectiveness/profitability
Business and work process orientation	The organization of work
Social compatibility	Work environment/HSE

Administrative comparability - The Norwegian sequential VET system generated logistic challenges, since first and second year apprentices are spread around in a number of workplaces. In collaboration with schools and training offices, we had to gather groups ranging from two to 20 apprentices, and negotiate with the companies about how much production loss could be traded for the general benefits of the MECVET project. We had to abandon the COMET guidelines of letting the testees work on two test tasks for four hours. As a rule, we settled for one task and a two-hour time frame, but if the apprentices had to take the last ferry to get back to their workplace, a shipyard on an island, we had to accept a shorter time frame. Another, but somewhat surprising experience, was that teachers and instructors were used to having exams with "comforting rounds" (trøsterunde), in which they clarified questions, suggested ways of framing the problem and so on, and they felt uncomfortable when asked to take a pure surveying role. Consequently, nearly all of the test administration had to be done by the research team – which was a costly affair, but gave us the opportunity to do field notes.

Rater comparability is hard to achieve on international large-scale assessments in the VET sector, where there is no "lingua franca" about what constitutes vocational expertise. As pointed out above, the Norwegian raters had a stronger basis in working life than the COMET raters in the other participating countries, which may be associated with less endurance when it comes to rater training and fine-grained assessments using rubrics and solutions spaces. In order to observe these effects, we triangulated the inter-rater reliability measures with video recordings of rater training sessions and collaboration between raters aiming to agree upon scores. We performed a test-retest rating to uncover systematic biases, e.g., in terms of group norms.

Outcome comparability - Compared with the other countries where the COMET instrument was piloted, the institutional involvement (from schools, policymakers, etc.) was low in Norway, and for the apprentices/students their achievements on the test had no impact on their scholastic standing or workplace ratings. Due to the forthcoming trade examination at the end of the second year of apprenticeship, the testees were more committed to the test, which in addition to work performance, includes planning, argumentation and documentation. Given the methodological angle of this article, the review above about the "intangibles of measurement" (Viswanathan, 2005) or "break-downs" of measurement standards represenkey findings from the MECVET study, while at the same time framing the interpretation of test-score distributions and the survey data to be presented in the next section.

5 Presentation and discussion of results on professional competence across countries, vocations and educational levels

Our ambition to do an international comparative study was toned down, not only due to the challenges of piloting the COMETmodel, but also because we did not obtainget access to raw data from Germany, South Africa and China. Thus, the tabulations of variation in competence levels between these countries in Table 2 below are primarily based on publicized data, and of course, this rescue strategy has a number of shortcomings. The most obvious is the lack of data on critical dimensions such as the variations in scores, the indexing of items, etc. Relevant meta-data information about sampling, the indexing of items, etc. is often not provided, with the implication that the transnational analyses are only suggestive. Our empirical foundation for making comparisons across the three vocations and educational levels in Norwegian VET is much stronger.

General vocational competence - Table 2 shows the differences in competence levels for

Table 2: Differences in competence levels among the four countries, based on Figure 16, page 46 in Rauner et al. (2013b). The Chinese figures are from VET schools.

Countries	Nominal competence	Level 1 Functional competence	Level 2 Processual competence	Level 3 Holistic competence
Germany (Bremen)	50%	42%	7%	2%
China	59%	38%	1%	2%
South Africa	70%	27%	3%	0%
Norway	70%	26%	3%	1%

electricians or electro-technicians (industrial and craft) among the four countries (Rauner et al., 2013b). The patterns are similar but elevated for the German and Chinese results on Functional competence, while the scores on Processual and Holistic competence are generally very low. The Norwegian results are almost identical with those of South African testees, but when compared with a subsample of electricians in Germany (Hessen) with no experience on the specific type of open test tasks typical of COMET, these were slightly lower than the South African and Norwegian results. Measures of variance indicate large differences between schools and training providers with the participating countries, but no transnational comparison on this dimension is provided in reports, etc.

