frontiers | Research Topics

Generic skills in higher education

Edited by
Heidi Hyytinen, Tarja Tuononen and Edith Braun

Published in
Frontiers in Education

ISSN 1664-8714
ISBN 978-2-8325-2215-8
DOI 10.3389/978-2-8325-2215-8

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Generic skills in higher education

Topic editors

Heidi Hyytinen — University of Helsinki Centre for University Teaching and Learning, Finland

Tarja Tuononen — University of Helsinki, Finland

Edith Braun — Justus-Liebig Universität, Germany

Citation

Hyytinen, H., Tuononen, T., Braun, E., eds. (2023). *Generic skills in higher education*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2215-8

Table of contents

frontiers | Frontiers in Education

TYPE Editorial
PUBLISHED 04 April 2023
DOI 10.3389/feduc.2023.1162156

OPEN ACCESS

EDITED AND REVIEWED BY
Maria J. Hernandez-Serrano,
University of Salamanca, Spain

*CORRESPONDENCE
Heidi Hyytinen
✉ heidi.m.hyytinen@helsinki.fi

SPECIALTY SECTION
This article was submitted to Higher Education, a section of the journal Frontiers in Education

RECEIVED 09 February 2023
ACCEPTED 28 March 2023
PUBLISHED 04 April 2023

CITATION
Hyytinen H, Tuononen T and Braun E (2023) Editorial: Generic skills in higher education. *Front. Educ.* 8:1162156. doi: 10.3389/feduc.2023.1162156

Editorial: Generic skills in higher education

Heidi Hyytinen[1]*, Tarja Tuononen[1] and Edith Braun[2]

[1]Centre for University Teaching and Learning (HYPE), Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland, [2]Institute for Educational Science, Professorship for Teaching and Learning in Higher Education, Justus-Liebig-University Giessen, Giessen, Germany

KEYWORDS

generic skill, higher education, generic skill development, assessment, teaching-learning

Editorial on the Research Topic
Generic skills in higher education

Over the past decade, the importance of generic skills, such as collaboration, critical thinking, problem solving, and communication skills has gained increased interest in educational policy discourses and in practice of higher education. Previous research has shown that generic skills are related higher education students' learning processes, academic achievement as well as learning of disciplinary knowledge and skills (Arum and Roksa, 2011; Tuononen et al., 2017; Hyytinen et al., 2021). Students need generic skills to construct and apply of their domain-specific knowledge and understanding (Hyytinen et al., 2019). To take an example, communications skills, enable students to make their ideas and conclusions visible to others (Braun, 2021; Kleemola et al., 2022). Together with domain-specific knowledge, generic skills are also vital in working life (Tuononen and Hyytinen, 2022; Iqbal et al., 2023). In transition to working life, generic skills are found to be related to development of expertise (Tuononen et al., 2017).

Although there is a growing consensus on the importance of generic skills, there is evidence many higher education students face challenges in generic skills. There is also a large variation in students' level of generic skills. Students' level of generic skills are associated with prior academic performance and by their socioeconomic background (Arum and Roksa, 2011; Kleemola et al., 2022). Surprisingly little is known to what extent and how these skills develop during higher education studies. Furthermore, there are no consensus the processes of implementing generic skills in teaching and learning in programmes (Hyytinen et al., 2019). Additionally, there is no unanimous agreement about what is meant by the concept of generic skills in the research community (El Soufi and See, 2019). Therefore, there is a wide variety of definitions and skills that researchers consider generic skills.

This Research Topic with twelve articles from different countries and research projects advocates better understanding about learning and teaching generic skills at the context of higher education. The aim of this Research Topic is to bring together research papers covering different aspects of generic skills, including theoretical or empirical papers that explore and outline the development of generic skills in various fields of education, empirical papers that investigate the learning and teaching environments that support the learning generic skills, papers that focus on the assessment of generic skills, and systematic review related to the topic of the Research Topic. Thus, the rationale for this Research Topic is to strengthen the current state of international research on generic skills with a view to the fundamentals established so far and to bring new insights into research on generic skills. This Research Topic also contributes to discussions about the importance of generic skills in the higher education.

In this special issue, theoretical, conceptual and methodological aspects and topics are addressed, including issues related to the validation and adaptation of a performance-based assessment of generic skills (see paper by Ursin et al., Nagel et al.) and value of assessing tertiary students' ability to reason relationally (Alexander et al.). Furthermore, this Research Topic includes a review study that examines theoretical, methodological, and empirical viewpoints on learning generic skills and synthesizes the current empirical evidence about the factors that enhance and impede student learning of generic skills (see Tuononen et al.). The majority of the articles in this Research Topic focus on learning and development of generic skills as well as what kind of role of teaching and learning environment and different learning experiences play in learning generic skills (see papers by Nielsen et al., Räisänen et al., Virtanen et al., Lee and Lee, Muukkonen et al., Slišāne et al.). The Research Topic also offers insights into the variation and challenges is generic skills (Kleemola et al.) and the fit between learning opportunities for generic skills available at universities and skills required in working life (Lohberger and Braun). The studies use various methods from a small-scale qualitative analysis of think-aloud data to the quantitative analysis of follow-up data.

Taken together, this issue of generic skills will serve a reference for articulating future directions in research and practice in the context of higher education. The papers included in this Research Topic also highlight new perspectives for the future research. For example, intervention and longitudinal studies focusing on the development of generic skills are needed. Moreover, in order to become more coherent research field new valid research instruments to measure generic skills are required.

Author contributions

HH: draft paper preparation. TT, EB, and HH: comments and revisions. All authors approved the final version of the editorial.

Acknowledgments

We would like to thank all authors, reviewers, journal editors, and colleagues involved in preparing this special issue.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Arum, R., and Roksa, J. (2011). *Academically Adrift: Limited Learning on College Campuses*. Chicago: University of Chicago Press. doi: 10.7208/chicago/9780226028576.001.0001

Braun, E. (2021). Performance-based assessment of students' communication skills. *Int. J. Chinese Educ.* 10, 1–12. doi: 10.1177/22125868211006202

El Soufi, N., and See, B. H. (2019). Does explicit teaching of critical thinking improve critical thinking skills of English language learners in higher education? A critical review of causal evidence. *Stud. Educ. Eval.* 60, 140–162. doi: 10.1016/j.stueduc.2018.12.006

Hyytinen, H., Toom, A., and Shavelson, R. J. (2019). "Enhancing scientific thinking through the development of critical thinking in higher education," in *Redefining Scientific Thinking for Higher Education: Higher-Order Thinking, Evidence-Based Reasoning and Research Skills*, eds. M. Murtonen, and K. Balloo (Palgrave Macmillan) 59–78. doi: 10.1007/978-3-030-24215-2_3

Hyytinen, H., Ursin, J., Silvennoinen, K., Kleemola, K., and Toom, A. (2021). The Dynamic relationship between response processes and self-regulation in critical thinking assessments. *Stud. Educ. Evaluat.* 71, 101090. doi: 10.1016/j.stueduc.2021.101090

Iqbal, J., Shaikh, A. A., Jamal, W. N., Akhtar,., K., Rahim, R., et al. (2023). Exploring the generic skills required for the employability and professional wellbeing of Pakistani Millennials: The employers' perspective. *Front. Psychol.* 5, 1070267. doi: 10.3389/fpsyg.2022.1070267

Kleemola, K., Hyytinen, H., and Toom, A. (2022). Critical thinking and writing in transition to higher education in Finland: do prior academic performance and socioeconomic background matter? *Eur. J. Higher Educ.* 2022, 1–21. doi: 10.1080/21568235.2022.2075417

Tuononen, T., and Hyytinen, H. (2022). Towards a successful transition to work - which employability factors contribute to early career success? *J. Educ. Work.* 35, 599–613. doi: 10.1080/13639080.2022.2126969

Tuononen, T., Parpala, A., and Lindblom-Ylänne, S. (2017). "The transition from university to working life. An exploration of graduates' perceptions of their academic competences," in *Higher Education Transitions: Theory and Research*, eds. E. Kyndt, V. Donche, K. Trigwell, and S. Lindblom-Ylänne (London: Routledge - Taylor and Francis Group) 238–253. doi: 10.4324/9781315617367-18

frontiers | Frontiers in Education

ORIGINAL RESEARCH
published: 09 May 2022
doi: 10.3389/feduc.2022.885987

The Challenge of Position-Taking in Novice Higher Education Students' Argumentative Writing

Katri Kleemola, Heidi Hyytinen and Auli Toom*

Centre for University Teaching and Learning, University of Helsinki, Helsinki, Finland

Argumentative writing is the central generic skill in higher education studies. However, students have difficulties in basic argumentation skills. Novice students do not necessarily receive adequate guidance, and their prior education may not have supported the requirements of higher education writing. Position-taking is at the core of argumentation, but students are often hesitant to make their point. Furthermore, they may have an incorrect and one-sided perception about an argument, leading them to avoid alternative positions in their argumentative writing. The study aims to explore starting level skills of novice students' argumentative writing, namely their position-taking. The participants were 196 first-year students from diverse fields of study in two Finnish higher education institutions. They were required to solve a problem and write an argumentative essay based on five documents that were given to them. The essays were analyzed using qualitative content analysis applying abductive approach. Substantial variation was detected in students' position-taking. We identified four groups of writers based on their position-taking. First two groups were more or less explicit in their position-taking. Most of the students (72%) belonged to these two groups. However, a minority of them were consistent in their position-taking. Writers in the third group (15%) implied their position, and writers in the fourth group (12%) stuck to summarizing sources without position-taking. The findings invite teachers to support novice students in their basic argumentation. Co-operation between faculty teachers and writing teachers is encouraged.

Keywords: argumentative writing, higher education, novice students, generic skills, position

OPEN ACCESS

Edited by:
Rafael Denadai, Plastic and Cleft-Craniofacial Surgery, A&D DermePlastique, Brazil

Reviewed by:
Jeffery Nokes, Brigham Young University, United States
Setyo Admoko, Surabaya State University, Indonesia

****Correspondence:***
Katri Kleemola
katri.kleemola@helsinki.fi

Specialty section:
This article was submitted to Higher Education, a section of the journal Frontiers in Education

Received: *28 February 2022*
Accepted: *21 April 2022*
Published: *09 May 2022*

Citation:
Kleemola K, Hyytinen H and Toom A (2022) The Challenge of Position-Taking in Novice Higher Education Students' Argumentative Writing. Front. Educ. 7:885987. doi: 10.3389/feduc.2022.885987

INTRODUCTION

Generic skills have been considered vital for success in higher education studies (Barrie, 2006; Shavelson, 2010; Hyytinen et al., 2019). They are universal expert skills, such as communication, problem solving and argumentation, and they are equally important in all fields, enabling learning discipline-specific skills and knowledge (Hyytinen et al., 2021a). The central generic skill is argumentation (Andrews, 2009; Mäntynen, 2009; Wolfe, 2011; Wingate, 2012). Argumentation, and more specifically argumentative writing, is required of the students from the moment they apply and enter a higher education institution, until graduation, in the form of essays, examinations, and dissertations (see Wolfe, 2011; Wingate, 2012). Even more important, it is not just a technical skill, to pull through assignments, but argumentation also facilitates learning (Asterhan and Schwarz, 2016; Iordanou et al., 2019; Kuhn, 2019). Research on generic skills often focuses on clusters of skills,

their importance, and students' experiences of them (e.g., Barrie, 2006; Tuononen et al., 2019; Virtanen and Tynjälä, 2019). However, such an approach offers few practical insights for higher education teachers who often struggle between teaching discipline-specific knowledge and supporting students in their generic skills. Instead, gaining a more detailed understanding of students' strengths and weaknesses in each generic skill, such as argumentation, will help in developing tools for teachers.

Several studies show that even advanced higher education students have gaps in their basic argumentation skills, such as combining claims and evidence, or presenting diverse viewpoints (Marttunen, 1994; Ivanič, 1998; Andrews et al., 2006; Laakso et al., 2016; Hyytinen et al., 2017, 2021b; Breivik, 2020). Students may be unsure about what an argument is (Andrews, 2009; Wingate, 2012; Breivik, 2020). They may also have difficulties in identifying rhetoric situations and their expectations and adapting their writing for the requirements of each assignment (Zimmerman and Risemberg, 1997; Johns, 2008; Roderick, 2019). It has been suggested that prior education does not provide sufficient argumentative skills, but students in higher education still feel that they do not receive adequate guidance or instructions on elements of argumentative writing (Andrews, 2009). Teachers often assume that students either already master these skills or learn as they go. Surprisingly, even though argumentative writing has been thought to be the Achilles' heel in the transition to higher education, little research has focused on actual novice students' starting level skills. We know a lot more about advanced students' or even senior scholars' argumentative skills. The present study focuses on novice students' basic skills in argumentative writing, namely describing the variation in the ways of their position-taking, which is viewed as the core of argumentation (Andrews et al., 2006; Wingate, 2012).

Argumentation and Argumentative Writing in Higher Education

The objective of an argument is to support one's claims and conclusions with reasons or evidence (Toulmin, 2003; Halpern, 2014). In academic contexts, the claims and conclusions are backed with prior research and/or empirical data (Swales, 1990; Wolfe, 2011). In argumentative guidebooks, an argument is often presented as a simple one or two sentence structure, but in practice it is often integrated in broader entities such as written essays or articles, or spoken addresses or debates (see Andrews, 2009). Most assignments that a higher education student—across disciplines—encounters during their studies require argumentative writing (Wolfe, 2011). Assignments that require argumentation have also been considered a valuable tool for learning. Such assignments have been found to be a particularly advantageous method when learning about complex topics with diverse viewpoints and complex skills such as critical thinking (Asterhan and Schwarz, 2016; Iordanou et al., 2019; Kuhn, 2019).

There is no template for constructing an argumentative text, but the writer must identify the requirements of the situation, and the best ways to fulfill those requirements (see Johns, 2008). A major decision in argumentative writing is related to choosing the placing of claims or conclusions and evidence. These rhetorical strategies are culture-specific to some degree; in other words, one strategy may be favored over another, across genres and communities. For instance, Finnish writers have been found to prefer to present all evidence and elements of uncertainty before their conclusion (final focus), in contrast to Anglo-American writers who prefer to present their inference first and then proceed to evidence (initial focus) (Mauranen, 1993; Mikkonen, 2010; see also Perelman and Olbrechts-Tyteca, 1969).

In addition to variation across cultures, argumentative skills have also been suggested to be, at least in part, discipline-specific (Andrews, 2009, 2015). Accordingly, there are disciplinary differences in the epistemologies that influence how to evaluate an argument (e.g., Hetmanek et al., 2018). However, beyond the varying conventions of cultures and disciplines, arguments and argumentative texts have more generic features. This includes development and presentation of one's position (Andrews, 2009; Wingate, 2012), and micro- and macrostructures of argumentation, such as claim or conclusion and evidence, and introduction, counterarguments, and discussion (Kuhn, 1991; Toulmin, 2003; Breivik, 2020). To learn the discipline-specific conventions of argumentation, it is necessary to master the generic features. Consequently, the ability to use generic features of argumentation is eminently important for novice students who are new to higher education. They are not yet integrated in their study program or academic writing community (Swales, 1990; Donald, 2002). However, despite the lack of relevant skills, novice students receive little guidance in argumentative writing. In the absence of proper guidance to academic requirements, they are tapping into the skills they have learnt in their prior education (Bereiter and Scardamalia, 1987; Andrews et al., 2006). In Finland, it has been suggested that argumentation is not sufficiently emphasized in the upper secondary school, and its final exams, the Matriculation Examination (Mäntynen, 2009; Komppa, 2012). However, evidence-based information about Finnish novice higher education students' argumentative writing is scarce. While we know that they have some problems in consistency of their arguments (Hyytinen et al., 2017), there is no research on more generic features in argumentative writing.

Position-Taking in Argumentative Writing

Taking a position is at the core of argumentation (Andrews, 2009; Wingate, 2012). Typically, the position is seen as the viewpoint the writer intends to support, or the main point the writer intends to make. The position conveys the writer's explicit presence in the text (Mauranen, 1993; Hyland, 2005). Additionally, to strengthen the argument, the position can be challenged with alternative positions (see Andrews, 2009). Failing to take a position can lead to problems in higher education studies where argumentation skills are vital (e.g., Wolfe, 2011). Such problems often go hand in hand with problems in deep learning and meaning construction (Biggs, 1988; see also Petrić, 2007).

In argumentative writing, the position is often expressed as a thesis, a holistic main claim that summarizes the writer's point of view (Kakkuri-Knuuttila and Halonen, 1998; Mikkonen, 2010; Wolfe, 2011). However, the position is not always expressed

as explicitly as a thesis. Indeed, it has been found that higher education students have challenges in emphasizing their position, and instead, they may lean toward research sources, as well as summarizing, and attributing (Lea and Street, 1998; Petrić, 2007; Mäntynen, 2009; McCulloch, 2012; Laakso et al., 2016; Lee et al., 2018). Consequently, they do not take a position, but they rather display their knowledge on the topic (see Petrić, 2007). Higher education students can feel inadequate for making a strong point (Ivanič, 1998; Andrews, 2009; Mendoza et al., 2022). They may avoid making a holistic statement like a thesis by making so called local arguments. These are claims that encompass a short proportion of the text, and do not summarize the point of view of the entire text (Mauranen, 1993; Wolfe, 2011). However, writers may also imply their position in more subtle ways than stating an explicit thesis or even making local arguments. Linguists talk about interactional features, referring to elements that convey writer's relation with their text (Hyland and Tse, 2004; Hyland, 2005). Writers might withhold (hedge) or emphasize (boost) their commitment, express their affective attitudes (attitude marker), or use first-person forms to remind reader of their presence in the text (self-mention) (Hyland, 2005).

