Assessing the Development of Collaborative Knowledge Work Competence: Scales for Higher Education Course Contexts

Hanni Muukkonen, Minna Lakkala, Pekka Lahti-Nuuttila, Liisa Ilomäki, Klas Karlsgren & Auli Toom


To link to this article: https://doi.org/10.1080/00313831.2019.1647284

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 14 Aug 2019.

Article views: 117

View related articles

View Crossmark data
Assessing the Development of Collaborative Knowledge Work Competence: Scales for Higher Education Course Contexts

Hanni Muukkonen a,b, Minna Lakkala b, Pekka Lahti-Nuuttila c,d, Liisa Ilomäki b, Klas Karlgren e and Auli Toom b

aFaculty of Education, University of Oulu, Oulu, Finland; bFaculty of Educational Sciences, University of Helsinki, Helsinki, Finland; cDepartment of Phoniatrics, Helsinki University Hospital, Helsinki, Finland; dDepartment of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland; eDepartment of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden

ABSTRACT

The necessity to learn competence for collaborative knowledge work during higher education (HE) is accepted widely, but continued work is required to explicate how to define and assess such competence. In this article, the development and validation of a questionnaire for assessing the development of collaborative knowledge work competence is based on object-bound collaborative knowledge creation practices. In total, 546 students responded to a questionnaire on Collaborative Knowledge Practices (CKP). The data were analysed for measurement invariance for two groups of HE students in media engineering and life sciences. Seven scales of the CKP were found to measure course-related learning of collaboration, integration of personal and collective efforts, development through feedback, persistent development of knowledge objects, understanding of different disciplines and related expertise, interdisciplinary collaboration, and using digital technology. The CKP questionnaire scales can be used as a generic self-evaluation tool for students on course-based learning outcomes.

ARTICLE HISTORY

Received 22 January 2018
Accepted 18 July 2019

KEYWORDS

Higher education; collaboration; assessment; competence development

Introduction

It has been recognized internationally that higher education (HE) studies should provide students with versatile expertise to effectively manage the challenges of working life in terms of interdisciplinary projects and tasks (Barrie, 2012; OECD, 2013; Virtanen & Tynjälä, 2018). Knowledge work is characterized by a wealth of information, changing collaborators and technologies, the contestable nature of what counts as valid knowledge, and the need to deal with uncertainty (Barnett, 2004). The necessity to learn related knowledge work competence during HE studies has been accepted, but continued work is required in order to more explicitly define such capabilities, identify the most effective pedagogical methods to teach them, and, further, discover how to assess the learning of them.

University education often emphasizes critical and analytical thinking, well-defined conceptual knowledge, and the acquisition of professional methods and practices (Bereiter, 2002; Muukkonen, Hakkarainen, & Lakkala, 2004; Slotte & Tynjälä, 2003). However, discipline-specific courses are less frequently pedagogically designed to systematically teach working collaboratively around complex,
open problems or iterative development of epistemic objects (Bovill, Bulley, & Morss, 2011; Goodyear & Zenios, 2007). This would be valuable because, in knowledge work, epistemic objects (or knowledge objects) are at the core of collaborative efforts in academic professions; for instance, in new product development, the production of research articles, or the creation of new guidelines for procedures. Students should learn how various tools, artefacts, and conceptual and practical knowledge are used and created in knowledge work processes (cf. Knorr Cetina, 2001; Miettinen, 2005; Rheinberger, 1997).

Continual efforts have been made to define generic or discipline-general competences and their measurement (e.g., Greiff et al., 2014; Hyytinen, Nissinen, Ursin, Toom, & Lindblom-Ylänne, 2015; Kember & Leung, 2009; OECD, 2013; Shavelson, 2010; Strijbos, Engels, & Struyven, 2015). Prior research and available assessment instruments have primarily focused on explicating individual dimensions of generic competences, including, e.g., critical and analytical thinking (e.g., Hyytinen et al., 2015; Kember & Leung, 2009), problem-solving (Neubert, Mainert, Kretzschmar, & Greiff, 2015), self-management of learning, communication skills, and information and digital literacy (e.g., Strijbos et al., 2015; Young & Chapman, 2010). Other research has shown that there are collective dimensions to be taken into account in competence development: collaborative idea generation and knowledge production, as well as epistemic, regulative, and metacognitive competence utilized in shared knowledge advancement (Damša, 2014; Damša, Kirschner, Andriessen, Erkens, & Sins, 2010; Hadwin, Järvelä, & Miller, 2011; Muukkonen & Lakkala, 2009; Panadero, Kirschner, Järvelä, Malmberg, & Järvenoja, 2015). These involve dimensions of competence required in collaborative working often referred to as interpersonal, reciprocal communication, and teamwork capabilities. Further, it has been suggested that the collective dimensions of competence do not develop without distinct exposure to educational practices that involve collaborative engagement in complex working processes around messy problems (Barrie, 2012; Damša, 2014; Lakkala, Toom, Ilomäki, & Muukkonen, 2015; Mu & Gnyawali, 2003).

What is lacking are the means to (a) follow how pedagogical practices are related to collective dimensions of competence development, (b) explicate how collaborative practices in knowledge creation could be defined for measuring competence development, and (c) provide course-based feedback for students and teachers in relation to learning such practices.

To answer how educational practices can advance being competent in collaborative knowledge work practices, there is a need to operationalize competence development related to knowledge work and knowledge creation from both the individual and collective perspective. The socio-cognitive research paradigm has generated well-developed means to measure the learning of various skills and well-defined competence dimensions (Binkley et al., 2012; OECD, 2013; Strijbos et al., 2015). At the same time, the socio-cultural research paradigm has emphasized the actual praxis, routines and reconfiguration of practices within activity systems (e.g., Greeno & Engeström, 2014; Miettinen & Virkkunen, 2005). In the latter paradigm, the idea of defining measurable competence dimensions is largely rejected based on the analytical perspective that “the phenomenon to be explained is transformation, change, and learning that occurs in the activity system as a whole, rather than the learning of any one participant only” (Greeno & Engeström, 2014, p. 140). In an attempt to integrate these positions, the present work is based on an argument that we ought to understand better how various pedagogical practices and their actual enactment by students is linked to their learning of more advanced means to engage in collaborative knowledge work practices. This entails examining the learning of complex practices in context instead of taking the assessment of learning out of the context within which the competence is actively used (cf. Shavelson, 2010). Hence, we ask students to evaluate their competence learning based on enacted collaborative practices. In parallel, it is important to collect data to understand the characteristics of the context, and the expected and produced outcomes of the learning activities.