Differential vocational competence - Although the three dimensions of vocational competence represent requirement levels, the concept of "professional problem-solving" is associated with the idea that flexible performance depends on high ratings on all of the eight criteria mentioned earlier. One main purpose with the competence profile is to provide diagnostics of strengths and weaknesses for differential didactic measures in schools and companies. The variation coefficients are higher for the Norwegian testees, compared with the re-

sults reported from the other countries for the industrial vocations. This could reflect a more heterogeneous sample of participating students and apprentices or rather broadly composed Norwegian rater teams – with a stronger representation of raters from industry compared with the other countries.

The evolution of vocational competence - The Norwegian data on competence development during the three years of the COMET piloting indicate a slight progression, but with interesting differences between the vocations. The electricians stagnated in the apprenticeship period, whereas the health-care workers had a marked improvement in scores from their first year in companies to their second. For the industrial mechanics, there was a drop from the second year in school to the first year in apprenticeship, but then a major recapitulation took place for this group during the second year in apprenticeship. Compared with the international COMET projects, the latter reported stagnation when measures were repeated after one year in combined school/company training (Germany and South Africa) or in transitions from school to work (China). Several explanations have been given for these patterns, most notably the effects of transitions to new curricular context (Fachschule), to restrictive learning environments in apprenticeship in China (Rauner, Piening, & Yingvi, 2015a).

Test motivation, learning environment and vocational identity - A hierarchical log-linear analysis of the Norwegian material indicates that both the learning support of the companies and the training offices contributed significantly to the performance levels on the COMET test – with the strongest effect to be traced back to the activities provided by the training offices (Lahn et al., in print). Test motivation also had an influence on the performance of the apprentices, though the perceived relevance of the test tasks for their jobs was less important than their reported effort used in solving the assignments. Yet, in all cases the differential interactions between these factors and the three vocations were highly significant.

If we compare these results with the German and Chinese COMET scores, the items that measure the quality of learning environments are somewhat elevated for the Norwegian apprentices compared with their German and Chinese counterparts (electro-technicians), and showed the same trend on, for example, the statement: "The work tasks I carry out fit my skills and capacities." Apprentices in their second year of apprenticeship were significantly more positive in this respect when compared with those in their first year. The Norwegian figures also indicated that apprentices who scored high on professional identity also reported strong support from the workplaces as learning environments. Similar results in German and Chinese studies were interpreted by Rauner et al. (2013b) as indications that high quality learning opportunities strengthened occupational commitment and professional identity. However, in our Norwegian sample, a strong support from the companies to competence development was not associated with an increase in occupational commitment and professional identify when comparing the first-year and second-year apprentices. It needs to be

added that in line with the international COMET results, we found low, non-significant positive correlations between the professional identity index and competence levels on the tests.

In summary, it is tempting to interpret the similarity of the Norwegian scores with the other countries on general and differential measures of professional competence as evidence for satisfactory equivalence. However, given the lack of data on variability etc., we cannot rule out that these results are due to chance factors when doing the transnational comparisons. The comparability between the three vocations is not subject to the same pattern of validity threats as the ones suggested by the typology for international studies. Firstly, we may expect a higher degree of homogeneity for nearly all categories, although the measurement instrument may favor differences in familiarity with the assessment format. This privileged the health care apprentices who were training for their case-based trade examination. Secondly, by collecting fidelity data on how the COMET model was implemented in MECVET, we could turn "noise" or contextual observations into rich data, thereby making it possible to identify systematic biases, e.g., the emergence of strong group norms among the raters of the electrician apprentices' solutions, resulting in high inter-rater reliability scores but unreasonably skewed distributions. As we experienced the challenges of implementing a psychometric logic, the MECVET study gradually turned into a case of ethnostatistics as defined by Gephart (1988) as studies that "… investigate the adequacy of basic technical and practical assumptions made in statistical analyses" (p. 30).