Discussion of diverse viewpoints, i.e., alternative positions, is an important yet challenging part of argumentation (Kuhn, 1991; Andrews, 2009; Wingate, 2012; Kuhn et al., 2016b). In its strongest form, a rebuttal, an explicit position is taken against some evidence. Just as they may be hesitant in their position-taking, as discussed above, even advanced higher education students may have challenges in introducing alternative positions in their argumentation (Laakso et al., 2016; Hyytinen et al., 2021b; Kuhn and Modrek, 2021). Even acknowledgment of alternative positions is difficult for many, not to mention rebutting them (Kuhn, 1991). This tendency has been called my-side bias, indicating an inability to see other alternatives (Perkins, 1989). However, these challenges may not be about an inclination to emphasize one's own opinion but instead they reflect the writer's incorrect perception of an argument (Wolfe and Britt, 2008; Wingate, 2012). Writers may see a good argument as a one-sided construction, and so they bring out all the supporting evidence, and leave out any contesting facts. The ability to develop rebuttals requires a basic understanding of position-taking, and usually, the presence of rebuttals is an indication of a higher overall quality of argumentation (Wolfe et al., 2009; Kuhn et al., 2016a).

The variation in position-taking can be a consequence of either hesitancy or uncertainty (Ivanič, 1998; Andrews, 2009; Lee et al., 2018) or an incorrect perception about position-taking and the characteristics of an argument (Wolfe and Britt, 2008; Andrews, 2009; Breivik, 2020). However, research has pointed out other reasons students may avoid position-taking. Cultural conventions, such as inclination to minimize writer presence in a text, can add to the challenge (Mauranen, 1993). Furthermore, the tendency to reward students for showing what they know instead of constructing new meanings discourages writers from developing their own positions (Bereiter and Scardamalia, 1987; Andrews, 2009). Overall, some writers intend to present all available evidence, in contrast to writers who intend to construct new information based on available evidence (Bereiter and Scardamalia, 1987; Biggs, 1988). Some interesting contrasting findings suggest that assignment directions could be a culprit behind the challenges. The writer may not correctly identify the requirements of an assignment. It has been found that even when explicit requests are made for position-taking, writers may fall back on a strategy of summarizing sources (Macbeth, 2006; Andrews, 2009; Paldanus, 2017). However, when the directions come across, higher education students and even younger adolescents are able to present an explicit position (Marttunen, 1994; Marttunen and Laurinen, 2004; Mikkonen, 2010).

The Research Gap and Objective of the Study

Argumentative writing is demanding, and novice students cannot be expected to master the conventions and hidden rules of the academic context (Swales, 1990; Ivanič, 1998; Macbeth, 2006; Johns, 2008; Andrews, 2009). To understand where novice students stand in their argumentative writing skills, all dimensions of argumentation need to be studied; implementing logic, rhetoric, and dialectic approaches. A tendency in the research of argumentation in the higher education studies is to focus on the validity and quality of arguments (cf. Andrews, 2009; Wingate, 2012). However, if students do not understand what an argument is or are not able to formulate their arguments on a textual level, teaching more abstract and often discipline-specific aspects of argumentation, such as validity, may be futile. Furthermore, novice students' preparedness for argumentative writing in the academic context has been questioned, but research-based understanding of the matter is insufficient. Thus, the present study investigates novice students' argumentative writing on a very basic level, namely focusing on their position-taking skills. Such investigation will bring insights into supporting novice students in their studies. The growing understanding of novice students' starting level skills will help not only writing teachers, but all teachers in higher education who are involved in giving writing assignments.

The aim of our study was to explore starting level skills of novice higher education students' argumentative writing. In more detail, we investigate how they take a position in an argumentative essay, and how they present alternative positions. Our specific research questions were:

RQ1: What kind of variation is there in students' presentations of their position?

RQ2: What kind of variation is there in students' presentations of alternative positions?

RQ3: What types of argumentative writers can be identified based on the findings in RQ1 and RQ2?

MATERIALS AND METHODS

Context of the Study

Higher education admissions in Finland are extremely competitive (OECD, 2019). Until recently, all applicants have participated in discipline-specific entrance examinations, but recently more emphasis was put on academic achievement

in prior education (Kleemola and Hyytinen, 2019). The upper secondary school in Finland consists of a general and a vocational track. The aim of the general track is to give students extensive general knowledge and to prepare them for further education either in higher education or in vocational training. While the general track introduces all subjects to all students, different emphases are allowed, such as mathematics, natural sciences, or languages. In contrast, the vocational track aims to give students vocational competence in their chosen field. Some general subjects are taught, but the focus is on vocational skills. It is noteworthy that both general and vocational tracks give eligibility for higher education admissions. Thus, novice students in higher education have varied academic backgrounds.

Participants and Data Collection

The data were collected in accordance with the ethical principles of research with human participants by the Finnish National Board on Research Integrity (2019). Students gave their consent to participation. The data were collected in a national project on higher education students' generic skills (Ursin et al., 2021). Based on instructions and templates provided by the project, administrators, and teachers in participating institutions invited students via e-mail. Approximately 25% of the invited students participated. In the national project, 1538 first-year students participated in the early stages of their first study term. Participation was voluntary, and students could withdraw from the research at any time. Sixty-nine students did not complete the assessment, and they were not included in any analyses. For the purposes of the present study, a subsample of the data was selected. The aim was to sample a subgroup that would represent a wide range of disciplines, and the variation in argumentative writing within the whole data. With these aims in mind, two large multidisciplinary higher education institutions in southern Finland were selected to represent the data. The selection of students in two institutions instead of a random sample across the 18 participating institutions ensured that there was not too much contextual variation in the students of the subgroup. One of the institutions was a research-intensive university (99 students) and one was a university for applied sciences (97 students). The subgroup of 196 Finnish-speaking students covered the whole range of performance levels in the national project, assessing generic skills. Thus, it was assumed that the subsample would represent the variation in argumentative writing within the whole data. Students represented a diversity of study fields, including healthcare, humanities, biosciences, engineering, natural sciences, and social sciences, covering most of the key disciplines. While it is possible that study fields attract students with different skillsets, the present study does not report disciplinary differences, as participating novice students were not yet exposed to different disciplinary cultures. Students' mean age was 24.83 (SD = 6.85) and median age 22.00. While this is slightly older than the average starting age, it is worthy to note that Finns typically start in higher education older than in other countries (OECD, 2019). The majority of the students (84%) had completed the general track in the upper secondary school.

Participants completed a computer-based Collegiate Learning Assessment (CLA +) that includes an open-ended performance task and a multiple-choice section. The open-ended task, used in the present study, required students to think critically and argue for their response in writing (see Klein et al., 2007; Kleemola et al., 2021). The task at hand was designed to activate argumentative writing skills in participants. Furthermore, it was suitable for assessing argumentative writing in the academic context: participants were required to develop a position by means of leaning on documents that were provided in the task. Thus, the task simulated the reality of academic contexts in that the text should be embedded in the existing research literature (see e.g., Swales, 1990). Performance tasks in general have been found to be motivating for students (Kane et al., 2005; Hyytinen and Toom, 2019; Hyytinen et al., 2021c). The task that was used in the study is confidential, but similar tasks are introduced by Shavelson (2010) and Hyytinen and Toom (2019). In the task, students were asked to take a role of an intern in a city government, where they would have to solve a problem and give a report concerning different life expectancies in two cities, Woodby and Brookdale. They were provided five documents, namely a blog post, a podcast transcript, a memorandum, a newspaper article, and an infographic. They were asked to give their response to the problem as an essay, and their proposals for action. They were reminded to discuss any counterarguments, and to support their own claims with information in the available documents. Students had 60 min to complete the task. The length of the responses ranged from one sentence of about 10 words to several pages of about 800 words.

Data Analysis

Qualitative content analysis was used to investigate students' texts. An abductive approach (e.g., Timmermans and Tavory, 2012) was adopted, namely the analysis was theory-driven to start with, but researchers kept an open mind to new discoveries based on the data. Both a group-level approach (RQ1 and RQ2) and an individual-level approach (RQ3) were used to gain a multifaceted view on the topic. The investigation proceeded in six main phases that are presented in **Table 1**. The process was non-linear in that the authors discussed findings and issues continuously during the process.

In the first phase of the analysis, all three authors became familiar with the data by reading students' responses. During this phase, general observations were made about the data and were discussed. In the second phase of the analysis, theory-based analysis criteria for identifying the position and alternative positions were created by the first author. The criteria are described in detail below. In phase three, the analysis criteria were implemented by the first author. In addition, data-based, novel features were noted. The episodes with positions and alternative positions were located and their variation was explored and reflected in light of existing research. In phase four, the second author analyzed independently 25% of the essays that were randomly selected. The first and second authors discussed the findings, negotiated their differences, and adjusted analysis criteria where needed and integrated the data-based findings in the criteria. In phase five, the first author re-analyzed the data according to the adjusted criteria from phase four. Finally, in phase six, all authors discussed the findings, and remaining

TABLE 1 | The main phases of the data analysis.

	Aim	Actions	Investigator
Phase 1	To become familiar with the data	The data were read and discussed	All authors
Phase 2	To prepare for the analysis	Analysis criteria were created based on theory and observations of the data	First author
Phase 3	To test and apply analysis criteria	The data were analyzed according to analysis criteria, variation, novel features, and problems were spotted	First author
Phase 4	To adjust analysis criteria and gain new perspectives	The data were analyzed by another investigator, findings and differences were discussed and negotiated; criteria were adjusted	First and second authors
Phase 5	To apply adjusted criteria	The data were re-analyzed	First author
Phase 6	To synthesize findings	The findings were discussed, final adjustments were made, and writer groups were created	All authors

analysis challenges. Additionally, groups of argumentative writers were identified on the basis of the findings.

The analysis criteria were created and adjusted during phases two and four. The criteria aimed for describing the variation in the students' texts instead of creating exclusive categories. To respond to RQ1, writers' positions were analyzed, and two types of episodes were traced, namely explicit thesis-statements and other position-indicating expressions (see also **Table 2** in "Results" section). The thesis was defined as the response for the main question (Mauranen, 1993). In academic texts, the question is the research question or research aim. In the present study, the task question "what is the reason for the different life expectancies in the two cities?" was considered to be the relevant question. Furthermore, the thesis should be a holistic claim, summarizing the main point of the whole text (Mauranen, 1993). During the analysis in phase three, it was found that some of the essays included a thesis-like statement that responded to a more generic question than the actual task question, having a different orientation (see more in "Results" section). In phase four it was determined that these were included as a thesis, but *thesis orientation* was to be examined. According to theories and prior studies, the *thesis location* could be in the introductory section of the text (initial focus), in the conclusion section of the text (final focus) or both (Nestorian order) (Perelman and Olbrechts-Tyteca, 1969; Mauranen, 1993). Thus, the thesis was traced in these sections. Expressions such as "in conclusion" were first looked for, but as Finnish writing tends to include less such metatext (Mauranen, 1993), the content of the sentences was examined more carefully. Consequently, typical expressions were "is caused by," "significant factors are," and "the reason behind—is." A closer analysis of the variation in the thesis-statements revealed that the *thesis precision* varied, and thus, it was integrated in the analysis criteria in phase four and examined. It was found that some thesis-statements were vague, and some included hedges that are interactional elements implying uncertainty (Hyland, 2005). In philosophical theories, such expressions have been considered to be modal qualifiers (Toulmin, 2003) that assess the degree of probability of the statement, namely how likely they think the statement to be true.

If a thesis could not be found, other position-indicating expressions were traced. *Interactional elements* (Hyland, 2005) were traced, namely attitude markers and self-mentions that appear without a thesis. Affective, and attitudinal expressions, such as "positive," "unreliable," "questionable," and "cannot be trusted" were traced. In addition, the first-person forms of pronouns "I" and verbs "I recommend" (Finnish verbs have an integrated first-person form which does not require the pronoun) were traced. During the analysis, it was found that in some essays there were explicit claims that could nevertheless not be identified as a holistic thesis. These were recognized as *local arguments* (Mauranen, 1993; Wolfe, 2011), that indicate a position while not expressing it holistically, in the way that the present study defined the thesis. Local arguments were integrated in the analysis criteria in phase four. Finally, in some essays we found *no position-taking*, neither explicit thesis nor other position-indicating expressions.

In analyzing alternative positions (RQ2), we traced episodes where the writer showed their awareness on rival explanations to the task question, namely "what else could be the reason behind the different life expectancies, but maybe is not true?" (see also **Table 3** in "Results" section). Such episodes show that the writer accepts the possibility of diversity in the positions, while they do not necessarily always express an explicit position against the evidence (Kuhn, 1991). Thus, we traced both *rebuttals* and expressions of *contradictory positions*. To trace these episodes, we looked for contrasting expressions such as "however," and "on the other hand." Furthermore, we looked for rebuttals, such as "is not significant," "does not affect," and "I can't agree." Additionally, as mentioned above in thesis-identification, content of the sentences was examined beyond metatext. Structures such as "X says Y, Z says not Y" were noted. During the analysis, variation was detected in the rebuttals, namely in the reasoning, hesitancy, and attitudes. These features were integrated in the analysis criteria in phase four.

To respond to RQ3, findings of RQ1 and RQ2 were examined carefully. At this point, the variation in the degree of writer's presence through their position-taking was recognized as the guiding theme. Four groups were created examining the aspects of position-taking, their differences, and similarities. The differences between the groups culminated in explicitness of their position. The process of creating the groups is also presented in the "Results" section (**Figure 1**).

Investigator triangulation was used in the data analysis (Denzin, 1970). Becoming familiar with the data in the first phase allowed all authors to evaluate findings in light of the entire dataset. While the first author performed the analyses in phases three and five, all authors discussed and evaluated findings and considered challenging features throughout the process. Additionally, in phase four, the independent analysis by the second author and discussions that followed ensured that all authors interpreted the analysis criteria similarly.

The excerpts that are presented in the "Results" section are translations from Finnish to English. Any typos were omitted in translations as they were not relevant to the present study. Some details concerning the content have been altered in order to preserve confidentiality of the task. These alterations do not influence the analysis of the position-taking that is the focus of the present study.

RESULTS

Variation in the Presentations of Writer's Position

Position-indicating episodes were identified in the essays and variation was analyzed. The sources of variation are summarized in **Table 2**. An explicit thesis was detected in 125 (64%) essays. The thesis-statements varied in their orientation, location, and precision. An explicit thesis could not be detected in 71 essays (36%). However, in some of these essays, a degree of writer presence and position-taking could be detected, namely interactional elements and local arguments.

Thesis Orientation

We first examined the orientation of the explicit thesis-statements. A thesis that adhered to the original definition, responding to the task question (explanation behind different life expectancies in two cities) was found in 88 essays (45%). A typical thesis is presented in example 1. A typical thesis that was oriented toward the task question explicates that the statement is about differences in the life expectancy in the two cities or about Woodby's higher life expectancy compared with that of Brookdale. Furthermore, a typical thesis lists the most important factor or factors behind the life expectancy. The thesis is underlined in the examples.

TABLE 2 | Variation in presentation of writer's position.

Thesis explicitness	Aspect of position-taking	Source of variation	*F*
Explicit thesis			125
	Thesis orientation	Thesis that responds to the task question.	88
		Thesis-like statement that responds to a generic question	38
	Thesis location	Initial focus thesis	62
		Final focus thesis	61
		Nestorian order	2
	Thesis precision	Uncertainty	12
		Vague thesis	13
No explicit thesis			70
	Position through interactional elements	Critique	4
		Commentary and reflection	23
		1st person expressions	22
	Position through local arguments	Thematic organization	26
	No writer presence		29

(1) — *Factors that influence Woodby's higher life expectancy are exercise and the level of education.* —

However, during the data analysis, it was discovered that some essays had a slightly different orientation. They included a thesis-like statement, making a holistic, summarizing statement based on the evidence, which did not respond to the task question. These thesis-statements (F = 38, 19%) responded to a more generic question of factors that influence life expectancy instead of addressing the differences between the two cities. Writers of these essays would write about the factors that help individuals live longer (example 2), or that shorten the life expectancy.

(2) — *In summary, it could be said that the level of education, sleep, exercise, and nutrition are keys to a long life.*

These students may have falsely understood the question, or they may have failed to analyze the task materials, as responding to the actual task question required deeper problem-solving across available documents. Furthermore, they may have assumed that this generic question was in fact, what was asked, due to most of the materials addressing it. In further analyses in the present study, both thesis orientations were treated as an explicit thesis.

Thesis Location

Next, we examined the variation in the thesis location. Half of the essays with an explicit thesis (F = 62) were using initial focus strategy, namely the thesis was situated in the introductory section of the essay. These thesis-statements were located either in the very beginning of the essay or after an orientating introduction. Some writers would open their essay by stating the reasons behind different life expectancies (example 3) and proceed by presenting the detailed evidence behind this statement. In contrast, some essays with the initial focus strategy opened with an introductory section where writers would describe general background information about the situation (example 4) or state their objectives of the text. After the introduction, they proceed to their thesis-statement, and follow with the detailed evidence.