The argument is expanded in the following sections by reviewing the existing literature on assessing dimensions of generic competence, delineating our theoretical and empirical basis for the development of collaborative knowledge work competence, presenting the instrument development and validation work, and, finally, discussing the implications and challenges of taking the practice perspective on competence learning in the context of HE.
First, we address the definition of competence, which is not a straightforward matter. The terms competence, competency, skill, capability, and ability are used interchangeably or with specific meanings in the literature. For generic dimensions of competence, we follow the definition of Strijbos et al. (2015): “generic competences are competences applicable across different professional contexts and beyond the field of study” (p. 20). Here, it is notable that the plural “competences” is used, presumably because these competences have been spelt out as a set of conceptual, people, and personal skills. On the other hand, Mulder (2012) defined competence as a “holistic concept, which describes a person’s ability to manage in a specific context” (p. 36). In the work of Marin-Garcia, Pérez-Peñalver, and Watts (2013), the relationship between competence, capacities and skills is further defined as involving increasing complexity of know-how. Competence is formed by a set of capacities that, in turn, are formed by several skills, all of which are required for more complex professional performances (Marin-Garcia et al., 2013; see also Keinänen, Ursin, & Nissinen, 2018). We follow this approach by suggesting that competence development is related to developing expertise in various practices (e.g., by taking part in interdisciplinary collaboration) and, particularly in the context of this study, in collaborative practices of knowledge work and knowledge creation.

Within the socio-cultural paradigm, knowledge practices are utilized as a broader concept encompassing both the routine and creative aspects of knowledge work, which involve the integration of individual and collective activities during working and co-creational processes (Hakkakainen, 2009; Schatzki, 2000). We take knowledge work competence to refer to the knowledge, skills and dispositions to act, study and work intentionally and effectively both individually and together with others in various contexts. Knowledge work competence enables the solving of complex problems and taking part in creating knowledge and novel solutions by using the community’s collective, technology-mediated efforts. According to this definition, competence is not equivalent to a specific skill or activity (e.g., “is able to post an expert comment to an online discourse”) that can be assessed as an acquired skill. Rather, we understand competence learning similarly to Mulder and Marin-Garcia and colleagues; that is, as being able to take part in, advance and monitor both individual and collective activities during knowledge work in progressively more expert ways.

Assessing Generic Aspects of Competence

Having set the argument above, it is nevertheless in order to acknowledge that studying collaborative learning and competence development are relatively challenging phenomena to be investigated by utilizing questionnaire surveys. This section presents prior instruments developed to measure generic dimensions of competence. Examples with details are provided in Table 1. We have included only those measurements that involve some aspect of collaboration, e.g., teamwork, interaction, or networks. We summarize the focus of the instrument, how and by whom the assessment is carried out, and the context of the assessment. For the context, the term general refers to a general assessment of one’s own competence level, which is not related to any specific learning activity or context at hand. Two examples derived from the context of a degree programme and two took place in a digital environment.

Kember and Leung (2009) and Sleap and Reed (2006) reported on questionnaires (see Table 1) for measuring the development of students’ generic capabilities within a degree programme. They aimed to identify relative strengths and weaknesses within a programme and then to model the types of teaching and learning environments that would be influential in terms of the development of generic capabilities. For assessment, these provide the means to carry out summative assessment based on large numbers but need formative assessment at a more detailed level to accompany them (Norman, Norcini, & Bordage, 2014).

Several frameworks and instruments for assessing more specific dimensions of competence development have been explicated, including innovation competence (Keinänen et al., 2018; Marin-Garcia et al., 2013), group-work skills (Cumming, Woodcock, Cooley, Holland, & Burns, 2015), and
discipline-generic competence (Greiff et al., 2014; OECD, 2013; Strijbos et al., 2015). However, the PISA 2014 measures are not included, for example, because they focus on an individual’s cognitive capacity when working alone (OECD, 2014). Reports on the OECD project AHELO (OECD, 2013) focused on the competences regarded as the most important in HE: critical thinking, discipline knowledge, problem-solving, teamwork, communication, professional skills, ethics and values, creativity, and learning to learn.

Another development area is assessment in digital environments, for instance in collaborative problem-solving in microworlds (Neubert et al., 2015) or through epistemic games to highlight the interplay between professional practices and generic competences (Rupp, Gushta, Mislay, & Shaffer, 2010; Shaffer, 2006) (see Table 1). Epistemic games address how students learn the skills necessary to think, act and interact in collaboration with others in productive ways while tackling complex tasks around authentic problems. The research focuses on developing digital games simulating professional knowledge practices as well as on automatic collection and analysis of qualitative and quantitative evidence about learners’ developing competences (Rupp et al., 2010). The present study shares the focus on knowledge practices but is not built within a certain digital environment.

The reviewed assessment instruments capture a wide range of generic dimensions and knowledge work related competences. From the perspective of our aim to examine learning gained from any regular coursework involving various forms of collaborative practices, these were considered insufficient due to their emphasis on the general and non-contextualized nature of what is assessed or working within a confined digital world. In contrast, the present study aims for a contextualized self-evaluation of the extent of competence learning on a completed course (cf. Zlatkin-Troitschanskaja, Shavelson, & Kuhn, 2015).