6 Concluding comments

In this article, we have explored perspectives on holistic professional competences, both in terms of conceptual and methodological issues – leaning strongly on the models and instruments of the COMET researchers. As pointed out above, their emphasis on holistic professional competences as multidimensional in an integrative and generative way is shared by many scholars. Yet, divergent views on the assessment of such skills have surfaced. We questioned the COMET rationale for a written case format, and noticed in our analysis of the Norwegian results that there was a one-to-one correlation between the scores on the criteria "(written) presentation" and the sum scores on "professional competence." These results were corroborated when we subjected a large sample of test solutions to a quantitative analysis of writing literacies, which raises a number of questions that need to be further explored. Are the COMET instruments biased towards the measurement of generic skills in order to strengthen comparability across countries? On the other hand, an ubiquitous "textualization" of work processes (Karlsson & Nikolaidou, 2016; Rein, 2017; Lahn & Nore, 2018) could tip the balance in favor of the generic dimensions of holistic professional competences. From a methodological point of view, one could argue that international large-scale assessment studies in this respect follow the path of VET-LSA and ASCOT, referred to earlier – to develop

advanced statistical models and automatic scoring in specific vocations (Klotz & Winther, 2017). This strategy may work for certain professional domains, but as advocated here, upscaled testing for the international or educational development of holistic professional competence at the aggregate level should include mixed methods strategies (Bazeley, 2017) for the generalization of results that attend to contextual features and diversity (Mislevy et al., 2013). In MECVET, such a strategy was embarked on by accident, but more effort should be invested in establishing flexible ex ante research designs for this type of studies that combine relevant statistical models and contextual analyses.

MECVET could be positioned on the international arena for VET research in the stream of activities initiated 10 years ago by the German projects of establishing a VET-PISA (Baethge et al., 2006), as well as models and tools for international large-scale assessment in this sector. At present, these ambitions seem to dampen-, and are replaced by projects that have a narrower focus. In this context, the MECVET results add to a growing realism on the feasibility of such enterprises given the complexity of VET – even in small countries. As mentioned above, Norway has approximately 180 vocations and Germany only slightly more. Furthermore, recent research on VET systems indicates stronger type differences between countries and cluster of countries (Pilz, 2016) than previous models.

References

Baartman, L.K.J., Bastiens, T.J., Kirchner, P.K., & van der Vleuten, C.P.M. (2007). Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks. Educational Research Review, 2 (1), 114–129.

Baethge, M., Achtenhagen, F., Arends, L., Babic, E., Baethge-Kinsky, V., & Weber, S. (2006). PISA-VET. A feasibility study. Stuttgart, Germany: Franz Steiner.

Baethge, M., & Arends, L. (2009). Feasibility Study VET-LSA. A comparative analysis of occupational profiles and VET programmes in 8 European countries – International report. Federal Ministry of Education and Research, Germany: Vocational Training Research, volume 8, Bielefeld: Bertelsmann Verlag.

Baumert, J. (2009). International comparisons: Strengths and weaknesses – and how to overcome the weaknesses. In F. Oser, E.G. John, U. Renold, E. Winther, & S. Weber (Eds.), VET Boost: Towards a Theory of Professional Competencies. Essays in Honor of Frank Achtenhagen (pp. 25–37). Rotterdam, Netherlands: Sense.

Bazeley, P. (2017). Integrating Analyses in Mixed Methods Research. Los Angeles, CA: Sage Publ.

Beckett, D. (2008). Holistic competence: Putting judgements first. Asia Pacific Education Review, 9 (1), 21–30.

Benner, P. (1982). From novice to expert. American Journal of Nursing, 82(3), 402–407.

Bremer, R., Rauner, F., & Röben, P. (2001). Experten-Facharbeiter - Workshops als Instrument der berufswissenschaftlichen Qualifikationsforschung. In F. Eicker & A.W. Petersen (Eds.), Mensch-

Maschine-Interaktion'. Arbeiten und Lernen in rechnergestützten Arbeitssystemen in Industrie, Handwerk und Dienstleistung (pp. 211–231). Baden-Baden, Germany: Nomos.

Brewer, L. (2013) Enhancing youth employability: What? Why? and How? Guide to core work skills. Geneva: International Labour Office.

Bybee, R. W. (1997). Achieving scientific literacy: From purposes to practices. Portsmouth, NH: Heinemann.

Cairns, L., & Malloch, M. (2017) Competence, Capability and Graduate Attributes. In M. Mulder (Ed.), Competence-based Vocational and Professional Education Bridging the Worlds of Work and Education (pp. 103–124). Heidelberg, Germany: Springer.

Carbonell, K.B., Könings, K.D., Siegers, M., & van Merriënboer, J.V.G. (2016). Measuring adaptive expertise: Development and validation of an instrument. European Journal of Work and Organizational Psychology, 25 (2), 167–180.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), Handbook of mixed methods in social and behavioral research (pp. 209–240). Thousand Oaks, CA: Sage.