(3) *It seems that two factors are above the rest behind the higher life expectancy of Woodby: exercise and the level of education.* —

(4) *In the region of Brookdale-Woodby, the changes in life expectancy and reasons behind it have been followed for a long time. At the moment, there is a lot of discussion about actual reasons why people in the Woodby region live longer than the average. Based on research findings, we also can give Brookdale residents tips on how to increase their life expectancy.*

There are many reasons for the long life expectancy in Woodby region. According to research findings, the two most significant differences between residents of Woodby and Brookdale are exercise and the level of education. —

The other half of the essays with an explicit thesis (F = 61) had a final focus strategy, namely the thesis was situated toward the end of the essay, as a conclusion (examples 5 and 6). Such a thesis was mostly preceded with the presentation of evidence that supports the thesis. In some essays, an alternative position

was presented after the supporting evidence, but before the thesis. In such cases, writers first listed the evidence that they thought was relevant to their own position and then listed the evidence they rebutted or considered inconsequential to the task question (see in more detail below in section "Variation in Presentations of Alternative Positions"). The thesis-statements with a final focus strategy may have been at the very end of the essay (example 5). However, most often it was followed by the proposals for action (example 6), namely what should the cities and their citizens do about the difference in life expectancies.

(5) — *It can be concluded that especially the large amounts of exercise and the high level of education seem to lengthen the life expectancy and are probably contributing to the differences in life expectancy in Woodby and Brookdale.*

(6) — *In other words, since Woodby residents are getting more exercise and have a higher level of education, they also live longer.*

Therefore, I recommend Brookdale residents to live healthier. —

In two essays, a Nestorian order was used, namely the thesis was repeated in the beginning and in the end (example 7). While the thesis was reworded, it was similar in contents in the beginning and in the end. Thesis-statements in the Nestorian order essays were not different from the above descriptions of initial and final focus strategies. However, with only two occurrences, inferences about thesis-statements in Nestorian order need to be taken with caution. Some of the essay text has been omitted from the example to save space.

(7) *Woodby has a higher life expectancy compared to its neighboring town of Brookdale. The strongest contributing factors seem to be the residents' level of education and the amount of exercise they are getting. In the Woodby that has the higher life expectancy, a larger proportion of residents —*

[supporting evidence]

[rebuttal of alternative positions]

— Neither seems the sleep they are getting to be significant contributor to the difference in the life expectancy: volume of sleeps seems to be equal in both towns.

Based on the materials, it seems that the longevity of Woodby residents has to do with healthy exercise and the level of education. Therefore, I recommend the Brookdale residents to focus on their exercise in accordance with instructions by the personal trainer Maria. In addition, the educational level of Brookdale residents should be raised.

Thesis Precision

The final aspect of the explicit thesis-statements to be examined was their precision. In most essays, the thesis-statements were plainly stating the conclusion. However, in a few rare cases interactional elements, namely hedges, were found (examples 8 and 9). Hedges are modal expressions that the writer uses to define either their uncertainty about the statement, or the degree of probability, namely how likely they think the statement to be true. Hedges such as probably, perhaps, or likely were identified in the essays. The hedges are outlined in the examples.

(8) — *I have been able to define two variables that most likely cause the difference in life expectancies, namely the amount of exercise and the level of education. —*

(9) — *The higher life expectancy in Woodby is probably mainly caused by the higher educational levels. —*

While the hedges often convey uncertainty of the writer, in some cases the thesis was very vague as an example 10 (many reasons). Such statements may have indicated uncertainty as well, but it was not expressed in an explicit way as in the hedges above. The vague thesis-statements complicated distinguishing between an essay with and without a thesis. While the writer of the example 10 is vague in their response, the writer in example 11 does not explicate the factors that are influencing the life expectancy, but instead refers to the rest of the essay, the evidence.

(10) — *There are possibly many reasons for the higher average life expectancy in Woodby compared with Brookdale. —*

(11) — *In my report, I examine factors that point to a longer life expectancy in Woodby, compared with Brookdale residents. —*

In such cases the deciding factor was that to be an explicit thesis, the sentence should stand for itself without the rest of the essay. Thus, the example 11 was considered not a thesis while the example 10 was a thesis.

Position Through Interactional Elements

Many essays lacked an explicit thesis. We could not identify a thesis in altogether 70 (36%) essays. However, when we analyzed how these writers related themselves to the available evidence, we found considerable variation. While some essays were strict summaries of available documents, some included various degrees of writer presence.

Even though no thesis could be identified, in a few of these essays, writers took a strong position toward reliability of the materials. For instance, in example 12, the writer questions the trustworthiness of the author of a document.

(12) — *In addition, it is notable that Doctor Dave's identity is open to question and therefore, his information is not reliable as a source. —*

Some essays with no thesis included a commentary or reflective sections. In example 13, the writer expresses their hesitancy about the evidence they are referring to by using the word "apparently." In contrast, in example 14, the writer reflects on their opinion of the facts that they cited, characterizing them as "positive." These are interactional elements, attitude markers to be precise. Such elements convey the writer's affective stand toward the facts, thus implying a position.

(13) — *Apparently the air quality in Woodby is excellent. —*

(14) — *The amount of sleep is according to statistics similar in both towns —. This is a very positive observation. —*

In several essays—despite having no thesis—the writer used first-person expressions, as in examples 15 and 16. It is worth mentioning that such essays also contained passive statements of

facts, so the first-person form did not cover the whole essay. First-person expressions are interactional elements similar to attitude markers. Such self-mentions indicate the writer's presence in the text. Self-mentions are vaguer in their position-taking than attitude markers. However, it is possible to interpret that the writer agrees with the facts in the first-person sentences.

(15) — *I would recommend Brookdale residents to exercise, to get an education and eat healthily.* —

(16) — *I noticed that all sources I browsed through, always mentioned the same problem.* —

Position Through Local Arguments

More than a third of the essays without an explicit thesis had a thematic organization where the writer considered each theme in the materials at a time. For instance, they dealt first with all aspects of exercise that are related to the life expectancy in the two cities, and then moved on to the next theme. In these essays, local arguments were made, but no holistic thesis, summarizing their findings, was present (example 17). The local arguments are underlined.

(17) *During the last 20 years, the Brookdale university has followed life expectancies in Woodby and Brookdale. The data shows that the life expectancy in Brookdale is 79 and in Woodby 84. What causes the 5-year difference in the life expectancies?*

Research shows that getting less exercise increases the risk for a premature death. In Brookdale, 35% of the residents do not exercise at all. Instead, in Woodby, 31% of residents exercise daily, and 29% of the residents exercise regularly. This surely has a positive influence on the life expectancy in Woodby.

The educational level of residents has also a bearing. According to research, 21% of Brookdale residents has a degree in higher education, compared with 34% of Woodby residents. The difference is not that large, but the higher educational level of Woodby residents very likely influences the life expectancy.

If Brookdale residents want to lengthen their life expectancy, they should exercise more, and get a better education.

In essays with local arguments, the writer clearly takes a position. However, they fail to summarize their observations in a holistic manner. They may consider local arguments to be sufficient as a response to the task, leaving the reader to make a synthesis.

Challenges in Distinguishing Between a Thesis and No Thesis

A complication for the thesis identification was the second part of the task, asking for recommendations for Brookdale residents to improve their life expectancy. A couple of writers had integrated their conclusions and recommendations like in example 18.

(18) — *Basing on the information presented above, I recommend Brookdale residents more exercise and going back to school.*

These cases were interpreted as no thesis, since the response to the task question was not explicit. However, it is feasible to assume that such statements implicate students' response and position.

TABLE 3 | Variation in presentation of alternative positions.

Nature of alternative position	Aspect of position-taking	Source of variation	*F*
Rebuttal			53
	Reasoning behind the rebuttal	Rebuttal with thorough reasoning	18
		Rebuttal because research says so	24
		Rebuttal without reasons	7
	Hesitant rebuttals	Unsure rebuttal	8
		U-turn afterward	6
	Attitudinal rebuttals	Beliefs	2
		Irony	4
Contradictory positions		Summarizing opposite sources	43

Variation in Presentations of Alternative Positions

In line with the research questions, the variation in ways that writers used to present alternative positions were analyzed. The findings are summarized in **Table 3**.

Alternative positions were presented in altogether 95 (48%) essays. Two main types of presenting alternative positions, namely rebuttals and contradictory positions, were identified. In rebuttals, the writer explicitly took a position against a piece of evidence within the materials. Of the essays that included alternative positions, 53 (56%) presented a rebuttal. In contrast, the rest of the essays presented contradictory positions. In this type of essay, the writer summarized opposing positions from different sources without taking a personal position toward the issue.

Variation in Rebuttals

In a stereotypical case, a rebuttal included a presentation of the piece of evidence, and an explicit rebuttal, and thorough reasons behind the writer's choice to dismiss the evidence (example 19).

(19) — *The amount of sleep seems to have a bearing on the life expectancy.* — *Nevertheless, when statistics about Brookdale and Woodby residents are examined, it turns out that there is no significant difference in their sleeping habits and therefore, it is not a sufficient reason to explain the higher life expectancy in Woodby.* —

However, not all essays that were identified as presenting a rebuttal were this explicit in their position. Some of the essays with rebuttals simply stated their refutation without expressing reasons behind that position, as an example 20. Often writers considered that "research" was a sufficient reason to rebut a piece of evidence, as an example 21. The rebutting expressions are underlined.

(20) — *The clean air in Woodby does not influence life expectancy despite news reports.* —

(21) — *According to the Woodby Times article, their air quality increases their life expectancy. There is no research proof on this, so I can't agree with the claim.* —

Interestingly, in a few of the essays the writer rebutted a piece of evidence, and afterward they made a U-turn, proceeding to rationalize, why that evidence could still be valid. Such is the case in example 22. This may indicate hesitancy about their position. In a similar vein, some writers were very unsure about their rebuttal, using word the *hardly* (example 23). This type of expression can be seen as an interactional hedge, even though in the example 22, the hedge is not expressed with a single word.

(22) — *It is worth mentioning, that in according to a rumor in Woodby their air quality is rejuvenating. The air quality is undeniably good, but research has not proven that the high life expectancy could be explained by the air. A good-natured belief about a distinctive quality of air does probably not harm the residents. It could even boost their moral and encourage them to take care of themselves, which would increase the life expectancy.* —

(23) — *The air quality of Woodby hardly has any significant influence on the life expectancy. According to research there is no essential difference, and residents' claims are based on individual cases instead of scientifically relevant data.* —

Another type of interactional rebuttal was presented by some writers who reasoned via attitudes. For some writers their own opinion was a sufficient reason for rebuttal (example 24). In addition, some writers took an ironic position toward the evidence they were rebutting, as did the writer mentioning "*magical air*" in example 25. While these are not attitude-markers in the sense of linguistic expressions, their aim is similar.

(24) — *In Woodby Times there is a claim that the air quality of Woodby could be the reason behind Woodby's long life expectancy. Of course, it could be beneficial compared with polluted air, but hypothetically I don't believe it is so much different from Brookdale air.* —

(25) — *A long life expectancy requires versatile nutrition instead of magical air.* —

Contradictory Positions

In contrast to the rebuttals, there was not a lot of variation in the essays with contradictory positions. The lack of variation was probably due to the summarizing nature of these statements. In addition, there was often minimal paraphrasing of the sources.

Typically, in these essays, the writer stated that one of the sources said something that another source opposed (example 26). Sometimes, writers presented more detailed evidence, as an example 27. However, in these cases the wording followed quite closely the original wording in the source documents, indicating problems in paraphrasing (see Hyytinen et al., 2017).

(26) — *The amount of sleep increases life expectancy according to Smith's memorandum, but Doctor Dave says otherwise. Need to investigate more.* —

(27) — *However, there is contradictory information about Woodby region's air quality. According to local Environmental Services unit, the air is rich in oxygen. In other towns with similar air quality, residents live longer than average. In their research, Woodby University has not found support on the theory about the air quality as the source for longevity, but they agree that the air is very fresh.* —

The distinction between a rebuttal and contradictory evidence was mostly straight-forward, but few cases proved challenging. Above in the example 21, the writer had rebutted the evidence citing research. A different wording changes the interpretation. In example 28, the writer cites a source that rebuts the claim instead of rebutting it themselves.

(28) — *An article in Woodby Times contemplates on the Woodby air quality as the secret of long life, but the claim is rebutted by a study on the air.* —

Using passive voice could also indicate hesitancy as in examples 22 and 23 above, and thus, analysis of such statements is not easy. In the present study, a strict view on explicit position-taking was adopted.

Types of Argumentative Writers

Based on the analysis of variation in presentations of position and alternative positions in students' essays, we identified that the degree of writer's presence in the text through their position-taking was a guiding theme across the findings. Therefore, we identified four kinds of writers. Variation in all detected aspects of position-taking was taken in consideration, however, the explicitness of position-taking was found to be the differentiating feature. The process of creating writer groups, and the differentiating features are presented in **Figure 1**.

Based on the analysis, we labeled the first group as *Writers with strong position*. These writers were consistently explicit in their position-taking. In this group, writers took an explicit position toward the evidence, stating an explicit thesis and an explicit rebuttal. Altogether 37 essays represented this group. There was variation in thesis orientation, both types were detected. Likewise, the thesis location varied. In regard to thesis precision, the majority of these essays presented an exact rather than a vague thesis (see example 10 above). Additionally, few of the thesis-statements in this group included hedges, expressing uncertainty. In addition to presenting a thesis, some of these essays also presented local arguments about each theme they discussed. There was more variation in rebuttals in the group. Rebuttal reasoning was varying, some writers presented more comprehensive reasons for their rebuttal than others, and some were more hesitant than others (see examples 19–23). All in all, interactional elements were present in most of the essays in this group. However, interestingly, some writers in this group withheld any other indications of their presence apart from the explicit thesis and the explicit rebuttal.

We labeled the second group as *Writers with volatile position*. These writers did take an explicit position but were more inconsistent compared to the first group. Writers either presented an explicit thesis or a rebuttal, but not both. Altogether 105 essays represented this group. The majority (F = 89) of these essays included only a thesis, but interestingly, some essays (F = 16) presented a rebuttal but not a thesis. Thesis orientation and thesis location varied in these essays just as in the first group. However, there was a difference in thesis precision. There was

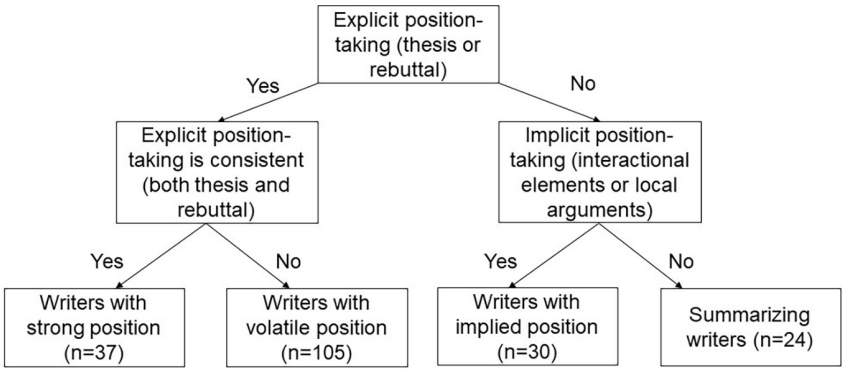


FIGURE 1 | Differentiating features between the writer groups: the process of creating the groups.

more vagueness and uncertainty compared to the first group. Half of those essays that presented only a rebuttal, had local arguments (see example 17), indicating that writers were willing to take a position, but they did not do it in a holistic manner as a thesis. The reasoning behind rebuttals were varied just as in the first group. Interestingly, most of the essays in this group that did not present a rebuttal, did not present any contradictory positions either. The presence of interactional elements was varied in this group.

The third group was labeled as *Writers with implied position*. Instead of being explicit in their position-taking as did writers in the first and second groups, these writers instead showed their presence in their essays using various ways to imply their position. They did not take an explicit position either in the form of a thesis or a rebuttal. Instead, the essays in this group included some interactional elements that implied their association with the evidence. Altogether 30 essays represented this group. Typically, these essays included critique (example 12), commentary or reflective sections (examples 13–14), and first-person expressions (examples 15–16). Approximately half of these essays had some local arguments (see example 17), implying some position toward the evidence. Additionally, about half of the essays in this group presented some contradictory positions.

Finally, the fourth group was labeled as *Summarizing writers*. These writers did not show their position toward the evidence in their essays, but they instead summarized the source materials, either by document or by theme. In other words, no thesis, no rebuttal, and no interactional elements were detected in the essays. Altogether 24 essays represented this group. The only indications toward their presence may have been the choice of summarized documents: if they thought some evidence was not relevant for the task question, they omitted it. Nevertheless, some of these essays included contradictory evidence (examples 26–27). This showed that at least some of these writers acknowledged the importance of diverse viewpoints.