### Table 1. Examples of instruments for assessing competence development involving collaboration and interaction.

<table>
<thead>
<tr>
<th>Reported in</th>
<th>Focus of instrument</th>
<th>Assessment by</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kember and Leung (2009)</td>
<td>Generic competences: critical and creative thinking, self-managed learning, adaptability, problem-solving, communication skills, interpersonal skills, and computer literacy</td>
<td>Questionnaire for measuring students’ generic capabilities development on a degree programme; evaluation of own competence level</td>
<td>Degree programme-related</td>
</tr>
<tr>
<td>SLEEP and REED (2006)</td>
<td>Specific work skills: personal work skills, business skills, and interactive work skills</td>
<td>Questionnaire on students’ self-assessment of skills development during university studies; evaluation of own competence level</td>
<td>Degree programme-related</td>
</tr>
<tr>
<td>LIZZIO and WILSON (2004)</td>
<td>Generic capabilities, e.g., self-management, interpersonal, adaptability and learning, problem-solving and decision-making, conceptual and analytical, team and group</td>
<td>Student self-assessment of how skills list items were characteristic of them, and their relevance for the present course, future work, and personal interest; capability profile</td>
<td>General</td>
</tr>
<tr>
<td>CUMMING et al. (2015)</td>
<td>Groupwork skills questionnaire; attitudes and conceptions</td>
<td>Students assess how likely they are to employ specific skills: task or interpersonal group-work skills</td>
<td>General</td>
</tr>
<tr>
<td>SHAFER (2006); RUPP et al. (2010)</td>
<td>Epistemic games learning simulation; designed to allow learners to develop discipline-specific expertise under realistic constraints</td>
<td>Complex performance-based assessments; learner performance as process and product data</td>
<td>In digital environment; experience of professional practices within a discipline</td>
</tr>
<tr>
<td>NEUBERT et al. (2015)</td>
<td>Complex problem-solving and collaborative problem-solving in interactive settings (multiple problem-solvers working on the same problem)</td>
<td>Researchers and practitioners assess individuals’ performance in computer-based microworlds; simulation of complex and collaborative problems that need to be actively explored and controlled</td>
<td>In digital environment</td>
</tr>
</tbody>
</table>
Theory-based Questionnaire Development: The Knowledge Creation Metaphor of Learning

As the theoretical background, the questionnaire item development builds on the socio-cultural theories of learning and particularly on the knowledge creation metaphor (Paavola & Hakkarainen, 2005). To expand the two metaphors of learning (acquisition and participation) put forward by Sfard (1998), Paavola and Hakkarainen (2005) added the knowledge creation metaphor. The acquisition metaphor of learning addresses assimilation of prevailing knowledge and the individual’s mental models and strategies of learning. Such practices are quite central in traditional HE courses, i.e., attending lectures, writing essays, and reading for exams. The participation metaphor highlights the adaptation to existing cultural and communal practices and the dialogical practices of learning (Lave & Wenger, 1991; Sfard, 1998); for example, field-training periods of study when students working in various institutions and under supervision become familiar with the practices, tools, and cultural knowledge of a particular working community.

The knowledge creation metaphor adds the object-centred approach to learning. The presence of artefacts, practices, and products – “objects” – and shared work on advancing them is fundamental. Knowledge creation is about setting up practices of working with knowledge and channelling the efforts of the participants toward the epistemic objects in ways that educe knowledge advancement; the development of ideas is one component of which (Hakkarainen, 2009; Paavola & Hakkarainen, 2005) and collective practices and products are others. The objects are seen to mediate knowledge advancement by structuring the efforts of the participants so that they work on these objects, negotiate meanings, expand, and version them using collaborative effort. Related theoretical frameworks, such as knowledge building (Scardamalia & Bereiter, 2006), have also argued for a better understanding and scaffolding of collective expertise and the epistemic practices of learning communities.

There exist prior findings on generic competence learning during collaborative knowledge creation. Damşa et al. (2010) reported on students’ learning about the regulative and epistemic aspects of collaboration, particularly the benefits of shared epistemic agency in relation to the advancement of shared outcomes. Other studies (Muukkonen & Lakkala, 2009; Muukkonen, Lakkala, & Hakkarainen, 2005) have used open-ended questions to solicit students’ self-evaluation of competence learning during courses. These have provided evidence of learning about domain-specific competences, e.g., designing software architecture or setting up a business plan, as well as domain-generic competences such as learning to coordinate teamwork, taking initiative, creating new knowledge, integrating theory and practice, and collaborating with work-life experts.

A study on inter-professional medical simulation training (Karlgren, 2012) suggested that teams may lack a clear and established practice for analysing teamwork and communication, and hence, may not work optimally in critical care situations. The applied simulation training and debriefing guided participants in identifying and analysing critical teamwork incidents based on conceptual guidelines for practices in critical care situations, and provided knowledge of efficient teamwork and communication. In another study on interdisciplinary collaboration on an HE course, the participants were found to struggle with professional means to work on business ideas, communication with potential customers, and the creation of functioning technical solutions (Kosonen, Muukkonen, Lakkala, & Paavola, 2012). These studies highlight the competence related to interdisciplinary collaboration and the adoption of profession-related analytical tools and practices.

These studies reported on competence related to aspects of knowledge work that concern professionals across disciplines, which are not usually emphasized in curricula. Disciplinary knowledge or discipline-specific competence are more routinely explicated in curricula and course objectives with available best practices for assessment, and they also vary in terms of the specific characteristics of the discipline. Therefore, an objective for instrument development was to build an instrument usable across a wide variety of educational contexts for courses using collaboration as a pedagogical method. This initiative attempts to generate feedback for students themselves, lecturers, and the
pedagogical designers on the course-specific outcomes, but in a format that can be compared across courses and examined separately from the learning of course content.

This article presents the development and substantiation process of the Collaborative Knowledge Practices questionnaire (CKP), which is theoretically based on the knowledge creation metaphor of learning and its six design principles (Paavola, Lakkala, Muukkonen, Kosonen, & Karlsgren, 2011). Questionnaire development was initiated as a result of the lack of an instrument to measure discipline-general competence development, especially in terms of collaborative knowledge practices. The CKP questionnaire was designed to evaluate self-assessed competence development during courses that use group work or collaborative assignments as instructional approaches. Empirically, this article examines construct validity in the questionnaire (see Evers, Verboon, & Klaeijsen, 2017; Jones & Paulhus, 2014; Marsh et al., 2009). Based on a dataset from HE courses, the research questions addressed were:

1. Can measurement invariance for the questionnaire be established between the two groups (media engineering and life sciences)?
2. If satisfactory scalar invariance can be established, are there any group differences in the latent factor means?