Deitmer, L., Hauschildt, U., Rauner, F., & Zelloth, H. (2013) (Eds.). The Architecture of Innovative Apprenticeship Dordrect, Netherlands: Springer Science CBusiness Media.

Dietzen, A. (2017). Knowledge Concepts in Competence-based VET. Research Perpsectives on Cognitivist and Social Constructivist Approaches. In M. Mulder (Ed.), Competence-based Vocational and Professional Education Bridging the Worlds of Work and Education (pp. 719-738). Heidelberg: Springer.

Dreyfus, S.E. (2004). The Five-Stage Model of Adult Skill Acquisition. Bulletin of Science Technology & Society, 24, 176–181.

Ericsson, K.A. (Ed.) (2009). Development of Professional Expertise: Toward Measurement of Expert Performance and Design of Optimal Learning Environments. Cambridge, UK: Cambridge Univ. Press.

Fischer, R., Hauschildt, U., Heinemann, L., & Schumacher, J. (2015). Erfassen beruflicher Kompetenz in der Pflegeausbildung europäischer Länder. In M. Fischer, F. Rauner, & Z. Zhao (Eds.), 2015. Kompetenzdiagnostik in der beruflichen Bildung. Methoden zum Erfassen und Entwickeln beruflicher Kompetenz: COMET auf dem Prüfstand, Berlin, Germany: LIT Verlag.

Fischer, M., Rauner, F. , & Zhao, Z. (Eds.) (2015). Kompetenzdiagnostik in der beruflichen Bildung. Methoden zum Erfassen und Entwickeln beruflicher Kompetenz: COMET auf dem Prüfstand. Berlin, Germany: LIT Verlag.

Gephart, R.J. (1988). Ethnostatistics. Newbury Park, CA: Sage Publ.

Haasler, B., & Erpenbeck, J. (2008). Assessing vocational competences. In F. Rauner, & R. MacLean (Eds.), Handbook of technical and vocational education and training research. Springer Media.

Hacker, W. (1973). Allgemeine Arbeits- und Ingenieurspsychologie. Berlin, Germany: VEB Deutscher Verlag der Wissenschaften.

Hager, P. (2017). The integrated view on competence. In M. Mulder (Ed.), Competence-based Vocational and Professional Education: Bridging the Worlds (pp. 203–228). Springer Media.

Hager, P., Gonczi, A., & Athanasou, J. (1994). General Issues about Assessment of Competence. Assessment & Evaluation in Higher Education, 19 (1), 3–16.

Havighurst, R.J. (1974). Developmental tasks and education (3. Ed.). New York: David Mc Kay Company.

Karlsson, A.M., & Nikolaidou, Z. (2016). The Textualization of Problem Handling: Lean Discourses Meet Professional Competence in Eldercare and the Manufacturing Industry. Written Communication, 33(3), 275–301.

Kirsch, I., Lennon, M., von Davier, M., Gonzalez, E., & Yamamoto, K. (2013). On the Growing Importance of International Large-Scale Assessments. In M. von Davier, E. Gonzalez, I. Kirsch, & K. Yamamoto (Eds.), The Role of International Large-Scale Assessments: Perspectives from Technology, Economy, and Educational Research (pp. 1–12). Dordrecht: Springer Science+Business Media.

Klotz, V.K., & Winther, E. (2017). Assessing Tomorrow's Potential: A Competence Measuring Approach in Vocational Education and Training. In D. Leutner, J. Fleischer, J. Grünkorn, & E. Klieme (Eds.), Competence Assessment in Education Research, Models and Instruments. (pp. 221–246). Cham, Switzerland: Springer International.

Kuger, S., & Klieme, E. (2016). Dimensions of Context Assessment. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), Assessing Contexts of Learning An International Perspective (pp. 3-38). Cham, Switzerland: Springer International.

Lahn, L.C. (2010). Professional learning as epistemic trajectories. In S. Ludvigsen & R. Säljö (Eds.), Learning across sites. New tools, infrastructures and practices. (pp. 53–68). Oxford, UK: Pergamon.