DISCUSSION

The present study gives unique insights into novice students basic level argumentative writing that has received little focus in earlier research. Findings show that there is a large variation in novice students' position-taking. On the bright side, the majority of the students are not entirely clueless about position-taking, but show inclinations to express their viewpoint, which is vital for argumentation. Some guidance by informed teachers might help these students improve their argumentative writing greatly. The findings invite higher education teachers to support novice students in their basic argumentation instead of assuming that they already master all relevant skills.

The findings are in line with earlier studies indicating that some higher education students find aspects of argumentative writing extremely difficult (Petrić, 2007; Laakso et al., 2016; Hyytinen et al., 2021b; Kuhn and Modrek, 2021). Most writers in the present study showed some position-taking regarding their supporting evidence, namely they presented a more or less comprehensive, explicit thesis, or at least made local arguments. Some degree of position-taking was detected in all the writer groups, except for *Summarizing writers* group. The variation in the degree of position-taking was not surprising. Earlier, it has been suggested that Finnish writers are more implicit in their argumentation, compared with Anglo-American writers (Mauranen, 1993). While being implicit is not always a disadvantage, as Mauranen (1993) points out, writers should be aware of requirements and consequences of their texts. Being implicit, namely letting the reader make conclusions, is an efficient way to activate reflectivity in the reader. However, when the writer needs to be sure that the reader comes to the intended conclusion, as in the present study, an explicit thesis-statement is essential. Student writers should learn to be aware of the requirements of each situation. Furthermore, they should learn to be able to identify each situation in order to fulfill its requirements (see Johns, 2008). It was expected that novice students would find discussing alternative positions challenging (Kuhn, 1991; Andrews, 2009; Wolfe et al., 2009; Kuhn and Modrek, 2021). However, half of the writers did present some version of an alternative position in their essay, indicating that many of them understood the importance of diverse viewpoints. In fact, the essays with contradictory evidence were detected in all four writer groups, while explicit rebuttals were much less frequent, and were

detected only in the *Writers with strong position* group and few of the essays in the *Writers with volatile position* group. Writers often incorrectly perceive alternative positions to be a shortcoming for an argument (Perkins, 1989; Wolfe and Britt, 2008; Wolfe et al., 2009), and learning how diverse viewpoints strengthen the message would benefit most students in all writer groups. Additionally, understanding similarities in position-taking regarding supporting and contradicting evidence could be helpful.

It is worth pointing out that in the present study, students used both initial focus and final focus strategies in their essays. This was surprising, as earlier studies have shown that Finnish writers have a strong preference on the final focus strategy (Mauranen, 1993; Mikkonen, 2010). It is possible that the format of the task with a direct question influenced this outcome: initial focus strategy may simply have sprung out of an urgency to respond to the task question. An alternative, intriguing explanation to this finding could be that the globalization and exposure to Anglo-American texts with initial focus may have influenced Finns' rhetoric preferences. However, further research with up-to-date data needs to be conducted in order to draw such conclusions.

Pedagogical Implications

The present findings invite higher education teachers to focus not only on advanced questions of argument validity but also on basic questions concerning how to build argumentation on a textual level, how to identify requirements in each situation, and how to introduce alternative explanations. All higher education teachers, not just writing teachers, should be aware that not all students have learnt basic argumentation in their prior education. Argumentation is difficult, and students need adequate, explicit guidance that focuses on basics (Andrews, 2009; Wingate, 2012; Paldanus, 2020). Fortunately, studies show that even small interventions such as tutorials or exposures to multifaceted texts can help students in their argumentative writing (Wolfe et al., 2009; Kuhn and Modrek, 2021). For instance, analysis of texts with explicit position-taking and summarizing strategies (see Paldanus, 2020), and asking guiding questions about the writer's position could be helpful (see Wingate, 2012). Such small interventions would nudge the *Writers with strong position* and *Writers with volatile position* toward stronger argumentation. However, interventions are not a magic bullet. If students have deficiencies in the basics of composition, as did students in the *Summarizing writers* group, they require more work and guidance. Furthermore, some of the challenges students have in their position-taking may be due to their self-doubt. Novice students—and even senior students—may feel they are not competent in expressing any position (Ivanič, 1998; Andrews, 2009; Mendoza et al., 2022), and teachers should address such perceptions. Giving more space for discussions and debates would benefit all students in developing their expertise and self-confidence. A vital task of higher education is not only to build expertise but to strengthen the sense of expertise in students.

Co-operation between faculty teachers and writing teachers could be beneficial in integrating learning of argumentation with discipline-specific studies. Supporting students in their position-taking and argumentation has wide-ranging benefits to other generic skills that are needed in higher education. Argumentative skills help in developing students' critical thinking, academic writing, and overall communicative skills, in addition to supporting knowledge acquisition (Wingate, 2012; Asterhan and Schwarz, 2016; Iordanou et al., 2019; Kuhn, 2019). However, argumentative assignments are not beneficial for learning if students do not receive guidance in the basics of argumentation (Iordanou et al., 2019).

Higher education teachers should be aware that prior education may give little guidance to argumentation and rhetoric (Andrews, 2009). The emphases are culture-specific, and for instance in the Finnish context, the upper secondary education does not focus on such skills (Marttunen and Laurinen, 2004; Mäntynen, 2009; Mikkonen, 2010; Komppa, 2012). The consequence of this shortcoming is that higher education students need even more support in their academic writing, and teachers should not assume that novice students are fully prepared to take on the academic genre.

In the present study, some students did not answer the question prompted in the task, but their response reflected a broader and more generic topic. Earlier research has similar observations. Students may have difficulties in understanding task assignments and what is expected of them (Bereiter and Scardamalia, 1987; Macbeth, 2006). Understanding principles of the argumentation is futile if students cannot identify situational requirements (Swales, 1990; Mauranen, 1993; Johns, 2008). Consequently, teachers across disciplines are encouraged to focus on clear and precise directions when giving assignments. Giving students opportunities to discuss assignment requirements can help in facing novel situations (see Johns, 2008).

Methodological Reflections

In the present study, two types of triangulation were used, namely investigator triangulation and theory triangulation (Denzin, 1970), strengthening the findings. In investigator triangulation, multiple researchers participated in the data analysis, ensuring that any alternative interpretations were considered and integrated in the analysis. In theory triangulation, multiple theoretical approaches were integrated, namely pedagogical, linguistic, and philosophical theories. This allowed for a practical approach, to support higher education teachers, and not to limit to one theoretical framework.

In the study design and the implementation of the assessment, some limitations were observed. In future studies, these points need to be addressed. The possible ambiguity of assignments needs to be acknowledged in the future. In the present task, students were prompted to write an essay, which is an ambiguous concept at best (see Johns, 2008), and is often associated with study assignments, or exams. While the intention of the task was for the student to take the role of an intern in a city government, the use of the word "essay" may have led some students to associate the assignment with their studies. Such association may have activated a knowledge-display mode instead

of argumentative writing (e.g., Bereiter and Scardamalia, 1987; Petrić, 2007). On the other hand, the task was not a part of students' real studies, and thus, not assessed as an assignment related to their studies. This may have had influence on students' motivation and effort they have put to the task. Additionally, students worked under time pressure, having 60 min to complete the task. This may have influenced their performance as they may have run out of time. The time limitation could even discourage students from engaging in complex cognitive processes (Bereiter and Scardamalia, 1987; Paldanus, 2020). In future studies of argumentative writing, students should be allowed to take their time, to obtain a realistic picture of their skills. However, working under time pressure may possibly reveal about which skills students can effortlessly use in a tight situation and which skills come less easily to them. In any case, all participants in the present study had the same time constraint. Finally, it is important to recognize that a study of end products, i.e., finished texts, does not tell us about strategies and processes students use while writing the text or decisions they make (e.g., Bereiter and Scardamalia, 1987; Hyytinen et al., 2021c). For instance, based on our findings, we do not know, if the *Summarizing writers* made a conscious decision about not expressing their position, or if *Writers with strong position* stumbled across their thesis-statement instead of goal-oriented writing. In future, combining study of texts with cognitive laboratories allowing investigating cognitive processes during writing, would bring a more thorough understanding of argumentative skills and strategies of novice students.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: The dataset is not readily available, as it may reveal information that comes under the copyright of the test developer. Requests to access these datasets should be directed to corresponding author.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

KK contributed to the theoretical background. KK and HH contributed to the data analysis. All authors contributed to writing, and revision of the manuscript, approved the submitted version, and contributed to the design of the study and data collection.

FUNDING

This work was supported by the Finnish Ministry of Education and Culture (Grant Nos. OKM/44/240/2018 and OKM/280/523/2017).

ACKNOWLEDGMENTS

We would like to thank the Council for Aid to Education (CAE) for providing the CLA + International, the participating institutions for administering the assessments and recruiting participants and the Kappas project team for their efforts.

REFERENCES

Andrews, R. (2009). *Argumentation in Higher Education*. Milton Park: Routledge.

Andrews, R. (2015). "Critical thinking and/or argumentation in higher education," in *The Palgrave Handbook of Critical Thinking in Higher Education*, eds M. Davies and D. Barnett (London: Palgrave Macmillan), 49–62. doi: 10.1057/9781137378057_3

Andrews, R., Bilbro, R., Mitchell, S., Peake, K., Prior, P., Robinson, A., et al. (2006). *Argumentative Skills in First Year Undergraduates: A Pilot Study*. York: The Higher Education Academy.

Asterhan, C. S. C., and Schwarz, B. B. (2016). Argumentation for learning: well-trodden paths and unexplored territories. *Educ. Psychol.* 51, 164–187. doi: 10.1080/00461520.2016.1155458

Barrie, S. C. (2006). Understanding What we mean by the generic attributes of graduates. *High. Educ.* 51, 215–241. doi: 10.1007/s10734-004-6384-7

Bereiter, C., and Scardamalia, M. (1987). *The Psychology of Written Composition*. New York, NY: Lawrence Erlbaum Associates.

Biggs, J. (1988). "Approaches to Learning and to Essay Writing," in *Learning Strategies and Learning Styles*, ed. R. Schmeck (New York, NY: Plenum Press), 185–228. doi: 10.1007/978-1-4899-2118-5_8

Breivik, J. (2020). Argumentative patterns in students' online discussions in an introductory philosophy course: micro-and macrostructures of argumentation as analytic tools. *Nord. J. Digit. Lit.* 15, 8–23. doi: 10.18261/ISSN.1891-943X-2020-01-02

Denzin, N. (1970). *The Research Act: A Theoretical Introduction to Sociological Methods*. Chicago, IL: Aldine.

Donald, J. (2002). *Learning to Think: Disciplinary Perspectives*. San Francisco, CA: Jossey-Bass.

Finnish National Board on Research Integrity (2019). *The Ethical Principles of Research with Human Participants and Ethical Review in the Human Sciences in Finland*. Helsinki: Finnish National Board on Research Integrity TENK.

Halpern, D. F. (2014). *Thought and Knowledge: An Introduction to Critical Thinking*, 5th Edn. London: Psychology Press.

Hetmanek, A., Engelmann, K., Opitz, A., and Fischer, F. (2018). "Beyond Intelligence and Domain Knowledge," in *Scientific Reasoning and Argumentation - The Roles of Domain-specific and Domain-general Knowledge*, eds F. Fischer, C. Chinn, K. Engelmann, and J. Osborne (Milton Park: Routledge), 205–226.

Hyland, K. (2005). Stance and engagement: a model of interaction in academic discourse. *Discourse Stud.* 7, 173–192. doi: 10.1177/14614456050 50365

Hyland, K., and Tse, P. (2004). Metadiscourse in academic writing: a reappraisal. *Appl. Linguist.* 25, 156–177. doi: 10.1093/applin/25.2.156

Hyytinen, H., Kleemola, K., and Toom, A. (2021a). "Generic skills and their assessment in higher education," in *Assessment of undergraduate students' generic skills in Finland*, eds J. Ursin, H. Hyytinen, and K. Silvennoinen (Jakarta: Ministry of Education and Culture), 14–18.

Hyytinen, H., Löfström, E., and Lindblom-Ylänne, S. (2017). Challenges in argumentation and paraphrasing among beginning students in educational sciences. *Scand. J. Educ. Res.* 61, 411–429. doi: 10.1080/00313831.2016.1147072

Hyytinen, H., Siven, M., Salminen, O., and Katajavuori, N. (2021b). Argumentation and processing knowledge in an open-ended task: challenges and accomplishments among pharmacy students. *J. Univ. Teach. Learn. Pract.* 18, 37–53. doi: 10.53761/1.18.6.04

Hyytinen, H., and Toom, A. (2019). Developing a performance assessment task in the Finnish higher education context: conceptual and empirical insights. *Br. J. Educ. Psychol.* 89, 551–563. doi: 10.1111/bjep.12283

Hyytinen, H., Toom, A., and Shavelson, R. (2019). "Enhancing scientific thinking through the development of critical thinking in higher education," in *Redefining Scientific Thinking for Higher Education*, eds M. Murtonen and K. Balloo (London: Palgrave Macmillan), 59–78. doi: 10.1007/978-3-030-24215-2_3

Hyytinen, H., Ursin, J., Silvennoinen, K., Kleemola, K., and Toom, A. (2021c). The dynamic relationship between response processes and self-regulation in critical thinking assessments. *Stud. Educ. Eval.* 71:101090. doi: 10.1016/j.stueduc.2021.101090

Iordanou, K., Kuhn, D., Matos, F., Shi, Y., and Hemberger, L. (2019). Learning by arguing. *Learn. Instr.* 63, 101–207. doi: 10.1016/j.learninstruc.2019.05.004

Ivanič, R. (1998). *Writing and Identity: The Discoursal Construction of Identity in Academic Writing*. Amsterdam: John Benjamins.

Johns, A. M. (2008). Genre awareness for the novice academic student: an ongoing quest. *Lang. Teach.* 41, 237–252. doi: 10.1017/S0261444807004892

Kakkuri-Knuuttila, M.-L., and Halonen, I. (1998). "Argumentaatioanalyysi ja hyvän argumentin ehdot," in *Argumentti ja Kritiikki. Lukemisen, Keskustelun ja Vakuuttamisen Taidot*, ed. M.-L. Kakkuri-Knuuttila (Helsinki: Gaudeamus), 60–113.

Kane, M., Crooks, T., and Cohen, A. (2005). Validating measures of performance. *Educ. Meas. Issues Pract.* 18, 5–17. doi: 10.1111/j.1745-3992.1999.tb00010.x

Kleemola, K., and Hyytinen, H. (2019). Exploring the relationship between law students' prior performance and academic achievement at University. *Educ. Sci.* 9:236. doi: 10.3390/educsci9030236

Kleemola, K., Hyytinen, H., and Toom, A. (2021). Exploring internal structure of a performance-based critical thinking assessment for new students in higher education. *Assess. Eval. High. Educ.* doi: 10.1080/02602938.2021.1946482 [Epub ahead of print].

Klein, S., Benjamin, R., Shavelson, R., and Bolus, R. (2007). The collegiate learning assessment: facts and fantasies. *Eval. Rev.* 31, 415–439. doi: 10.1177/0193841X07303318

Komppa, J. (2012). *Retorisen Rakenteen Teoria Suomi Toisena Kielenä - Ylioppilaskokeen Kirjoitelman Kokonaisrakenteen ja Kappalejaon Tarkastelussa*. Helsinki: University of Helsinki.

Kuhn, D. (1991). *The Skills of Argument*. Cambridge, MA: Cambridge University Press.

Kuhn, D. (2019). Critical thinking as discourse. *Hum. Dev.* 62, 146–164. doi: 10.1159/000500171

Kuhn, D., Hemberger, L., and Khait, V. (2016b). Dialogic argumentation as a bridge to argumentative thinking and writing. *J. Study Educ. Dev.* 39, 25–48. doi: 10.1080/02103702.2015.1111608

Kuhn, D., Hemberger, L., and Khait, V. (2016a). Tracing the development of argumentive writing in a discourse-rich context. *Writ. Commun.* 33, 92–121. doi: 10.1177/0741088315617157

Kuhn, D., and Modrek, A. (2021). Mere exposure to dialogic framing enriches argumentive thinking. *Appl. Cogn. Psychol.* 35, 1349–1355. doi: 10.1002/ACP.3862

Laakso, H., Kiili, C., and Marttunen, M. (2016). Akateemisten tekstitaitojen ohjaus yliopisto-opiskelijoiden tiedonrakentamisen tukena. *Kasvatus* 47, 139–152.

Lea, M. R., and Street, B. V. (1998). Student writing in higher education: an academic literacies approach. *Stud. High. Educ.* 23, 157–172. doi: 10.1080/03075079812331380364

Lee, J. J., Hitchcock, C., and Elliott Casal, J. (2018). Citation practices of L2 university students in first-year writing: form, function, and stance. *J. Engl. Acad. Purp.* 33, 1–11. doi: 10.1016/j.jeap.2018.01.001

Macbeth, K. (2006). Diverse, unforeseen, and quaint difficulties: the sensible responses of novices learning to follow instructions in academic writing. *Res. Teach. Engl.* 41, 180–207.

Mäntynen, A. (2009). "Lukion tekstitaidoista akateemiseen kirjoittamiseen," in *Tekstien Pyörityksessä. Tekstitaitoja Alakoulusta Yliopistoon*, eds M. Harmanen and T. Takala (Helsinki: Äidinkielen opettajain liitto), 153–159.