The first research question examines the hypothesized structure of the CKP and what type of measurement invariance it holds. The second research question examines the entire data based on the two disciplinary fields. Establishment of measurement invariance is needed to reliably interpret group differences (Cheung & Rensvold, 2002). Based on the characteristics of the studied fields and courses, we expected the media engineering students to demonstrate higher means on technology use and, due to the intensive project courses, also higher means on the integration of efforts. For the life sciences, as a result of having more courses involving mixed participation of students from different degree programmes, we expected higher means on learning to understand various practices and interdisciplinarity. For the other factors, we had no expectations.

Method

Research Contexts

The data were collected from a research-intensive university and two applied sciences universities. The questionnaire was used on courses with student collaboration and some use of digital technologies. In media engineering, 24 courses at an applied sciences university were included (25–30 students, mostly in their first year of studies, 15 ECTS (European Credit Transfer System)). In life sciences, altogether eight courses were included: two first-year courses (30 and 31 students, 15 ECTS) at an applied sciences university; four first-year courses (74, 60, 144, and 117 students) and two third-year courses (33 and 30 students) of three or five ECTS at a research university.

The contexts are interesting to compare because there was distinct variation in their instructional design reflecting the educational emphasis in both domains. The media engineering (ME) courses emphasized intensive domain-typical project work, whereby students worked together to advance various artefacts and products as a team, such as project documentation, game scenarios or website designs, by using technology. The life sciences (LS) courses had a very varied scope of activities, including lectures with small-scale collaboration particularly in the large-class settings of more than 50 students, but also projects with interdisciplinary participation and customers outside of university. The LS courses were less technology-intensive than the ME courses.
Participants

Students taking part in 32 higher education courses in ME and LS representing the STEM disciplines (science, technology, engineering and mathematics) were invited to participate in the investigation. The data consisted of student self-evaluations \( (N = 546) \), with a response rate of 48.5%. ME students were, on average, younger than the LS students (see Table 2). There were more male than female students in ME, while in the LS there were fewer males than females. Six students reported their gender as other. Altogether, 44.5% of the students were female. In one course for international students \( (n = 14, 2.6\%) \), an English-language questionnaire was used.

Students were asked to answer the questionnaire and to provide their informed consent electronically at the end of their course. The researchers provided the link, and the lecturers forwarded the e-questionnaire to the students and encouraged their participation.

Questionnaire

The development of the questionnaire was based on the knowledge creation metaphor. Its six design principles (DP; Paavola et al., 2011) were used to define the characteristics of pedagogical practices that are considered central in educational settings expected to promote students’ knowledge work competence:

(DP1) Organizing activities around shared “objects”. The principle emphasizes the epistemic and organizational aim of anchoring the collective and individual efforts to a shared object. In education, an object can be, for instance, a report, concept, product, service, presentation, or documentation of a procedure.

(DP2) Supporting integration of personal and collective agency and working through developing shared objects. The aim is to organize learning activities in such a way that students take responsibility not only for their own learning but also, importantly, for the collective process and learning.

(DP3) Emphasizing development and creativity through transformations and reflection. The epistemic objects of knowledge work are characteristically open-ended and evolving (Knorr Cetina, 2001; Rheinberger, 1997). This open-endedness poses challenges for creating a shared understanding. One means of supporting sense-making and structuring the phenomenon is to utilize various ways to document and represent the phenomenon or object at hand, e.g., by drawing, filming or modelling, and reflecting on the transformational aims and processes.

(DP4) Fostering sustained processes of knowledge advancement. The objective of sustained processes is to deepen an understanding of a phenomenon and create plans, products, services, or practices that have further use after the course. Often such demand exists for objects that stem from a community’s developmental aims or questions.

(DP5) Promoting cross-fertilization of knowledge practices and artefacts across communities and institutions. During HE studies, it is central to embed complex questions, procedures, methods, and expert practices of work-life into educational settings. Cross-fertilization through interdisciplinarity may take various forms: interdisciplinary teamwork, lecturers with complementary expertise or methodologies, professionals as expert members, or customers participating in courses.

Table 2. Participants’ age and gender.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Students N</th>
<th>Courses N</th>
<th>M (SD)</th>
<th>Median</th>
<th>Male/Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Media engineering</td>
<td>273</td>
<td>24</td>
<td>22.5 (3.5)</td>
<td>22.0</td>
<td>217/50</td>
</tr>
<tr>
<td>Life sciences</td>
<td>273</td>
<td>8</td>
<td>24.2 (6.2)</td>
<td>22.0</td>
<td>82/191</td>
</tr>
<tr>
<td>Total</td>
<td>546</td>
<td>32</td>
<td>23.4 (4.9)</td>
<td>22.0</td>
<td>299/241</td>
</tr>
</tbody>
</table>
(DP6) Providing flexible tools for developing artefacts and practices. To mediate the above collaboration practices, digital tools should enable flexible mobility and synchronous and asynchronous collaboration.

To develop the CKP questionnaire, activities and practices related to the aims described by the design principles were explicated as 48 items. The students were asked to evaluate how well each item corresponded to what they had learned during the course: “During the course I have learned …”, e.g., “to develop ideas further together with others” or “to present my expertise to representatives of another field”. The items were rated on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree); responding “Not applicable to this course” (0) was also possible.

These practices were identified in prior research (e.g., Lakkala et al., 2015; Muukkonen & Lakkala, 2009). The intent was to ask about one’s learning based on taking part in various practices. Hence, the aim was to find items reflecting learning of competence to engage in carrying out the collaborative practices described by the design principles. It is important to note that not all of the design principles are focal to a given course; indeed, only one or two may be significant and thus, in terms of outcomes, self-assessed learning should be compared to the expected learning outcomes of that particular course. Based on this, the CKP outcomes can be compared across types of course and disciplinary context.

The participants provided information about their age, gender, and degree course. Based on the pilot course responses, items were examined through an explorative factor analysis, comparison with theoretical design principles, and a cultural adaptation process (Beaton, Bombardier, Guillemin, & Ferraz, 2000). Items that were ambiguous or produced single-item factors were excluded. The second iteration involved the remaining 27 items and answers to the statements from both iterations were used as data in the validation of the present study.