Lahn, L.C. & Nore, H. (2018). ePortfolios as hybrid learning arenas in vocational education and training (VET). In Choy, G.B. Wärwick, & V. Lindberg (Eds.), Integration of vocational education and training experiences: Purposes, practices and principles. Springer Publ. Lahn, L.C., Nore, H., Johannessen, H., & Bakken, R. (in print). Sluttrapport (Final report) for Measuring Competence Development in VET (MECVET). Kjeller, Norway: OsloMet.

Le Deist, J., & Winterton, J. (2005). What Is Competence?, Human Resource Development International, 8 (1), 27–46.

Leont'ev (1981). Problems of the development of mind. Moscow, Russia: Progress Press.

Lum, G. (2013) Competence: A tale of two constructs, Educational Philosophy and Theory, 45(12), 1193-1204.

Maddox, B., Zumbo, B.D, Tay-Lim, B., & Qu, D. (2015). An Anthropologist Among the Psychometricians: Assessment Events, Ethnography, and Differential Item Functioning in the Mongolian Gobi, International Journal of Testing, 15 (4), 291–309.

Maxwell, J.A., & Chmiel, M. (2014). Generalization in and from Qualitative Analysis. In U. Flick (Ed.), The SAGE Handbook of Qualitative Data Analysis (pp. 540–549). London, UK: Sage.

Mieg, H.A. (2001). The Social Psychology of Expertise: Case Studies in Research, Professional Domains, and Expert Roles. Mahwah, NJ: Lawrence Erlbaum.

Mislevy, R.J., Haertel, G., Cheng, B.H., Ructtinger, L., DeBarger, A., Murray, A., Rose, D., Gravel, J., Colker, A.M., Rutstein, D., & Vendlinski, T. (2013). A "conditional" sense of fairness in assessment. Educational Research, & Evaluation, 19 (2/3), 121–140.

Monnier, M. (2015). Difficulties in Defining Social-Emotional Intelligence, Competences and Skills - a Theoretical Analysis and Structural Suggestion. International Journal for Research in Vocational Education and Training (IJRVET), 2(1), 59–84.

Mulder, M. (2007). Competence – the essence and use of the concept in ICVT. European Journal of Vocational Training, 40 (1), 5–21.

Mulder, M. (Ed.) (2017). Competence-based Vocational and Professional Education Bridging the Worlds of Work and Education. Heidelberg, Germany: Springer.

Oesterreich, R., & Volpert, W. (1986). Task Analysis for Work Design on the Basis of Action Regulation Theory. Economic and Industrial Democracy, 7 (4), 503–527.

Olsen, O.J. (2009). Feasibility Study VET-LSA. National report from Norway. Bergen, Norway: Department of Sociology/University of Bergen.

Pellegrino, J.W., Hilton, M.I., & National Research Council (Eds.) (2012). Education for Life and Work: Developing Transferable Knowledge and Skills in the 21. Century. Washington DC: The National Academic Press.

Pilz, M. (2016). Typologies in Comparative Vocational Education: Existing Models and a New Approach. Vocations and Learning, 9, 295–314.

Ragin, C.C. (1987). The Comparative Method. Moving Beyond Qualitative and Quantitative Strategies. Berkeley, CA: University of California Press.

Rauner, F., Heinemann, L., & Maurer, A. (2011). COMET China: Implementation of learning areas. Testing a competence model as a basis for test tasks and learning tasks (pilot project large-scale competence diagnostics in Beijing VET schools and VET colleges). Bremen, Germany: Working Paper of FG I:BB, Universität Bremen.

Rauner, F., Heinemann, L., Hauschildt, U., & Piening, D. (2012). Project Report: COMET-Pilot Test South Africa including the Pilot Test Vocational Identity/Occupational Commitment and Results of the QRC Pilot Project,). Bremen, Germany: Working Paper of FG I:BB, Universität Bremen.

Rauner, F., Heinemann, L., & Hauschildt, U. (2013a). Measuring Occupational Competences: Concept, Method and Findings of the COMET Project. In L. Deitmer, U. Hauschildt, F. Rauner, & H. Zelloth (Eds.), The Architecture of Innovative Apprenticeship (159-175). Dordrecht, Netherlands: Springer Science CBusiness Media.