Marttunen, M. (1994). Assessing argumentation skills among Finnish university students. *Learn. Instr.* 4, 175–191. doi: 10.1016/0959-4752(94)90010-8

Marttunen, M., and Laurinen, L. (2004). Lukiolaisten argumentointitaidot – perusta yhteisölliselle oppimiselle. *Kasvatus* 35, 159–173. doi: 10.3114/fuse.2019.03.06

Mauranen, A. (1993). *Cultural Differences in Academic Rhetoric: A Textlinguistic Study*. Bern: Peter Lang.

McCulloch, S. (2012). Citations in search of a purpose: source use and authorial voice in L2 student writing. *Int. J. Educ. Integr.* 8, 55–69. doi: 10.21913/IJEI.v8i1.784

Mendoza, L., Lehtonen, T., Lindblom-Ylänne, S., and Hyytinen, H. (2022). Exploring first-year university students' learning journals: conceptions of second language self-concept and self-efficacy for academic writing. *System* 106:102759. doi: 10.1016/j.system.2022.102759

Mikkonen, I. (2010). *"Olen sitä mieltä, että." Lukiolaisten yleisönosastotekstien rakenne ja argumentointi*. Jyväskylä: University of Jyväskylä.

OECD (2019). *Education at a Glance 2019. OECD Indicators*. Paris: OECD.

Paldanus, H. (2017). "Historian Esseevastauksen Funktionaalinen Rakenne," in *Kielitietoisuus eriarvoistuvassa yhteiskunnassa - Language awareness in an increasingly unequal society. AFinLAn vuosikirja Suomen soveltavan kielitieteen yhdistys AFinLA*. eds S. Latomaa, E. Luukka, & N. Lilja (Helsinki: AFinLAn vuosikirja), 219–238.

Paldanus, H. (2020). *Kuka Aloitti Kylmän Sodan? Lukion Historian Aineistopohjaisen esseen Tekstilaji Tiedonalan Tekstitaitojen Näkökulmasta*. Jyväskylä: University of Jyväskylä.

Perelman, C., and Olbrechts-Tyteca, L. (1969). *The New Rhetoric*. Notre Dame, IN: University of Notre Dame.

Perkins, D. (1989). "Reasoning as it is and could be: an empirical perspective," in *Thinking Across Cultures: The Third International Conference on Thinking*, eds D. Topping, D. Crowell, and V. Kobayashi (Mahwah, NJ: Lawrence Erlbaum Associates), 175–194.

Petrić, B. (2007). Rhetorical functions of citations in high- and low-rated master's theses. *J. Engl. Acad. Purp.* 6, 238–253. doi: 10.1016/j.jeap.2007.09.002

Roderick, R. (2019). Self-Regulation and rhetorical problem solving: how graduate students adapt to an unfamiliar writing project. *Writ. Commun.* 36, 410–436. doi: 10.1177/0741088319843511

Shavelson, R. (2010). *Measuring College Learning Responsibly: Accountability in a New Era*. Redwood City, CA: Stanford University Press.

Swales, J. M. (1990). *Genre analysis. English in academic and research settings*. Cambridge, MA: Cambridge University Press.

Timmermans, S., and Tavory, I. (2012). Theory construction in qualitative research. *Sociol. Theory* 30, 167–186. doi: 10.1177/0735275112457914

Toulmin, S. E. (2003). *The Uses of Argument*. Cambridge, MA: Cambridge University Press.

Tuononen, T., Parpala, A., and Lindblom-Ylänne, S. (2019). Graduates' evaluations of usefulness of university education, and early career success–a longitudinal study of the transition to working life. *Assess. Eval. High. Educ.* 44, 581–595. doi: 10.1080/02602938.2018.1524000

Ursin, J., Hyytinen, H., and Silvennoinen, K. (eds) (2021). *Assessment of undergraduate students' generic skills in Finland: Findings of the Kappas! project*. Chiyoda: Ministry of Education and Culture.

Virtanen, A., and Tynjälä, P. (2019). Factors explaining the learning of generic skills: a study of university students' experiences. *Teach. High. Educ.* 24, 880–894. doi: 10.1080/13562517.2018.1515195

Wingate, U. (2012). 'Argument!' helping students understand what essay writing is about. *J. Engl. Acad. Purp.* 11, 145–154. doi: 10.1016/j.jeap.2011.11.001

Wolfe, C. R. (2011). Argumentation across the curriculum. *Writ. Commun.* 28, 193–219. doi: 10.1177/0741088311399236

Wolfe, C. R., and Britt, M. A. (2008). The locus of the myside bias in written argumentation. *Think. Reason.* 14, 1–27. doi: 10.1080/13546780701527674

Wolfe, C. R., Britt, M. A., and Butler, J. A. (2009). Argumentation Schema and the Myside Bias in Written Argumentation. *Writ. Commun.* 26, 183–209. doi: 10.1177/0741088309333019

Zimmerman, B. J., and Risemberg, R. (1997). Becoming a self-regulated writer: a social cognitive perspective. *Contemp. Educ. Psychol.* 22, 73–101. doi: 10.1006/ceps.1997.0919

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Education

ORIGINAL RESEARCH
published: 11 May 2022
doi: 10.3389/feduc.2022.885825

Linguistic, Contextual, and Experiential Equivalence Issues in the Adaptation of a Performance-Based Assessment of Generic Skills in Higher Education

Jani Ursin[1], Heidi Hyytinen[2], Kaisa Silvennoinen[3] and Auli Toom[2]*

[1] Finnish Institute for Educational Research, University of Jyväskylä, Jyväskylä, Finland, [2] Centre for University Teaching and Learning, University of Helsinki, Helsinki, Finland, [3] Department of Education, University of Jyväskylä, Jyväskylä, Finland

This qualitative study investigated the various linguistic, contextual, and experiential equivalence issues embedded in a performance-based instrument aimed at assessing generic skills in higher education. A rigorous translation and adaptation process (American English to Finnish) was conducted on one instrument, namely *Collegiate Learning Assessment* (*CLA+*) *International*. The data were obtained from cognitive laboratories (*n* = 13), with think-alouds and follow-up interviews conducted among Finnish undergraduate students. Content logs were created, and the data were analyzed thematically. The findings revealed that linguistic and contextual equivalence issues were more prominent than experiential ones. The findings underline how important – and potentially problematic – it is for a test to measure the same construct in a different language and culture. To achieve adequate measurement equivalence, a detailed qualitative analysis of linguistic, contextual, and experiential equivalence should be conducted as part of test adaptation.

Keywords: cross-cultural adaptation, equivalence, higher education, translation, performance-based assessment, generic skills and competences

OPEN ACCESS

Edited by:
Maria J. Hernandez-Serrano, University of Salamanca, Spain

Reviewed by:
Jacqueline P. Leighton, University of Alberta, Canada
Edith Braun, Justus-Liebig-Universität Gießen, Germany

****Correspondence:***
Jani Ursin
jani.p.ursin@jyu.fi

Specialty section:
This article was submitted to Higher Education, a section of the journal Frontiers in Education

Received: *28 February 2022*
Accepted: *25 April 2022*
Published: *11 May 2022*

Citation:
Ursin J, Hyytinen H, Silvennoinen K and Toom A (2022) Linguistic, Contextual, and Experiential Equivalence Issues in the Adaptation of a Performance-Based Assessment of Generic Skills in Higher Education. Front. Educ. 7:885825. doi: 10.3389/feduc.2022.885825

INTRODUCTION

International (comparative) assessments of learning outcomes such as generic skills have become popular in many countries. Several such assessments have been conducted by the Organization for Economic Co-operation and Development (OECD), including the Programme for International Student Assessment (PISA), and the Programme for the International Assessment of Adult Competencies (PIAAC). In higher education there have also been initiatives to measure learning outcomes from a comparative perspective, notably the OECD Assessment of Higher Education Learning Outcomes (AHELO), which investigated what students at the end of their first (bachelor level) degree know and are able to do (see Tremblay et al., 2012). The AHELO was a performance-based assessment that included two complementary components to assess generic skills: selected-response questions (SRQs), and an open-ended performance task (PT). Nonetheless, the AHELO measurement of generic skills has been criticized as being inadequately contextualized and as disproportionately "American" in an international context, with consequent issues of content validity and reliability (Tremblay et al., 2012; Shavelson et al., 2019). In fact, such challenges

are typical of cases in which assessments are developed in a given country and then transferred to other contexts.

Despite this, adapting an existing test has many advantages as opposed to developing and validating a completely new instrument, bearing in mind the resources required, such as high-level expertise in the skills or knowledge being assessed, deep contextual and cultural understanding, and time and money (Ercikan and Lyons-Thomas, 2013; Schendel and Tolmie, 2017). Adapted tests are therefore used frequently, especially in cross-national comparative studies (Hambleton and Patsula, 1998; Hambleton and Lee, 2013). Nevertheless, it is impossible to evaluate such studies and draw conclusions on the findings without a carefully implemented and fully reported translation and adaptation process, with careful attention to equivalence issues (van Widenfelt et al., 2005; Borsa et al., 2012). To successfully adapt a test instrument from one cultural setting to another requires more than merely translating the original test into the target language on a word-for-word basis (Borsa et al., 2012; Ercikan and Lyons-Thomas, 2013; Ercikan and Por, 2020; Ronderos et al., 2021). Typically, the translation and adaptation process includes phases of translation, reconciliation, back translation, expert reviews, pretesting, and evaluation of the final structure (e.g., Karlgren et al., 2020). Especially in selected-response question formats, the evaluation of the final structure is often conducted *via* factor analysis. However, such a quantitative approach is insufficient as a sole indicator of validity when applied to inherently complex performance-based test instruments that include detailed instructions and a number of complex reference documents. Qualitative analyses are therefore needed in efforts to adapt performance-based assessments of generic skills, having in view possible culturally embedded meanings that are difficult to detect by purely statistical means (Ronderos et al., 2021). In fact, most previous studies on the adaptation and validity of performance-based assessments of generic skills in higher education have been quantitative in nature (e.g., Zlatkin-Troitschanskaia et al., 2019; Kleemola et al., 2021), with only a few qualitative analyses of validity (Schendel and Tolmie, 2017; Karlgren et al., 2020).

The present study sought to fill this gap. It did so by analyzing students' response processes while they were carrying out assessment tasks, relating these to the various linguistic, contextual, and experiential equivalence issues embedded in a performance-based instrument aimed at measuring generic skills, namely the *Collegiate Learning Assessment* (*CLA+*) *International.* This measures undergraduate students' generic skills, including problem solving, reasoning, critical reading and evaluation, and written communication. The instrument was originally developed in the United States, then implemented in the Finnish higher education context. Here, linguistic equivalence refers to the notion that words should mean the same thing in the target language. Contextual equivalence refers to an instrument and its parts having the same relevance and being understood in a similar fashion irrespective of the context, and experiential equivalence means that instrument and its parts need to have a similar intention or function in the target culture. Our research question was as follows: *What kinds of linguistic, contextual, and experiential differences can be found in the adaptation of the CLA+ International into Finnish higher education?*

THE CROSS-CULTURAL ADAPTATION OF A TEST INSTRUMENT

In order to obtain reliable results and improve the validity of generic skills measurements, the instrument must be adapted with great care to ensure its usability in a new cultural context. It has often been noted that adapting a test instrument from one cultural setting to another requires more than just translating the original test into the target language on a word-for-word basis (Hambleton, 2005; Borsa et al., 2012; Ercikan and Lyons-Thomas, 2013). Hence, test adaptation aims "to maintain equivalence in content and cultural meaning between the original and the translated/adapted test, thus fostering the comparability of scores across individuals from [. . .] different cultural groups" (Hernández et al., 2020, p. 390).

Test translation and test adaptation are intertwined as concepts and processes. However, according to Hambleton (2005), they refer to different things. *Test adaptation* has been understood as a broad term referring to the various activities that are needed when preparing to use a test in another language and culture; by contrast, *test translation* can be seen as merely one phase of this process (Hambleton, 2005; Ercikan and Lyons-Thomas, 2013; International Test Commission, 2018). This aspect is dealt with in greater detail by Ronderos et al. (2021; see also Hambleton and Patsula, 1998), with translation being seen as the creation of linguistically equivalent versions of a test, in contrast with adaptation, which involves cultural considerations such as equivalence of the construct, similarity of test administration, speed of response, and familiarity with the item format. For its part, the term *cross-cultural adaptation* can be used to describe a process in which not just language, but also other aspects related to cultural adaptation, are taken into consideration in translating and adapting a test instrument to a new cultural context (Beaton et al., 2000; Ercikan and Por, 2020).

The research literature presents a large number of guidelines and suggestions for adapting test instruments for use in another culture and for evaluating the quality of this process (see Beaton et al., 2000, 2007; Hambleton, 2005; Gudmundsson, 2009; Borsa et al., 2012; Ercikan and Lyons-Thomas, 2013; Hambleton and Lee, 2013; International Test Commission, 2018; Oliveri and Lawless, 2018; Hernández et al., 2020). Nonetheless, researchers have noted that there is no clear agreement on the ideal adaptation method (Borsa et al., 2012; Epstein et al., 2015). The test adaptation process may vary depending on the instrument and its intended use (Gudmundsson, 2009; Borsa et al., 2012; Hernández et al., 2020).

Typically, test adaptation includes the phases of (1) considering whether the measured construct can be captured by a test in another cultural context, (2) translating/adapting the test (by competent translators) and deciding on the kinds of accommodations needed in order to use the test in another language and culture, (3) evaluating the quality and equivalence of the translations, and (4) pretesting the adapted test (see

Beaton et al., 2000, 2007; Hambleton, 2005; Borsa et al., 2012). The first step refers to how far the intended construct has a similar meaning in different cultures (Hambleton, 2005). Such construct equivalence between source and target cultures is crucial, as without it, cross-cultural comparisons are impossible (Hambleton, 2005; Ercikan and Lyons-Thomas, 2013).

Secondly, test adaptation guidelines emphasize the importance of an accurate translation process. The recommendation is to use multiple trained translators who are familiar with both the source and the target languages and cultures (Hambleton, 2005; Beaton et al., 2007). To obtain translation accuracy, translators should be provided with sufficient information on the nature of the instrument being adapted (Hambleton, 2005; Arffman, 2013). In addition to forward translation, there have been recommendations also to use back translation (i.e., having the adapted test translated back into its original language) as a step to evaluate the quality and validity of translations (e.g., Beaton et al., 2000, 2007; Borsa et al., 2012). However, in their review regarding cross-cultural adaptation methods and guidelines, Epstein et al. (2015) noted that back translation has generated considerable controversy: some practitioners have regarded it as an essential part of cross-cultural adaptation whereas others make no such recommendation, especially in cases where the adaptation team speak both the source and the target language.

Thirdly, different translations made independently by translators should be synthesized and then evaluated by an expert group (Beaton et al., 2000, 2007; Borsa et al., 2012). In this way, possible equivalence issues and sources of translation/adaptation errors can be identified. The equivalence between two language versions of the test may be lacking for a variety of reasons. For example, translations may change the content or meaning of test items. In order to maintain item equivalence, it is essential to consider to what extent and in what way this change has taken place (Ercikan and Lyons-Thomas, 2013). The aim of translations is not just to find words but also expressions and concepts that have both linguistically and culturally similar meanings in the target culture (Hambleton, 2005). Literal translation is unlikely to be the optimal way to proceed, as it can lead to errors in terms of test content, linguistic, or cultural factors (van Widenfelt et al., 2005; Karlgren et al., 2020) – a phenomenon also referred to as "unwanted literal translation" (Arffman, 2012). Indeed, all translation requires some degree of adaptation, as translations depend on characteristics of the target language including "its interplay with the intended meaning of a test item and the features of the source and target culture and population" (Ronderos et al., 2021, p. 66).

Finally, the adapted test and its functionality should be pretested in practice within the intended target group (Beaton et al., 2000, 2007). In addition to examining the content and characteristics of test items, pretesting makes it possible to evaluate other factors related to the test, such as the functionality of the instructions (Borsa et al., 2012; Hambleton and Lee, 2013). Indeed, pretesting an instrument is particularly crucial for performance-based assessments of generic skills, which include open-ended questions (such as performance tasks), detailed instructions, and several qualitative background documents on which students must base their answers. Properly conducted pretesting of a performance-based assessment will help to reveal possible sources of error that might threaten the validity of the instrument. In the performance-based assessment of generic skills it is pivotal that the questions should not contain unfamiliar words or complicated structures that would produce comprehension problems (Johnson et al., 2009).

As indicated above, equivalence is imperative in translation and in the adaptation of a test from one culture to another. Equivalence refers to the requirement that different language versions should be comparable to each other, and measure the same construct (Arffman, 2013). Overall the literature presents various forms and categorizations of equivalence. According to Karlgren et al. (2020; see also Borsa et al., 2012) one needs to check whether words mean the same thing (*semantic equivalence*), whether colloquialisms or idioms need to be replaced (*idiomatic equivalence*), and whether the "same" word holds a different conceptual meaning in the culture (*conceptual equivalence*). *Experiential equivalence* is also important. This means that items have to be replaced by something addressing a similar intention or function in the target culture; for example, knife and fork may need to be replaced with chopstick if that is the common utensil used for eating in target culture. Furthermore, participants in different cultures may not be equally familiar with certain test item types, such as selective-response questions (Hambleton and Patsula, 1998; Hambleton, 2005; Ercikan and Por, 2020). This aspect relates to *item equivalence* (Herdman et al., 1998). The structure of the test instrument or the way in which the test is administered are also important factors to consider from the perspective of cultural adaptation (e.g., Herdman et al., 1998; Hambleton, 2005; Schendel and Tolmie, 2017). This is known as *operational equivalence* (Herdman et al., 1998). In addition, *measurement equivalence* – meaning that the two versions should not differ significantly in their psychometric properties – is often viewed as a distinct form of equivalence (Herdman et al., 1998; Epstein et al., 2015).