**Data Analysis**

Questionnaire data from 32 courses were screened. The responses of ten students were excluded, based either on missing data ($n = 2$) or because they provided too many (ten or more) “not applicable” responses to items ($n = 8$).

Measurement invariance of ordered categorical CKP items was analysed using exploratory structural equation modelling (ESEM) (Asparouhov & Muthén, 2009; Marsh et al., 2009). ESEM allows less restrictive measurement than traditional confirmatory models, for example cross-loadings. The criteria are less stringent than for confirmatory factor analysis (CFA), which is unlikely to result in a fitting model if ambiguity exists in the number of factors to extract, more than a few cross-loadings, and moderate factor correlations (Jones & Paulhus, 2014; Marsh et al., 2009; Marsh, Morin, Parker, & Kaur, 2014). We first carried out an explorative oblique Geomin rotated factor analysis with WLSMV estimation and THETA parameterization using the Mplus statistical package (version 7.2; Muthén & Muthén, 1998–2011) to assess the proper number of factors. Based on this initial analysis, we then proceeded with seven factors to analyse invariance between two student groups with ESEM, again using oblique Geomin rotation, weighted least square mean and variance adjusted (WLSMV) estimation and THETA parameterization (Asparouhov & Muthén, 2009; Marsh et al., 2009; Millsap, 2011; Millsap & Yun-Tein, 2004). In the evaluation of configural, metric, and scalar invariance we used goodness-of-fit indices rather than traditional $\chi^2$-difference testing as they are not affected by sample size and, specifically, CFI is also independent of model complexity and the overall fit measures (Cheung & Rensvold, 2002). It has been reported that, for models with more than 400 cases, the chi-square is almost always statistically significant. Chi-square is affected by the size of the correlations in the model, so that the larger the correlations, the poorer the fit (Kenny & McCoach, 2003). More than 500 cases were included, and larger factor correlations were found. Hence, it was more suitable to use goodness-of-fit indices.

We have chosen to use the term null model for the total group ESEM rather than baseline model because, during the invariance examination, both configural and metric models can also be
considered as baseline models at respective points. For the total group ESEM, the null model is one with zero loadings, zero factor covariances, and free threshold estimates. For the invariance models, equality of the latent means is also included in the null model.

We report three oft-used indices – the comparative fit index (CFI), Tucker–Lewis index (TLI) and root mean square error of approximation (RMSEA) – as they are appropriate with WLSMV estimation. These indices take into account model complexity. The CFI is considered a measure least affected by sample size. Both CFI and TLI express incremental fit by comparison to a null model of uncorrelated variables. Population discrepancy is considered by the RMSEA as it is sensitive to increasing misfit due to sample size increases. A value > 0.95 is recognized as indicative of good fit for CFI and TLI (Hu & Bentler, 1999). For RMSEA, 0.01 indicates an excellent fit and 0.05 a good fit, with a generally agreed cut-off point of 0.07.

**Results**

**Establishing Measurement Invariance**

First, we examined whether invariance for the questionnaire structure could be established between the two domains: media engineering and life sciences.

Three models – configural, metric, and scalar – were compared to examine the invariance (Table 3). The configural model does not impose any invariance constraints but does impose a restriction whereby the number of factors must be the same between two groups. The metric model tests the equality of factor loadings constraining them to be the same. The scalar model adds the test of the equality of item thresholds with ordered categorical items. The scalar model also provides a good fit for the data (CFI = 0.979; TLI = 0.975; RMSEA = 0.045). Standardized loadings, latent means, and factor correlations for both groups in the scalar model are presented in Table 4. Many factor correlations were of different magnitudes between the groups.

The ESEM allows items to load on several factors. Here, several items had cross-loadings especially on factors F1 (learning to collaborate), F2 (integrate working), F7 (understand disciplines), and F5 (persistent development). When we looked at factor correlations and cross-loadings results together, we found a reasonable accord with the theoretical model. For example, items 4 (to ask questions relating to the practices of another field) and 28 (how useful it is to learn the working practices of other fields and organizations) have reciprocal cross-loadings on F7 (understand disciplines) and F3 (interdisciplinary collaboration) that are quite justifiable. Interestingly, the factors have substantial correlation in the ME group.

The exploit technology factor (F6) had only the theoretically designated items loading high on it, but this factor did have quite a high factor correlation with F5, persistent development, in the ME group, which suggests that the sustained efforts to develop artefacts and objects were made using technology.

Still, some controversial loadings emerged. Item 11 (to define sub-goals for the collaborative work) did not load very highly on F2 (integrate efforts) but instead loaded on F1 (learning to collaborate) and F5 (persistent development). This could be explained by the fact that students considered defining sub-goals from the perspective of how to engage in collaboration and its sustained processes as knowledge advancement rather than how to combine individual and collaborative activities during collaboration.

The CKP items and scales based on the factors are presented in Table 5. Overall, the factor analysis results provided a good fit for the data with seven factors. However, due to the oblique rotation solution, the factor loadings were partially overlapping. Therefore, on a few items (L8, L10, L16), theory-based decisions were made on their placement on the scales, to focus on the collaboration on shared object with items 8 and 10 and the continued efforts on item 16. The relationship to design principles is also presented in Table 5, showing that design principle 3, emphasizing development and creativity through transformations and reflection, has minor representation in the questionnaire.
<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>NFP</th>
<th>$\chi^2$ / df / $p$ for Difference</th>
<th>CFI</th>
<th>$\Delta$ CFI</th>
<th>TLI</th>
<th>$\Delta$ TLI</th>
<th>RMSEA</th>
<th>90% CI for RMSEA</th>
<th>Model description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total group</td>
<td>474.498</td>
<td>183</td>
<td>276</td>
<td>na</td>
<td>0.982</td>
<td>na</td>
<td>0.966</td>
<td>Na</td>
<td>0.054</td>
<td>0.048–0.060</td>
<td>Total group exploratory structural equation model</td>
</tr>
<tr>
<td>Configural model</td>
<td>606.332</td>
<td>366</td>
<td>552</td>
<td>na</td>
<td>0.985</td>
<td>na</td>
<td>0.971</td>
<td>Na</td>
<td>0.049</td>
<td>0.042–0.056</td>
<td>Item residual variances = 1; Factor means = 0; Factor variances = 1</td>
</tr>
<tr>
<td>Metric model</td>
<td>813.107</td>
<td>506</td>
<td>352</td>
<td>275.223/140 / &lt;0.001</td>
<td>0.981</td>
<td>−0.004</td>
<td>0.973</td>
<td>0.002</td>
<td>0.047</td>
<td>0.041–0.053</td>
<td>Loadings invariant; Item residual variances = 1; Factor means = 0; Factor variances = 1 in ME, free in LS</td>
</tr>
<tr>
<td>Scalar model</td>
<td>948.972</td>
<td>607</td>
<td>311</td>
<td>197.53/101 / &lt;0.001</td>
<td>0.979</td>
<td>−0.002</td>
<td>0.975</td>
<td>0.002</td>
<td>0.045</td>
<td>0.040–0.051</td>
<td>Loadings and thresholds invariant; Item residual variances = 1; Factor means = 0 in ME, free in LS; Factor variances = 1 in ME, free in LS</td>
</tr>
</tbody>
</table>