Rauner, F., Heinemann, L., Maurer, A., Haasler, B., Erdwien, B., & Martens, T. (2013b). Competence-Development and Assessment in TVET. Theoretical Framework and Empirical Results (COMET). Dordrecht, Netherlands: Springer Verlag.

Rauner, F., Heinemann, L., Maurer, A. Ji, L., & Zhao, Z. (2014). Messen beruflicher Kompetenzen. Band III. Drei Jahren KOMET Testerfahrung. Berlin, Germany: LIT Verlag.

Rauner, F., Piening, D., & Yingvi, Z. (2015a). Stagnation der Kompetenzentwicklung – und wie sie überwunden werden kann. Bremen, Karlsruhe, Oldenburg, Weingarten, Germany: A+B Forschungsberichte Nr. 18/2014 A+B Forschungsnetzwerk.

Rein, V. (2017). Towards the Compatibility of Professional and Scientific Learning Outcomes: Insights and Options in the Context of Competence Orientation. International Journal for Research in Vocational Education and Training (IJRVET) 4(4), 325–345

Rutkowski, L., & Rutkowski, D. (2018). Improving the Comparability and Local Usefulness of International Assessments: A Look Back and A Way Forward, Scandinavian Journal of Educational Research, 62 (3), 354–367.

Rychen, D.S., & Salganik, L.H. (2003). Key Competencies for a Successful Life and a Well-Functioning Society. Cambridge, MA: Hogrefe & Huber Publ.

Seeber, S., & Nickolaus, R (2010). Kompetenz, Kompetenzmodelle und Kompetenzentwicklung in der beruflichen Bildung. In R. Nickolaus, G. Pätzold, U. Reinisch, & T. Tramm (Eds.), Handbuch Berufs- und Wirtschaftspädagogik (pp.. 247–257). Bad Heilbrunn: Klinkhardt.

Smith, E., & Comyn, P. (2003). The development of employability skills in novice workers. Adelaide: NCVER.

Tigelaar, D. E. H., & van der Vleuten, C. P. M. (2014). Assessment of professional competence. In S. Billett, C. Harteis, & H. Gruber (Eds.), International handbook on research in professional and practice-based learning, (pp. 1237–1269). Dordrecht: Springer.

van de Vijver, F.J., & He, J. (2015). Bias Assessment and Prevention in Non-cognitive Outcome Measures in Context Assessments (pp. 229-253). In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), Assessing Contexts of Learning An International Perspective. Cham, Switzerland: Springer International.

van der Vleuten, C., Sluijsmans, D., & Joosten-ten Brikke, D. (2017). Competence Assessment as Learner Support in Education. In M. Mulder (Ed.), Competence-based Vocational and Professional Education Bridging the Worlds of Work and Education (pp. 607–630), Heidelberg, Germany: Springer.

Viswanathan, M. (2005). Understanding the intangibles of measurement in the social sciences. In G. Walford, E. Tucker, & M. Viswanathan (Eds.), The Sage Handbook of Measurement (pp. 285–312). Los Angeles, CA: Sage.

Weinert, F. E. (2001) Concept of competence: a conceptual clarification. In Rychen, D.S. and Salganik, L.H. (Eds), Defining and Selecting key Competencies (pp. 45–66). Göttingrn: Hogrefe.

Winterton, J. (2012). Varieties of Competence: European Perspectives. In M. Pilz (Ed.), The Future ofVocational Education and Training in a Changing World, (pp. 455–480). Cham: Springer Int. Publ.

Zhao, Z., Zhang, Z., & Rauner, F. (2016). KOMET-Based Professional Competence Assessments for Vocational Education and Training (VET) Teachers in China. In M. Pilz (Ed.), Vocational Education and Training in Times of Economic Crisis Lessons from Around the World (pp. 231–254). Cham, Switzerland: Springer International Publ.

Biographical notes

Dr Leif Chr. Lahn is a professor at the Department of Education, University of Oslo, and adjoint professor at the Department of Vocational Teacher Education, OsloMet – Oslo Metropolitan University. His research areas are workplace learning and ICT, vocational education and mixed methods research.

Hæge Nore is a professor at the Department of Vocational Teacher Education, OsloMet – Oslo Metropolitan University. Her research interests are vocational education, assessment in VET and recognition of informal and unformal competences.