Because the concept of equivalence has various forms and meanings and many of them are closely linked to each other, we see is as useful to summarize the forms of equivalence that we apply. These are: (1) *linguistic equivalence*, incorporating elements from semantic, idiomatic and conceptual equivalence and referring to the notion that words should mean the same thing in the target language, (2) *contextual equivalence*, meaning that an instrument and its parts have the same relevance and are being understood in a similar fashion irrespective of the context, and (3) *experiential equivalence*, based on the notion that an instrument and its parts should have a similar intention or function in the target culture.

COLLEGIATE LEARNING ASSESSMENT INTERNATIONAL AS A PERFORMANCE-BASED TEST INSTRUMENT

This study utilized the test instrument *Collegiate Learning Assessment* (*CLA+*) *International*. CLA+ International is a subject-independent performance-based assessment developed

by the United States-based Council for Aid to Education (CAE), which measures undergraduate students' generic skills. For our part, we understand generic skills as universal expert skills needed in studies and working life. In higher education, higher-order skills such as analytical reasoning skills and problem-solving skills are typically valued more highly than practical generic skills. Performance-based assessment aims to cover generic skills in an authentic manner by giving an opportunity for students to demonstrate their skills measured in the assessment task (Shavelson, 2010; Hyytinen et al., 2021). Performance-based assessment refers to a variety of task types, such as open-ended performance-task and document-based selected-response questions. Typically, a performance-based assessment will challenge students to use their higher-order thinking skills to create a product or to complete a process (Braun et al., 2020). Indeed, a performance assessment can be viewed as "an activity or set of activities that requires test takers [. . ..] to generate products or performances in response to a complex, most often real-world task" (Davey et al., 2015, p. 10). Thus, students actively participate in the problem-solving exercise and may even learn during the performance-based assessment (cf. Kane, 2013), rather than passively selecting answers (Palm, 2008; Hyytinen et al., 2021).

In line with this definition of performance-based assessment, CLA+ International includes three components. First of all, a student has 60 min to respond to a performance task which measures analysis and problem solving, writing effectiveness, and writing mechanics. The performance task includes an open-ended question in which students are asked to produce a justified solution to a presented real-life problem, utilizing in their written response different source materials from an online Document Library. Responding to the performance task requires students to simultaneously use a range of generic skills, as they need to analyze and evaluate information, make conclusions, and provide evidence for their own solution or recommendation (Shavelson, 2010; Zahner and Ciolfi, 2018; Hyytinen et al., 2021). In this study, the performance task was about comparing life expectancies in two cities, and students had to consider whether some measures were needed to increase the life expectancy in one of the cities. In their responses, the students had to present a solution to the problem and to give a recommendation for possible measures. The source materials that students needed in order to formulate their response contained five different source documents: a blog text, a transcribed podcast, a memorandum, a newspaper article, and infographics (see Ursin et al., 2021).

Thereafter, students had 30 min to answer 25 selected-response questions. Ten of the questions were relevant to the background document, which dealt with the secretion of proteins in the brain. These questions measured scientific and quantitative reasoning. A second set of ten questions, based on a letter about nanotechnology sent by a reader to an imaginary journal, measured critical reading and evaluation. The last five questions, which related to an opinion piece on women in combat, assessed the student's ability to analyze arguments, including possible logical fallacies. At the end of the test, the students filled in a background information survey (Ursin et al., 2021). Because the test tasks are still used internationally, the performance task and selected-response questions used in this study cannot be published or described in a more detailed manner. However, similar test tasks are presented by Shavelson (2010) and Tremblay et al. (2012).

TRANSLATION AND ADAPTATION OF COLLEGIATE LEARNING ASSESSMENT INTERNATIONAL TO THE FINNISH CONTEXT

The CLA+ test was translated into Finnish. The translation and adaptation of the test progressed through four main steps as specified in the guidelines of the International Translation Committee (International Test Commission, 2018; cf. Hambleton and Patsula, 1998). In the first phase, the test was translated from English into Finnish by a qualified translator with knowledge of large-scale assessments in the field of education. In the second phase, two trained translators in Finland reviewed, confirmed and, if necessary, proposed changes or corrections to the translations independently of each other. In the third phase, the project team in Finland decided on the final versions of the translations on the basis of the translators' proposals. The reconciled revisions were then verified by the test developer in the United States. The translated test was then pretested in Finnish in "cognitive laboratories," ensuring that the translation and adaptation phase had not changed the meaning, the level of difficulty, or the internal structure of the test (see Ursin et al., 2021). The suitability of the test for the Finnish context was ensured in detail. The translation and adaptation of the test instrument did not include a phase of back translation, since, as noted above, previous studies (e.g., Epstein et al., 2015) have indicated that it may not be a necessary step, especially if the research personnel speak both the source and the target language, which was the case in this study.

AIMS, MATERIALS, AND METHODS

The main aim of the study was to identify various equivalence issues in adapting the CLA+ International instrument. More specifically, we focused on the differences to be found in the adaptation of CLA+ International from the United States context to Finnish higher education (cf. Hambleton, 2005; Borsa et al., 2012; Karlgren et al., 2020), in line with our categorization of issue types. We see the differences as involving:

(1) Linguistic issues (whether words mean the same thing in the target language);
(2) Contextual issues (whether an instrument or its parts has the same relevance and are being understood in a similar fashion irrespective of the context);
(3) Experiential issues (whether the instrument or its parts have a similar intention or function in the target culture).

The data came from 13 cognitive lab events with think-alouds and follow-up interviews, conducted on a target group of students (**Table 1**). The participants, who all were

TABLE 1 | Demographics of the participants in the cognitive labs.

Gender	Field of study	Type of higher education institution
Male	Humanities and Arts	University
Female	Humanities and Arts	University
Male	Science	University
Female	Science	University
Male	Humanities and Arts	University
Female	Health and Welfare	University of Applied Sciences
Female	Health and Welfare	University of Applied Sciences
Female	Health and Welfare	University of Applied Sciences
Female	Services	University of Applied Sciences
Male	Engineering, manufacturing, architecture, and construction	University of Applied Sciences
Male	Arts	University of Applied Sciences
Female	Arts	University of Applied Sciences
Female	Arts	University of Applied Sciences

white Caucasians, represented two large multidisciplinary higher education institutions in southern Finland. One of the institutions was a research-intensive university, and the other was a professionally oriented university of applied sciences. Participation in the research was voluntary, and informed consent was obtained from the participants. The cognitive labs made it possible to collect authentic data on participants' ongoing thinking processes and behaviors while they were working on a task (van Someren et al., 1994; Leighton, 2017, Leighton, 2019). The data were collected individually from all the participants by following a similar procedure. At the start of each lab, the participants were instructed and trained to think aloud as they were solving the tasks. Verbalization took place when the participant first completed the performance task, and thereafter during 25 selected-response questions in a secured online environment. To avoid bias in the data collection, a neutral form of the think-aloud protocol was used (van Someren et al., 1994; Leighton, 2017). It follows that the participants were not interrupted while they were performing the tasks. "Keep talking" was the only probe given during the lab if the participant was silent for a long time. The researchers sat in the back of the room and kept their distance from the participants while they completed the tasks. The neutral form of the think-aloud protocol ensured that the probing questions were not asked until the follow-up interview. In the second phase, after the think-aloud, a short follow-up interview was conducted. This included both targeted questions (based on the observations during the think-aloud phase) and general questions posed to all participants (covering notably the clarity of the instructions, the comprehensibility of the test, how interesting the test was, the strategy used for answering). The first- and second-named authors collected the data.

Each lab lasted around 2 h and was videotaped and recorded by a camera and a table microphone. In addition, notes were taken by the researchers. The verbalizations of each participant during the cognitive labs were transcribed verbatim. After that, content logs were created in which accurate descriptions of non-verbal actions, a summary of events, and transcriptions of the verbalizations of each participant were combined into one text file (Oranje et al., 2017; see **Table 2**). A content log provides an overview of the video data, and it can be used to locate sequences and events for further analysis. The log externalizes and visualizes participants' thinking processes and behaviors associated with the assessment constructs and progress in the task. A strength of the log is that it encompasses all the input provided by a test-taker, i.e., direct quotes, assertions, behavior, and actions that took place during the think-aloud process. The log includes information on the sequence, timing, and variety of the test-taker's response behaviors and actions. Furthermore, it combines both verbal and non-verbal response processes (Hyytinen et al., 2014; Oranje et al., 2017).

The transcripts and content logs were analyzed using a thematic approach in which, initially, similar notions were systematically coded under preliminary content categories. Subsequently, final categories were formed on the basis of a relational analysis (Braun and Clarke, 2006). Finally, the preliminary categories were further elaborated on a theoretical basis, with special attention paid to issues of contextual, linguistic, and experiential equivalence (Hambleton, 2005; Borsa et al., 2012; Karlgren et al., 2020). Furthermore, numbers of occurrence were calculated in order to reach an understanding of how typical a given category might be. The first and third author of this paper did the initial coding; this was then revised against the coding made by the second and fourth authors. Thereafter, the final categories were discussed and agreed with all the authors. Translated and anonymized excerpts from the cognitive laboratories were selected for each category for illustrative purposes.

RESULTS

All of the participants (n = 13) experienced equivalence issues while taking the test. The analysis identified several linguistic, contextual, and experiential issues (**Table 3**). The equivalence issues identified related mainly to how questions were formulated, and how materials were comprehended; also to how the instructions were presented, and how certain concepts were understood.

Linguistic Equivalence

Most of the issues related to linguistic differences between Finnish and English. By linguistic equivalence we refer to the notion that the meaning of the words and phrases should not have altered in the translation and adaptation from English into Finnish. One language-related difference concerned the phrasing of the questions. Efforts had been made to keep the equivalence between English and Finnish phrases as close as possible, but this occasionally created situations where it was difficult for a student to understand the translated question. An example from a cognitive lab reads as follows:

> *[The participant] reads the question, ponders for a moment what it says ("the following criteria, that is, these [criteria] except one of them, is that so?")* (SRQ item 3 – ID17).

TABLE 2 | An example of a content log (Hyytinen et al., 2021).

Time	Duration	Code of student: ID2
0:00:00–0:00:27	0:00:27	Logging into the test plus the privacy notice Glances through the privacy notice and asks how to move on. Asks the same thing also at the summary of the test. Glances through the summary of the test
0:00:27–0:08:14	0:07:47	Performance task
0:00:27–0:01:08	0:00:41	Reads and glances through the general instructions for the performance task
0:01:08–0:08:14	0:07:06	Moves to the actual performance task. First, quickly reads the task instruction and some of the documents. Moves directly to writing the answer, does not plan it beforehand. Browses the documents. Concentrates on the infographic. Using that as a basis, says that "the physical activity habits of the residents should be improved." Does not substantiate the answer more precisely, compare the information in the documents, or evaluate the reliability of their content aloud. Completes the answer in no more than 8 min and moves on to the SRQ items Written response: *The physical activities of Brookdale's inhabitants should be improved. The inhabitants must be told about a healthy diet. The education level must be improved*
0:08:14–0:59:52	0:51:38	SRQs
0:08:14–0:10:39	0:02:25	Moves to the SRQ section. Browses through the SRQ instructions. Asks for help on how to move on
0:10:39–0:15:32	0:04:53	Glances at the first question and items, then the document provided for the first SRQ set. Moves back to the first question and items, then identifies the relevant section from the source document. Compares the items to the document. Thinks aloud which item (A–D) would most weaken the main claim of the document. Says that "option A could be true based on the document, hence A is not the right answer." Selects option D. Moves to the second question

TABLE 3 | Linguistic, contextual, and experiential equivalence issues in the data (with number of occurrence).

	Linguistic equivalence (*n* = 29)	Contextual equivalence (*n* = 20)	Experiential equivalence (*n* = 4)
Questions (*n* = 19)	Phrasing of the questions (*n* = 19)		
Materials (*n* = 17)	Difficulties in understanding the text (excessive use of abbreviations) (*n* = 6)	Differences in understanding a reliable source of information (*n* = 6)	
		Proper interpretation of a figure (map) (*n* = 5)	
Instructions (*n* = 13)	Difficult linguistic expressions (*n* = 4)	Multitude of instructions (*n* = 5)	
		Usefulness of instructions (*n* = 4)	
Concepts (*n* = 4)			Difficulties in understanding the meaning of concepts in the Finnish context (*n* = 4)

The way the question was posed was not a typical way of presenting a question in Finnish, thus making it rather difficult to understand. Nonetheless, changing the formulation of question into a more "Finnish" formulation might also have impacted on the difficulty of the question (made it less difficult). Hence, no substantial changes were made to the formulation of this particular question.

Another language-related issue was how the instructions were given in the online test environment. This resulted in situations, for example, where students were unsure how to move forward in the test platform because they were confused about the linguistic expression and symbol represented (in the original) by "mark complete." The following excerpt from a cognitive lab exemplifies this:

> *[The participant] is wondering for a moment how to move forward from the instructions, until she clicks on "mark complete"* (SRQ – ID19).

"Mark complete" was initially translated into Finnish (literally) as "merkitse valmiiksi" which is not a typical (although a possible) way of expressing that one can now move on to the next page in the online platform. To make the instruction more understandable it was ultimately changed to the more conventional "valmis" (meaning "completed" in English), thus avoiding the pitfall of unwanted literal translation (Arffman, 2012).

The final language issue related to the background materials used in the test instrument. These were required for a student to answer the questions. Typically, Finnish does not use abbreviations as readily as English. One of the SRQ documents (regarding secretion of proteins in the brain) included an excessive use of difficult abbreviations (from the point of view of the Finnish language) making it challenging and occasionally frustrating for students to understand the text. This is reflected in the following excerpt from a cognitive lab: [*The participant] notes that she is too tired to concentrate on a text filled with abbreviations* – (SRQ item 1 – ID14). Although some other students also reported challenges related to the use of abbreviations, this example of frustration emerged as an extreme case, and ultimately no changes were made to the background document.

Contextual Equivalence

Several context-related equivalence issues became visible in the analysis. By contextual equivalence we refer to the requirement for an instrument and its parts to have the same relevance and

meaning in the target culture as it has in the culture of departure. The first contextual equivalence issue related to the background materials on which the students had to base their answers. On a few occasions the students wondered what a reliable source of information might actually be ultimately leading to a question whether reliable source of information is comprehended similarly among undergraduate students in United States and in Finland. One of the participants felt that some of the documents were ridiculous, almost to the point of neglecting the document altogether:

> *Wonders aloud about those [documents]; notes that the articles, on the basis of which the report should be made, seem a bit "silly." Questions the relevance of Document 4, does not find it a reliable/relevant source and is going to ignore it* (Performance task – ID12).

This had an impact on the quality of the answer, as students were informed in the instructions for the performance task that their answers would be judged on how thoroughly the information was covered, including mention of potential counterarguments. Hence, completely ignoring some of the documents might have resulted in poorer test scores. Another example relating to contextual equivalence concerned the interpretation of the figures in one of the background documents. Thus, in order to interpret one of the figures correctly, a student should be familiar with the intermediate compass points (such as South-East) on a map representing the United States. As the use of intermediate compass points to describe Finland as a country is not as typical as it is in the United States, this led to challenges for some students in attempting to answer a question. One of the students reacted as follows:

> *[R]eads the question and examines the figure. Is not sure about the compass points and comments that this task makes no sense if you don't recall [the compass points]* (SRQ item 10 – ID3).

The second contextual equivalence issue related to how instructions in the background documents were formulated. As compared to common practice in the United States, in Finland matters are typically presented without much guidance or orientation to the reader. Consequently, metatext tends to be used much less in the Finnish context than in the Anglo-American context (see Mauranen, 1993). In several background documents, multiple instructions were given, including lengthy guidance. This caused confusion to some of the participants. The following example from a cognitive lab illustrates the challenges due to the excessive use of orientation text:

> *[The participant] moves on to the instructions for the performance task, reads/goes through it. Notes that there is much to read in the instructions* (Performance task – ID13).

Another participant was uncertain whether all the instructions in the test were actually needed or relevant:

> *[The participant] is reading the privacy notice and asks [from the researchers carrying out the cognitive lab] if one can just accept it. Reads the summary of the test and asks for specifications about task duration. Asks if the instruction section can be skipped* (ID3).