Notes: df = degrees of freedom; NFP = number of free parameters; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean squared error of approximation; CI = confidence interval for RMSEA; na = not applicable; ME = Media engineering students; LS = Life sciences students.
Table 4. ESEM solution for Scalar invariance model: standardized loadings, factor means, variances and correlations.

<table>
<thead>
<tr>
<th>Items</th>
<th>Media engineering</th>
<th>Life sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>L1</td>
<td>0.147</td>
<td>0.104</td>
</tr>
<tr>
<td>L2</td>
<td>0.029</td>
<td>−0.094</td>
</tr>
<tr>
<td>L3</td>
<td>0.081</td>
<td>−0.184</td>
</tr>
<tr>
<td>L4</td>
<td>−0.091</td>
<td>−0.104</td>
</tr>
<tr>
<td>L5</td>
<td>0.089</td>
<td>−0.034</td>
</tr>
<tr>
<td>L6</td>
<td>0.259</td>
<td>0.398</td>
</tr>
<tr>
<td>L7</td>
<td>−0.131</td>
<td>0.033</td>
</tr>
<tr>
<td>L8</td>
<td>0.270</td>
<td>0.216</td>
</tr>
<tr>
<td>L9</td>
<td>0.077</td>
<td>0.258</td>
</tr>
<tr>
<td>L10</td>
<td>0.305</td>
<td>−0.024</td>
</tr>
<tr>
<td>L11</td>
<td>0.326</td>
<td>0.115</td>
</tr>
<tr>
<td>L12</td>
<td>0.307</td>
<td>0.430</td>
</tr>
<tr>
<td>L13</td>
<td>−0.138</td>
<td>0.216</td>
</tr>
<tr>
<td>L14</td>
<td>−0.271</td>
<td>0.117</td>
</tr>
<tr>
<td>L15</td>
<td>0.108</td>
<td>0.271</td>
</tr>
<tr>
<td>L16</td>
<td>0.019</td>
<td>0.410</td>
</tr>
<tr>
<td>L17</td>
<td>0.028</td>
<td>0.356</td>
</tr>
<tr>
<td>L18</td>
<td>0.063</td>
<td>−0.016</td>
</tr>
<tr>
<td>L19</td>
<td>0.119</td>
<td>−0.040</td>
</tr>
<tr>
<td>L20</td>
<td>0.235</td>
<td>0.504</td>
</tr>
<tr>
<td>L21</td>
<td>0.178</td>
<td>−0.053</td>
</tr>
<tr>
<td>L22</td>
<td>0.516</td>
<td>0.125</td>
</tr>
<tr>
<td>L23</td>
<td>0.170</td>
<td>0.148</td>
</tr>
<tr>
<td>L24</td>
<td>0.321</td>
<td>0.394</td>
</tr>
<tr>
<td>L25</td>
<td>0.024</td>
<td>0.093</td>
</tr>
<tr>
<td>L26</td>
<td>−0.031</td>
<td>0.005</td>
</tr>
<tr>
<td>L27</td>
<td>0.150</td>
<td>0.263</td>
</tr>
</tbody>
</table>

(Continued)
Table 4. Continued.

### Factor Means

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>F6</th>
<th>F7</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>F6</th>
<th>F7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.380</td>
<td>−0.726</td>
<td>0.251</td>
<td>−0.721</td>
<td>−0.583</td>
<td>−0.599</td>
<td>0.392</td>
</tr>
<tr>
<td>p</td>
<td>0.015</td>
<td>&lt;.001</td>
<td>0.009</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Factor Variances (diagonal) and Correlations

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>F6</th>
<th>F7</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>F6</th>
<th>F7</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>1</td>
<td>1.451</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2</td>
<td>0.578</td>
<td>1</td>
<td>0.083</td>
<td>0.334</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3</td>
<td>0.243</td>
<td>0.226</td>
<td>1</td>
<td>−0.114</td>
<td>0.561</td>
<td>0.512</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F4</td>
<td>0.438</td>
<td>0.454</td>
<td>0.319</td>
<td>1</td>
<td>0.260</td>
<td>0.862</td>
<td>0.432</td>
<td>1.462</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F5</td>
<td>0.543</td>
<td>0.439</td>
<td>0.194</td>
<td>0.407</td>
<td>1</td>
<td>0.234</td>
<td>0.449</td>
<td>0.394</td>
<td>0.572</td>
<td>1.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F6</td>
<td>0.403</td>
<td>0.453</td>
<td>0.145</td>
<td>0.452</td>
<td>0.544</td>
<td>1</td>
<td>−0.135</td>
<td>0.228</td>
<td>0.201</td>
<td>0.336</td>
<td>0.089</td>
<td>1.462</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F7</td>
<td>0.057</td>
<td>0.089</td>
<td>0.635</td>
<td>0.259</td>
<td>0.212</td>
<td>0.313</td>
<td>1</td>
<td>0.430</td>
<td>0.439</td>
<td>0.128</td>
<td>0.306</td>
<td>0.118</td>
<td>−0.124</td>
<td>2.175</td>
</tr>
</tbody>
</table>

Note: Absolute factor loadings ≥ 0.25 are bold.
Differences between the Groups

Group mean differences in latent variables show that media engineering had statistically significantly higher means on F2 (integrate efforts), F4 (feedback), F5 (persistent development), and F6 (exploit technology), while life sciences had higher means on F1 (learning to collaborate), F3 (interdisciplinary collaboration), and F7 (understand disciplines). Based on course descriptions from the lecturers, the ME courses emphasized intensive project work, whereby students needed to work together to advance the various documents and products of their team by using technology. Hence, they provided higher scores on integrating efforts and learning feedback practices, persistence in development efforts, and exploiting technology. The LS courses involved a more varied scope of activities, including lectures with small-scale collaboration but also projects with interdisciplinary participation and customers outside of university. These apparently provided, on average, fewer opportunities for

Table 5. Scales, related design principles, reliability coefficients, and items of the collaborative knowledge practices questionnaire.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Design principle</th>
<th>Alpha (Cronbach)</th>
<th>Items during the course I have learned …</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning to collaborate on shared objects</td>
<td>DP1</td>
<td>0.77</td>
<td>L8 to coordinate the development of products (e.g., plans, reports, models) together with others. L10 to take responsibility for shared group work. L24 to plan collaborative work. L26 to develop ideas further together with others.</td>
</tr>
<tr>
<td>Integrating individual and collaborative working</td>
<td>DP2</td>
<td>0.72</td>
<td>L6 to understand how important the expertise of others is when developing products. L11 to define sub-goals for the collaborative work. L12 to understand the benefits of working in collaboration. L20 to accomplish challenging tasks in collaboration with others.</td>
</tr>
<tr>
<td>Development through feedback</td>
<td>DP2</td>
<td>0.72</td>
<td>L19 to receive feedback on my products (e.g., plans, reports, models) for developing them further. L23 to comment on the work of others. L27 to understand the value of commenting on work in progress. L29 to finalize products with patience.</td>
</tr>
<tr>
<td>Persistent development of knowledge-objects</td>
<td>DP4</td>
<td>0.75</td>
<td>L1 to evaluate the development of a shared product. L2 to work on products that are used later by others or myself. L9 to work on the shared products by improving them through revisions. L16 to evaluate how much effort is needed to develop a product.</td>
</tr>
<tr>
<td>Understanding various disciplines and practices</td>
<td>DP5</td>
<td>0.75</td>
<td>L3 new aspects about the practices of different organizations. L15 the practices of people with different kinds of expertise. L25 about the practices of work-life experts. L28 how useful it is to learn the working practices of other fields and organizations.</td>
</tr>
<tr>
<td>Interdisciplinary collaboration and communication</td>
<td>DP5</td>
<td>0.71</td>
<td>L4 to ask questions relating to the practices of another field. L13 to present my expertise to representatives of another field. L18 to collaborate with representatives of other fields.</td>
</tr>
<tr>
<td>Learning to exploit technology</td>
<td>DP6</td>
<td>0.79</td>
<td>L5 to use technology to advance collaborative work. L7 to use various digital applications and use them together whenever needed. L14 to understand the possibilities of digital technology better than before. L17 to develop products collaboratively by using technology.</td>
</tr>
</tbody>
</table>
intensive integration, feedback or technology use, but the students evaluated learning to collaborate, understand disciplines, and interdisciplinary collaboration more highly.

**Discussion**

Development of the questionnaire resulted from an identified lack of research instruments with which to measure discipline-general competence development, particularly in terms of collaboration, and measurements that could capture development related to the learning objectives of a certain course or teaching unit. We carried out explication and interpretation work to operationalize the design principles as corresponding items related to practices of knowledge work. We also noted that the learning practices examined may involve such a wide variety of detail in terms of activities that not all items are meaningful in each course context. Based on the statistical analysis, those aspects of reporting learning from collaborative practices that were perceived as meaningful across the many courses in the data were highlighted and included in the presented scales.

In the questionnaire, knowledge work competence was structured as object-bound collaboration, integration of personal and collective efforts, development through feedback, persistent development, understanding of different disciplines and related expertise, interdisciplinary collaboration, and using flexible tools and technology. First, a statistical analysis was carried out to establish measurement invariance for the questionnaire. Two groups of students were compared and a scalar model provided a good fit for the data, indicating strict factorial invariance (see Marsh et al., 2009), with acceptable correlations between factors. However, some of the items were actually related to several of the theory-based factors. These cross-loadings, especially on factors 1–2 and 4–5, are understandable, as the items examine various aspects of collaboration, i.e., planning and coordinating (factor 1), integrating individual and collective efforts (factor 2), giving and receiving feedback (factor 4), and iteratively developing the shared knowledge objects and products (factor 5). Factors 3 and 7 are both related to cross-fertilization of practices but are successful at distinguishing between two aspects: getting acquainted with professional practices and work-life experts (factor 7) and actually personally engaging in interdisciplinary collaboration (factor 3). Factor 6 has a distinct focus on exploiting technology during the course.

Group mean differences were reflective of the typical pedagogical settings of the courses as well as the discipline-specific expertise emphasized generally in the fields studied (cf. Jones, 2009; Steedle & Bradley, 2014). This speaks for the fact that self-assessment was not conducted at a general level but provided context-related information about learning. Overall, this met the original goal of the questionnaire development; that is, to provide an instrument that captures the connection between pedagogical practices and competence learning. More detailed examination of data and students’ feedback to developers has also yielded the observation that we should consider students’ prior work experience: some participants wrote that they did not award high scores to their own learning in the questionnaire because they were already competent in these practices as a result of their prior work experience. Future studies could address the relationship between age and prior work experience and pedagogical practices and self-assessed learning. This fits together with the competence-based framework whereby students may acquire certain competences at various points in their lives.