Experiential Equivalence

There were also some experiential equivalence issues found in the data. By experiential equivalence we mean that the instrument and its parts should have a similar intention or function in the target culture. There were a few concepts such as "drinking water" and "ordinary diet" which might have been experienced differently by the Finnish participants. In Finland, drinking water is typically the same as tap water (which is high-quality, drinkable, and of similar taste across the country), but this is not the case in the United States. Although this point was not particularly crucial from the point of view of answering the question posed in the performance task, it might have led to a different understanding of the concept of "drinking water" depending whether one was an undergraduate student in the United States or in Finland. One of the participants pondered what the "ordinary Finnish diet" mentioned in Document 5 might actually mean (Performance task – ID13). "Ordinary" was initially translated as "tavallinen" in Finnish, but was ultimately changed to "perinteinen," meaning "traditional" in English. Nonetheless, the issue remained whether an ordinary/traditional diet means the same thing in the United States and in Finland.

DISCUSSION

By analyzing students' response processes during tasks, our study aimed to identify various linguistic, contextual, and experiential equivalence issues embedded in the CLA+ International translation and adaptation process from the United States to the Finnish context. In our study, linguistic and cultural equivalence issues emerged interestingly as more crucial than experiential ones (cf. Hambleton, 2005; Borsa et al., 2012; Karlgren et al., 2020). The issues of linguistic equivalence were associated with the formulation of questions, difficulties in understanding some linguistic expressions in the instructions of the test, and challenges in comprehending one of the SRQ documents due to an excessive use of abbreviations. Contextual equivalence issues were related to the interpretation of a figure (how to make sense of a map of the United States), and to the abundance and utility of the instructions posed in the test instrument, with difficulties also in understanding what a reliable source of information could consist of. There were only a few issues of experiential equivalence, linked to difficulties in comprehending the meaning of certain concepts (such as "drinking water") in the Finnish context. Nevertheless, our findings show that linguistic, contextual, and experiential factors need to be taken into account in interpreting the performance-based assessment of generic skills. All these aspects affect how students interpret the task, instructions, questions and materials used in assessments, and how they formulate their responses (Ercikan and Por, 2020). If students face unfamiliar or completely new ways of presenting a test in a situation, this may demand additional capacity from them, and thus influence their performance in the test.

The findings confirm previous notions of what is needful for translations: not merely to find words, but also expressions and

concepts that have both linguistically and culturally a similar meaning and intention in the target culture (Hambleton, 2005; Arffman, 2012; Ercikan and Por, 2020). An example of this occurred in the way a student was given instructions when answering questions. Because Finnish texts tend to include only minimal explicit metalanguage to orient the reader, as compared to Anglo-American texts (Mauranen, 1993; Kleemola et al., 2022), the multitude of instructions in some instances raised concern over whether the instructions actually embodied the same intention in Finnish cultural context. Some of these findings may seem minor, but they could have a considerable impact in the test situation and on students' test performance.

Our findings also indicated the extent to which equivalence issues can be intertwined. For example, how a word or term is translated (linguistic equivalence) might also change how it is understood in different contexts (experiential equivalence). An example in our data was the term "ordinary diet," which was confusing to Finnish participants, and led to a revision of the translation (to "traditional diet") in the final version of the test. One can then ask whether undergraduate students in United States higher education would understand the term "traditional diet" in a similar fashion to their counterparts in Finland. Another example of the intertwined nature of equivalence issues in our data was about how the Finnish undergraduate students can make sense of a map of United States when the use of intermediate compass points in Finland is not as typical as in the United States. While this is contextual equivalence issue (whether the map of the United States has the same meaning in Finland as in the United States) this is also "experiential" issue insofar as it relates to geography, and the large area of the American land mass, and the geographical variations it contains. This experience of United States as a country is something that the Finnish participants are lacking. Overall, our findings showed the extent to which the translation and adaptation of CLA+ International from American to Finnish context involved a process of carefully balancing between content, language, and experiential factors (see van Widenfelt et al., 2005; Karlgren et al., 2020).

Our findings contribute to the assessment literature by suggesting a need for greater recognition of equivalence and validity issues in translated and adapted performance-based assessments of generic skills in higher education. This is important in order to guarantee collecting high-quality research data. This is of crucial importance as opposed to pure selected-response questions, performance-based assessments typically include a complex set of background documents and instructions. As shown in our study also, these increase the likelihood of cultural, linguistic, and experiential equivalence issues in the test instrument (cf. Ercikan and Por, 2020). Consequently, in performance-based assessments it is crucial to identify equivalence issues if one is seeking to diminish their effect on participants' test-taking. Our findings importantly support the notion that to ensure that a test measures the same construct in a different language and culture, a qualitative analysis of equivalence issues is a necessary part of test adaptation, together with psychometric evidence (e.g., Ercikan and Pellegrino, 2017). However, it is important to note that without cognitive laboratory data, it would not have been possible to gather authentic data on participants' ongoing response processes while they worked on a task. A key observation of our study to the assessment literature is that qualitative analyses of cognitive laboratory data are of enormous help in revealing possible challenges in the validity and equivalence of an adapted test instrument. In the long run, such research is crucial for the development of the generic skills research field, which at present lacks robust replicable instruments (Braun et al., 2012; El Soufi and See, 2019; Tuononen et al., 2022).

LIMITATIONS OF THE STUDY AND FUTURE RESEARCH

The findings of our study can be used to improve the quality of a translated and adapted generic skills assessment instrument. However, certain limitations in the study should be taken into account. The first of these relates to the relatively small amount of data obtained, given that the data comprised 13 cognitive labs with think-alouds. This nevertheless resulted in around 26 h of recorded data, and it can be claimed that saturation was reached in terms of sufficiency of the data. In the future, cognitive labs could be carried out with a more versatile group of undergraduate students (i.e., from different disciplinary backgrounds), though one has to bear in mind that the setting up of cognitive labs followed by detailed analysis of the data (including the creation of content logs) requires considerable resources. A further limitation concerns the think-aloud method. It is possible that the participants' ability to verbalize their thoughts might have biased the think-aloud data. Note, however, that in order to minimize bias in the data collection, we followed a formalized procedure. This included instructions and explanations to participants on thinking aloud, a brief training session, and a neutral protocol that avoided probing questions. In this way, we endeavored to ensure the reliability of the verbal data (van Someren et al., 1994; Leighton, 2017). A third limitation is related to the three forms of equivalence (linguistic, contextual and experiential) that we used in our paper. The different forms of equivalence are intertwined to the extent that it is difficult to make a clear-cut analytical distinction between the different manifestations of equivalence. Furthermore, contextual and experiential equivalence is strongly related to the characteristics of the participant; if participants represent a sub-culture or belong to a particular ethnic group, they may have a different understanding of an instrument (or parts of it) from that of the majority of the population in the target culture. A fourth limitation is linked to the fact that the translated and adapted test instrument included only one performance task and one set of selected-response questions. A more reliable picture of the equivalence issues would have been obtained by including more than just one of each type of task. Hence, in future it would be important to study the equivalence issues pertaining to different kinds of performance tasks and selected-response questions, since these might enter into the performance-based assessment of generic skills in higher education. Furthermore, it would also be important to develop tasks in an international context by a knowledgeable team of experts, and to study whether

such tasks would include fewer equivalence issues than those developed in a single country. Note also that linguistic, cultural, and experiential equivalence issues appear to be closely bound up with the methodological and technical aspects of a test instrument (Hambleton, 2005).

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because they contain information that could potentially identify the participants, and thus compromise their anonymity. The datasets may also reveal information on the test that comes under the copyright of the test developer. Requests to access the datasets should be directed to the corresponding author.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the Local Legislation and Institutional Requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

FUNDING

The present study was supported by the Finnish Ministry of Education and Culture. This study was carried out through the funding of a project entitled *Korkeakouluopiskelijoiden oppimistulosten arviointi Suomessa* (KAPPAS!).

ACKNOWLEDGMENTS

We would like to thank Donald Adamson for the helpful comments and proof-reading the manuscript.

REFERENCES

Arffman, I. (2012). Unwanted literal translation: An underdiscussed problem in international achievement studies. *Educ. Res. Int.* 2012, 1–13. doi: 10.1155/2012/503824

Arffman, I. (2013). Problems and issues in translating international educational achievement tests. *Educ. Meas. Issues Pract.* 32, 2–14. doi: 10.1111/emip.12007

Beaton, D., Bombardier, C., Guillemin, F., and Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 25, 3186–3191. doi: 10.1097/00007632-200012150-00014

Beaton, D., Bombardier, C., Guillemin, F., and Ferraz, M. B. (2007). *Recommendations for the Cross-Cultural Adaptation of Health Status Measures.* Available online at: http://dash.iwh.on.ca/sites/dash/files/downloads/cross_cultural_adaptation_2007.pdf (accessed May 20, 2021).

Borsa, J. C., Damaìsio, B. F., and Bandeira, D. R. (2012). Cross-cultural adaptation and validation of psychological instruments: Some considerations. *Paideìia* 22, 423–432. doi: 10.1590/1982-43272253201314

Braun, E., Woodley, A., Richardson, J. T. E., and Leidner, B. (2012). Self-rated competences questionnaires from a design perspective. *Educ. Res. Rev.* 7, 1–18. doi: 10.1016/j.edurev.2011.11.005

Braun, H. I., Shavelson, R. J., Zlatkin-Troitschanskaia, O., and Borowiec, K. (2020). Performance Assessment of Critical Thinking: Conceptualization, Design, and Implementation. *Front. Educ.* 5:156. doi: 10.3389/feduc.2020.00156

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Davey, T., Ferrara, S., Shavelson, R., Holland, P., Webb, N., and Wise, L. (2015). *Psychometric Considerations for the Next Generation of Performance Assessment.* Available online at: https://www.ets.org/Media/Research/pdf/psychometric_considerations_white_paper.pdf (accessed Feb. 17, 2022)

El Soufi, N., and See, B. H. (2019). Does explicit teaching of critical thinking improve critical thinking skills of English language learners in higher education? A critical review of causal evidence. *Stud. Educ. Eval.* 60, 140–162. doi: 10.1016/j.stueduc.2018.12.006

Epstein, J., Santo, R. M., and Guillemin, F. (2015). A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J. Clin. Epidemiol.* 68, 435–441. doi: 10.1016/j.jclinepi.2014.11.021

Ercikan, K., and Lyons-Thomas, J. (2013). "Adapting tests for use in other languages and cultures," in *APA Handbook of Testing and Assessment in Psychology, Testing and Assessment in School Psychology and Education*, eds K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, et al. (Washington, DC: American Psychological Association), 545–569. doi: 10.1037/14049-026

Ercikan, K., and Pellegrino, J. W. (2017). "Validation of score meaning using examinee response processes for the next generation of assessments," in *Validation of Score Meaning for the Next Generation of Assessments: The Use of Response Processes*, eds K. Ercikan and J. W. Pellegrino (New York, NY: Routledge), doi: 10.1111/jedm.12256

Ercikan, K., and Por, H. H. (2020). "Comparability in multilingual and multicultural assessment contexts," in *Comparability of Large-Scale Educational Assessments: Issues and Recommendations*, eds A. Berman, E. Haertel, and J. Pellegrino (Washington, DC: National Academy of Education), 205–225. doi: 10.31094/2020/1

Gudmundsson, E. (2009). Guidelines for translating and adapting psychological instruments. *Nord. Psychol.* 61, 29–45. doi: 10.1027/1901-2276.61.2.29

Hambleton, R. K. (2005). "Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures," in *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*, eds R. K. Hambleton, P. F. Merenda, and C. D. Spielberger (Mahwah, NJ: Lawrence Erlbaum), 3–38.

Hambleton, R. K., and Lee, M. K. (2013). "Methods for translating and adapting tests to increase cross-language validity," in *The Oxford Handbook of Child Psychological Assessment*, eds D. H. Saklofske, C. R. Reynolds, and V. L. Schwean (Oxford: Oxford University Press), 172–181. doi: 10.1159/000477727

Hambleton, R. K., and Patsula, L. (1998). Adapting tests for use in multiple languages and cultures. *Soc. Indic. Res.* 45, 153–171. doi: 10.1023/A:1006941729637

Herdman, M., Fox-Rushby, J., and Badia, X. (1998). A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Qual. Life Res.* 7, 323–335. doi: 10.1023/a:1024985930536

Hernández, A., Hidalgo, M. D., Hambleton, R. K., and Gómez-Benito, J. (2020). International Test Commission guidelines for test adaptation: A criterion checklist. *Psicothema* 32, 390–398. doi: 10.7334/psicothema2019.306

Hyytinen, H., Holma, K., Shavelson, R. J., and Lindblom-Ylänne, S. (2014). The complex relationship between students' critical thinking and epistemological beliefs in the context of problem solving. *Frontline Learn. Res.* 6:1–15. doi: 10.14786/flr.v2i4.124

Hyytinen, H., Ursin, J., Silvennoinen, K., Kleemola, K., and Toom, A. (2021). The dynamic relationship between response processes and self-regulation in critical

thinking assessments. *Stud. Educ. Eval.* 71:101090. doi: 10.1016/j.stueduc.2021.101090

International Test Commission (2018). ITC guidelines for translating and adapting tests (Second edition). *Int. J. Test.* 18, 101–134. doi: 10.1080/15305058.2017.1398166

Johnson, R., Penny, J., and Gordon, B. (2009). *Assessing Performance. Designing, Scoring, and Validating Performance Tasks*. New York, NY: Guilford Press.

Kane, M. (2013). The argument-based approach to validation. *Sch. Psychol. Rev.* 42, 448–457. doi: 10.1080/02796015.2013.12087465

Karlgren, K., Lakkala, M., Toom, A., Ilomäki, L., Lahti-Nuuttila, P., and Muukkonen, H. (2020). Assessing the learning of knowledge work competence in higher education – cross-cultural translation and adaptation of the Collaborative Knowledge Practices Questionnaire. *Res. Pap. Educ.* 35, 8–22. doi: 10.1080/02671522.2019.1677752

Kleemola, K., Hyytinen, H., and Toom, A. (2021). Exploring internal structure of a performance-based critical thinking assessment for new students in higher education. *Assess. Eval. High. Educ.* doi: 10.1080/02602938.2021.1946482

Kleemola, K., Hyytinen, H., and Toom, A. (2022, in press). The challenge of position-taking in novice higher education students' argumentative writing. *Front. Educ.* doi: 10.3389/feduc.2022.885987

Leighton, J. P. (2017). *Using Think-Aloud Interviews and Cognitive Labs in Educational Research*. Oxford: Oxford University Press.

Leighton, J. P. (2019). The risk–return trade-off: Performance assessments and cognitive validation of inferences. *Br. J. Educ. Psychol.* 89, 441–455. doi: 10.1111/bjep.12271

Mauranen, A. (1993). Contrastive ESP rhetoric: metatext in Finnish-English economics texts. *Engl. Specif. Purp.* 12, 3–22. doi: 10.1016/0889-4906(93)90024-I

Oliveri, M. E., and Lawless, R. (2018). *The Validity of Inferences from Locally Developed Assessments Administered Globally (Research Report No. RR-18-35)*. Princeton, NJ: Educational Testing Service, doi: 10.1002/ets2.12221

Oranje, A., Gorin, J., Jia, Y., and Kerr, D. (2017). "Collecting, analyzing, and interpreting response time, eye-tracking and log data," in *Validation of Score Meaning Using Examinee Response Processes for the Next Generation of Assessments*, eds K. Ercikan and J. W. Pellegrino (New York, NY: Routledge), 34–44.

Palm, T. (2008). Performance assessment and authentic assessment: a conceptual analysis of the literature. *Pract. Assess. Res. Eval.* 13, 1–11. doi: 10.7275/0qpc-ws45

Ronderos, N., Shavelson, R. J., Holtsch, D., Zlatkin-Troitschanskaia, O., and Solano-Flores, G. (2021). International performance assessment of critical thinking: framework for translation and adaptation. *J. Supranat. Policies Educ.* 13, 62–87. doi: 10.15366/jospoe2021.13.003

Schendel, R., and Tolmie, A. (2017). Beyond translation: adapting a performance-task-based assessment of critical thinking ability for use in Rwanda. *Assess. Eval. High. Educ.* 42, 673–689. doi: 10.1080/02602938.2016.1177484

Shavelson, R. J. (2010). *Measuring College Learning Responsibly: Accountability in a New Era*. Redwood City, CA: Stanford University Press.

Shavelson, R. J., Zlatkin-Troitschanskaia, O., Beck, K., Schmidt, S., and Marino, J. P. (2019). Assessment of university students' critical thinking: Next generation performance assessment. *Int. J. Test.* 19, 337–362. doi: 10.1080/15305058.2018.1543309

Tremblay, K., Lalancette, D., and Roseveare, D. (2012). *Assessment of Higher Education Learning Outcomes AHELO. Feasibility Study Report. Volume 1: Design and Implementation*. Paris: OECD.