The relationship between the latent factors and constructed scales with the six design principles is not one to one (see Table 5). The development was initiated with the aim of presenting each design principle with a corresponding scale, but we realized that this was not a feasible goal. The design principles and CKP scales address two sides of the same coin. The theoretical design principles provide guidelines for course planning and organization for educational settings that intend to foster such educational practices and related learning outcomes. The CKP scales describe the types of generic competence that are potentially learned through participation in such social and cultural practices in educational settings. As described in the previous paragraph, within the practices, many of the activities are intertwined and perceived as overlapping. However, the analysis was able to establish strict factorial invariance for this model of seven factors, so it can be considered as a good point...
of departure to continue studying learning of collaborative knowledge practices and related competence development in various pedagogical contexts.

The validity of the instrument was examined based on a set of 32 courses, which involved considerable variation of educational practices in terms of pedagogical methods and learning activities in HE studies. The 27-item version of the instrument demonstrated acceptable psychometric properties. Based on a comparison of ESEM with structural equation modelling (SEM) and manifest regression analysis (MRA), ESEM is recommended when non-ignorable cross-factor loadings exist (Mai, Zhang, & Wen, 2018). In our study, ESEM produced good fit, although factor correlations were quite high. Also, several cross-factor loadings existed, which were examined item by item to provide a theory-based evaluation of the meaning of the cross-loadings. These may relate, first, to the challenge of explicating and differentiating the practices on a theoretical–empirical axis and, second, on the respondents’ interpretation of the questionnaire items and personal engagement in the practices. Hence, more data from various contexts is needed to examine their interplay and identify further revision needs for the instrument. However, in light of the aim of developing an instrument that provides contextual information about competence development in a certain educational setting, the results indicated that the instrument can be used for self-evaluation of competence development and comparison of pedagogical practices across different contexts.

**Limitations of the study**

In the Likert scale, we initially provided the option “not applicable to this course” (0) because we expected that this would enable the participants to express their opinion that such learning was not targeted and did not take place on their course. In subsequent analysis, we noticed the challenges of this formulation. The 0 score does not relate to the continuum of the rest of the scale. We ran the analysis as categorical data. Further, although certain items on a course received several 0 scores, other respondents provided moderately high scores for the items, so we could not draw clear conclusions about that specific course. We replaced the “not applicable” responses with “strongly disagree” in the scale in order to aid interpretation of the data. Using an alternative that was more extreme than “strongly disagree” would have biased the results. In the following data collection, we reformulated the Likert scale to “I learned during the course …” “not at all” (1) to “very much” (5).

From the data collection point of view, self-reported competence development is quite critically scrutinized in the current research debate (e.g., Stes, Min-Leliveld, Gijbels, & Van Petegem, 2010). Commonly, concerns have been raised about the construct validity of instruments and disregard for the context-relatedness of competence learning; these shortcomings were addressed by developing and examining the validity of the instrument and always situating the data collection within a course. Further, regarding the limitations of the use of the questionnaire, it is only meaningful in settings wherein there are some collaboration activities for students, e.g., during interactive lectures and in groups, teams, peer commenting, or project work; that is, where students have an opportunity to develop their collaborative knowledge work competence. In other words, the questionnaire items are not relevant within a course context in which no collaboration activities take place.

**Conclusions**

The first research question focused on establishing measurement invariance for the theory-based structure of the CKP questionnaire between media engineering and life sciences student subgroups. Based on the results we may conclude that the analysis provided evidence of strict measurement invariance. This permits future examination of course and group differences based on the scales. The second research question examined whether any group differences existed in the latent factor means. The results indicate that there were group differences between the disciplinary fields. These were in line with the scope of pedagogical aims and practices in the set of 32 courses examined. Further validation in different disciplinary, pedagogical, and cultural contexts is needed.
The question of what “competence” means remains a challenging issue and particularly the approach adopted here – linking the socio-cognitive and socio-cultural paradigms through the questionnaire – certainly requires further elaboration through research in various HE settings. Compared to the assessment instruments reviewed following the socio-cognitive paradigm, the responses to the CKP questionnaire are always related to learning of the course practices and not general self-evaluation of their own competence level or specification of a certain proficiency level based on tests. Compared to the socio-cultural paradigm, this study has not investigated the knowledge practices in educational settings as they evolved. Rather, it has examined the students’ perceptions of learning that has taken place within a given course. Further research is needed to compare teacher and student evaluations to relate pedagogical design and competence learning in combination with other types of criteria, including content learning, peer assessment, and authentic assessment (Stokking, van der Schaaf, Jaspers, & Erkens, 2004). Inclusion and examination of other materials, e.g., teacher interviews, descriptive documentation of the course syllabus, and expected learning outcomes is needed to address the question concerning the relationship between pedagogical practices and students’ learning about knowledge work competence more comprehensively (see, e.g., Holvikivi, Lakkala, & Muukkonen, 2016; Kymäläinen et al., 2018; Laakkonen & Muukkonen, 2019; Muukkonen, Lakkala, Toom, & Ilomäki, 2017). The validated questionnaire also included open-ended questions about students’ experiences of the course, but these were not investigated in the present study. In addition, a teachers’ version of the questionnaire has been developed, which asks teachers to describe the pedagogical design and practices of their courses and their evaluation of student learning using the same items. Open-ended questions and teachers’ descriptions could be included in the analyses in further studies to produce an even more comprehensive view of the usefulness of the developed instrument in various authentic educational settings.

We consider the contextuality of the CKP questionnaire as both a strength and a challenge of the instrument. Courses and educational practices have various objectives. It is neither feasible nor meaningful to take all of the competence dimensions reviewed or explicated here as a goal for any single course, but students’ competence learning ought to be developed through a series of different kinds of course and engagement. How the competence of a student, or a cohort of students, progresses could be examined through repeated use of the CKP questionnaire together with other generic measures and content learning assessment methods.

**Disclosure Statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was supported by the Academy of Finland [grant number 285806] and the European Commission [grant number 402765 543154-LLP-1-2013-1-FI-KA3-KA].

**ORCID**

Hanni Muukkonen http://orcid.org/0000-0002-1086-1352  
Minna Lakkala http://orcid.org/0000-0002-8289-4292  
Pekka Lahti-Nuuttila http://orcid.org/0000-0001-5463-1738  
Liisa Ilomäki http://orcid.org/0000-0002-5429-7853  
Klas Karlgren http://orcid.org/0000-0003-1255-7650  
Auli Toom http://orcid.org/0000-0002-3261-3376

**References**