Tuononen, T., Hyytinen, H., Kleemola, K., Hailikari, T., Männikkö, I., and Toom, A. (2022). *Systematic review of learning generic skills in higher education - enhancing and impeding factors.*

Ursin, J., Hyytinen, H., and Silvennoinen, K. (eds) (2021). *Assessment of Undergraduate Students' Generic Skills in Finland: Findings of the Kappas! Project*. Helsinki: Ministry of Education and Culture.

van Someren, M. W., Barnard, Y. F., and Sandberg, J. A. C. (1994). *The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes*. London: Academic Press.

van Widenfelt, B. M., Treffers, P. D. A., de Beurs, E., Siebelink, B. M., and Koudijs, E. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clin. Child Family Psychol. Rev.* 8, 135–147. doi: 10.1007/s10567-005-4752-1

Zahner, D., and Ciolfi, A. (2018). "International comparison of a performance-based assessment in higher education," in *Assessment of Learning Outcomes in Higher Education: Cross-national Comparisons and Perspectives*, eds O. Zlatkin-Troitschanskaia, M. Toepper, H. A. Pant, C. Lautenbach, and C. Kuhn (Berlin: Springer International Publishing), 215–244. doi: 10.1007/978-3-319-74338-7

Zlatkin-Troitschanskaia, O., Shavelson, R. J., Schmidt, S., and Beck, K. (2019). On the complementarity of holistic and analytic approaches to performance assessment scoring. *Br. J. Educ. Psychol.* 89, 468–484. doi: 10.1111/bjep.12286

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Education

HYPOTHESIS AND THEORY
published: 19 May 2022
doi: 10.3389/feduc.2022.883370

Relational Reasoning in Tertiary Education: What Is Its Value and How Can It Be Assessed and Trained?

Patricia A. Alexander†, Nancy Gans† and Alina J. K. Maki†*

Department of Human Development and Quantitative Methodology, University of Maryland, College Park, College Park, MD, United States

The goal of this article is to demonstrate the value of incorporating relational reasoning assessment and training in tertiary education. To accomplish this, the authors organize the article into three sections. The first section overviews the nature of relational reasoning, and its different forms, developmental trajectories, and assessment. How relational reasoning predicts performance in various academic domains and fields of practice is also considered. The second section focuses on the role that relational reasoning plays in the scientific domains that are foundational to tertiary education and professional practice—the natural, social, applied, and formal sciences. In the final section, the authors describe an ongoing design experiment in which relational reasoning assessment and training are integrated into a university course.

Keywords: relational reasoning, analogy, anomaly, antinomy, antithesis, tertiary education

OPEN ACCESS

Edited by:
Heidi Hyytinen,
University of Helsinki, Finland

Reviewed by:
Auli Toom,
University of Helsinki, Finland
Zanda Rubene,
University of Latvia, Latvia

****Correspondence:***
Patricia A. Alexander
palexand@umd.edu

†ORCID:
Patricia A. Alexander
orcid.org/0000-0001-7060-2582
Nancy Gans
orcid.org/0000-0003-0495-0583
Alina J. K. Maki
orcid.org/0000-0002-1277-3540

Specialty section:
This article was submitted to Higher Education, a section of the journal Frontiers in Education

Received: *25 February 2022*
Accepted: *29 April 2022*
Published: *19 May 2022*

Citation:
Alexander PA, Gans N and Maki AJK (2022) Relational Reasoning in Tertiary Education: What Is Its Value and How Can It Be Assessed and Trained? Front. Educ. 7:883370. doi: 10.3389/feduc.2022.883370

The purpose of this article is to demonstrate the theoretical, empirical, and practical value of assessing tertiary students' ability to reason relationally by means of a novel and fluid measure (Diamond, 2013). Further, we will draw on the extant literature and an ongoing classroom-based design experiment to illustrate how the assessment of relational reasoning and its subsequent training in the context of a university course can serve multiple purposes. Specifically, we will describe how the administration of a generic and fluid measure can result in a profile of tertiary students' analogical, anomalous, antinomous, and antithetical reasoning capabilities. Moreover, with their profile as a starting point, tertiary students can be given explicit instruction in the forms of relational reasoning and the underlying cognitive processes they require. They can also be shown how their ability to reason relationally can bolster their academic performance in specific domains and be invaluable to their future career success. As a starting point for this discussion, we explain what it means to reason relationally and the distinct forms of this cognitive ability that have been identified. We also describe how relational reasoning develops and what this capability predicts in learning and achievement.

AN OVERVIEW OF RELATIONAL REASONING

Relational Reasoning Defined

In the neuroscience, developmental, cognitive science, and psychological literature, *relational reasoning* is defined as the ability to recognize complex, meaningful patterns within bodies of seemingly unrelated information (Spearman, 1927; Cattell, 1940; Singley and Bunge, 2014). The student taking classic English literature who sees parallels in Shakespeare's *Julius Caesar* and contemporary political intrigues; the physics major who grasps the underlying association between

the calculus and physics; or medical residents who conclude that the case they are diagnosing cannot be classified as juvenile diabetes are all exhibiting relational reasoning. Gentner and Gentner (1983) and Dumas et al. (2014) distinguish the higher-order patterns associated with relational reasoning from simpler, linear patterns like number sequences (e.g., 3, 8, 13, 18 ____) by noting that these higher-order patterns must be based on *relations-among-relations* (e.g., 3: 21: 4: ____). In essence, the identification of multiple associations must occur for any of the aforementioned "insights" to emerge.

Alexander and Baggetta (2014) also differentiate between *relational thinking* and *relational reasoning*, because thinking relationally may transpire without awareness or intentionality on the part of the individual. When a young child intuitively recognizes that an unfamiliar animal (chihuahua) is, in fact, a dog, that child has recognized a link between some new creature and the idea of "dog." Yet that realization was more or less unconscious and involved little cognitive effort. Relational reasoning, in contrast, requires the effortful and intentional harnessing of information that can result in an intricate association between objects, ideas, or events that extend and deepen understanding (Alexander and Baggetta, 2014). Scientists puzzling over whether Pluto was a planet or not had to wrestle with the presence or absence of determinative attributes before concluding that this astronomical body should be classified as a "dwarf planet." These scientists' intentions and the level of effort it took to appropriately classify Pluto clearly position this example as relational reasoning.

Although our focus in this article is squarely on relational reasoning, we want to make clear that relational thinking is also essential for human functioning (Alexander, 2019). Further, the more intuitive or System 1 thinking works in concert with the more effortful System 2 processes implicated in relational reasoning (Stanovich, 2010). What is core to relational thinking *and* relational reasoning is their dependency on the perception of and attention to similarities and dissimilarities among objects, ideas, or events—some subtle and some dramatic; some concrete and some abstract (James, 1893; Cattell, 1940). As we will see, the nature of these similarities and dissimilarities is what defines the forms of relational reasoning.

Relational Reasoning Classified and Categorized

Within the neuroscience, developmental, cognitive science, and psychological literature, relational reasoning is conceptualized in a fairly consistent manner that corresponds to the definition we proffered earlier. Differences within those literature including how relational reasoning is positioned within the neurological architecture, its development over the lifespan, and its operationalization in empirical research have direct relevance to the assessment and training of relational reasoning in higher education (Dumas et al., 2013; Alexander, 2016). For example, there are those who regard relational reasoning ability as a higher-order executive function that continues to develop into early adulthood with the myelination of the prefrontal cortex, a region associated with complex problem solving and decision making (Dumontheil et al., 2010; Krawczyk et al., 2011; Diamond, 2013). Therefore, based on their neurophysiological development, those enrolled in colleges and universities seem well positioned to benefit from relational reasoning assessment and training.

There is also some dispute in the literature over the degree to which relational reasoning ability is affected by social and educational experiences, as well as by neurophysiological changes (Carlson, 2009; Bunge and Leib, 2020). Decoupling neurological and biological factors from what is concomitantly occurring socially and educationally continually proves challenging. The lack of a single measure of relational reasoning that can be reliably used with differently aged participants further complicates this matter. Recently, Chae and Alexander (2021a) found themselves in a unique position to shed some light on this conundrum when they were able to test the relational reasoning capabilities of three groups of South Koreans who varied significantly in age and in the course of their formal education. These researchers administered the same fluid ability measure, the *Test of Relational Reasoning-Junior* (TORRjr; Alexander and The Disciplined Reading and Learning Research Laboratory [DRLRL], 2018), to young adolescents in school, older adults (ages 50+) whose schooling happened in a typical timeframe, and older adults (ages 45+) now completing their middle-school or high-school education.

These researchers hypothesized that if social and educational factors are not significant forces in relational reasoning development, then the two groups of older adults should perform comparably. On the other hand, if social and educational experiences do play an important role in relational reasoning development, then the older adults who did not attend school until decades later should perform the worst of the three groups—which is precisely what Chae and Alexander (2021a) found. This outcome implies that college students who are not only nearing their neurophysiological prime but who also are furthering their education should have ample opportunities to reason independently and collaboratively.

Relational Reasoning Forms

Despite the largely shared conceptualization of relational reasoning that populates the literature, the manner in which this cognitive capability has been operationalized is far more contentious (Alexander et al., 2016a; Baggetta and Alexander, 2016). For the most part, the operationalization of relational reasoning has focused only on analogical reasoning or the discernment of patterns of similarity (Dumas et al., 2013). Similarly, the assessment of relational reasoning has largely been relegated to one measure, the Raven's Matrices (Raven, 1941), a fluid ability measure composed entirely of figural analogy problems. Consequently, minimal data have been gathered on salient patterns that also place differential emphasis on dissimilarities.

To address this significant shortcoming, Alexander and The Disciplined Reading and Learning Research Laboratory [DRLRL] (2012) and Alexander et al. (2016a) set out to devise a fluid measure that gauged individuals' ability to recognize multiple relational forms. Drawing on cross-disciplinary literature, including philosophy, mathematics, logic, and intellectual

assessment (e.g., James, 1893; Cattell, 1940; Russell and Lackey, 1973), four relational forms were identified: analogies (similarity), anomalies (aberrance), antimonies (exclusivity), and antitheses (opposition). With these forms identified, the process began to construct novel, figural items that could reliably and validly capture those relational manifestations. The outcome of this multi-year effort was the *Test of Relational Reasoning* (TORR; Alexander and The Disciplined Reading and Learning Research Laboratory [DRLRL], 2014). The TORR is a psychometrically sound standardized instrument suitable for adolescents and adults that has been shown to be invariant for females and males from diverse cultural and ethnic backgrounds (Dumas and Alexander, 2016, 2018).

The TORR consists of four 8-item scales, each representing one form (see **Figure 1**). The *analogical reasoning* scale, as with the Ravens, is composed of figural problems displayed in a 3 × 3 matrix. Respondents are directed to find the option that conforms to the pattern indicated. The *anomalous reasoning* problems the identification of the figure within a given set that deviates from the others. *Antinomous reasoning* is defined as the ability to recognize a true binary distinction, where ideas, objects, or events either fit within a specific category or not (e.g., living versus non-living objects). For each *antinomous reasoning* item, respondents are shown a set of related figures. They are then directed to find the set from among the options that can have no figure in common with the given set. The final scale of the TORR assesses antithetical reasoning. While comparisons made on the antinomous reasoning scale represent binary or dichotomous distinctions, those on the *antithetical reasoning* scale capture opposing but continuous differences (e.g., tall versus short). The antithesis items, therefore, depict certain features of a given figure (x) being switched to create a new figure (y). Respondents are asked to select the option that represents the reverse of the process conveyed in the given problem.

Developmental Trends by Form

The value of assessing multiple forms of relational reasoning rather than relying solely on the measurement of analogical reasoning extends beyond achieving a more accurate conceptual-operational mapping of this foundational ability. Particularly as it pertains to the assessment and training of tertiary students, there is evidence that these forms reveal important trends by age and by domain of study. While we will reserve discussion of relational reasoning forms and academic domains for later in this article, here we want to highlight findings that speak to performance trends by form for samples that run the gamut from young children to senior citizens. The studies to which we will refer all assessed analogical, anomalous, antinomous, and antithetical reasoning, albeit by means of diverse methodologies.

For instance, Jablansky et al. (2016, 2020) analyzed longitudinal data on relational reasoning collected from New Zealand students in kindergarten through Grade 12. The children and youth were participants in a project designed to promote their technological literacy. What Jablansky et al. (2016, 2020) documented was a salient shift in the reasoning forms that were more or less prevalent in the language of younger and older students as they discussed the more and less familiar tools. Specifically, even though students at each grade level engaged in analogical, anomalous, antinomous, and antithetical reasoning to some extent, younger students relied primarily on analogical and anomalous reasoning. The older students, conversely, made greater use of antinomous and antithetical reasoning when analyzing the two technological tools. Interestingly, when presented with a familiar tool, younger students were able to show some ability to reason antinomously and antithetically, while older students found far less reason to reason analogically or anomalously (Jablansky et al., 2016, 2020). These findings suggest that while reasoning relationally may require some relevant background knowledge if experiences are too familiar or routine there may be limited impetus to reason relationally.

More recently, Chae and Alexander (2021b) and Zhao et al. (2021) also had the opportunity to examine relational reasoning performance in students at different grade levels. For both investigations, students completed the TORRjr translated into Korean or Chinese, respectively. These studies showed shifts in reasoning performance by form that paralleled the longitudinal study by Jablansky et al. (2020). Younger students performed higher for analogical and anomalous reasoning than for antinomous and antithetical reasoning, whereas older students performed higher for antinomous and antithetical reasoning.

Finally, the investigation by Chae and Alexander (2021a) that looked at the relational reasoning among adolescents and typically schooled and atypically schooled older adults reinforced the pattern seen across this collection of studies by suggesting that analogies may represent the easiest mode of relational pattern for participants of all ages. Even the atypically schooled adults in the Chae and Alexander (2021a) study, whose overall performance was significantly below that of the young adolescents and typically schooled adults, scored near the midpoint on the analogy scale. Further, the young adolescents and typically schooled adults scored comparably on the TORRjr. However, the current students performed significantly better than these older adults on the antinomy scale, while the reverse was true for the anomaly scale where the older, typically schooled adult prevailed.

Overall, these developmental trends would suggest that, based on their age and educational background, tertiary students would be expected to possess the ability to reason analogically, anomalously, antinomously, and antithetical to at least an adequate degree. Of course, we would expect that there would be variability in performance across the scales, with somewhat greater difficulty exhibited for the antinomous and antithetical items over the analogy and anomaly items prior to any training.

IMPORTANCE OF RELATIONAL REASONING IN TERTIARY EDUCATION

As we indicated in the prior overview of relational reasoning, students populating colleges and universities are well positioned to benefit from the assessment and training of this essential cognitive ability. For one thing, the myelination of the prefrontal cortex, the brain region central to problem-solving, decision-making, and self-regulation, is nearing an end (Dumontheil et al., 2010; Krawczyk, 2012). For another, these tertiary students

are expanding their knowledge and are routinely engaged in academic and social interactions that may require them to reason relationally (Carlson, 2009; Bunge and Leib, 2020). Moreover, during this period of their lives, college students are determining which career paths to pursue and then preparing for those careers. Relational reasoning assessment and training can be invaluable for examining the higher-order thinking abilities of these students (Alexander, 2016). Their scores on the TORR can serve as an initial profile of their relational reasoning strengths or areas of need. Further, students' specific reasoning profiles can be compared to those who are already studying or practicing in their chosen fields (Dumas et al., 2014; Fountain, 2016; Jablansky, 2020).

It is important to keep in mind that relational reasoning, as with any cognitive or intellectual capacity, is malleable (Hsu et al., 2014; Alexander et al., 2016b). Thus, performance for the different forms of relational analysis can be improved through training, relevant experiences, and repeated practice. Consequently, as we will discuss in the final section of this article, it is only a first step to assess the relational reasoning abilities of college and university students. It is quite another to use those resulting data to help these students hone their abilities to reason analogically, anomalously, antinomously, and antithetically.

Relational Reasoning in Domains and Disciplines

Those engaged in relational reasoning research have investigated the degree to which this ability predicts performance in specific academic domains (mathematics or literacy) or fields of practice (medical diagnosis or engineering design). Within higher education, academic domains correspond roughly to fields of study around which content and learning experiences are organized and delivered. For example, mathematics curricula and courses are often sequenced to capture the increasing complexity of underlying concepts and procedures (Schmidt et al., 2005). Thus, precalculus is taken before calculus, and a statistics course on general linear models likely precedes a course on multivariate mixed models. Other domains such as history or literature may have content structured chronologically or by genres (Orrill and Shapiro, 2005).

There can be noticeable differences in how students experience the content that has implications for relational reasoning, as well. In some fields like mechanical engineering or architecture, where professionals often work in teams, university students may work on projects with other students under the guidance of knowledgeable instructors (Dumas et al., 2016; Kavousi et al., 2020). Further, laboratory experiences or simulations can be integral to domains where tertiary students learn specific techniques or procedures viewed as central to a future career, as in nursing training or chemistry. In contrast, other fields such as journalism or history may be more focused on individual learning and production with fewer structured group collaborations.

Also, the kinds of prototypic problems that populate courses from different domains can differ markedly. For instance, there may be more reliance on memorization, recall, and well-structured problems (i.e., clearly has a right/wrong answer) for students in the biological sciences, and more interpretative and evaluative tasks (i.e., ill-structured problems) for students in the social sciences (Alexander, 2006; Reed, 2016). Such variations in the structure or delivery of content can mean that certain forms of relational reasoning occur more often than others or unfold in a different pattern (Dumas et al., 2014; Jablansky, 2020). However, it can be presumed that all forms of relational reasoning have a role to play in learning and performance within tertiary education. Here we will look at the studies that have explored the association between relational reasoning overall and by form in selected fields of study and professional practice.

Given the enormity of domains that tertiary students can pursue, we have chosen to organize this brief exploration around four branches of science; the *natural sciences* that deal with nature in some fashion; the *social sciences*, which focus on people, society, and culture; the *applied sciences* such as engineering, statistics, architecture, and medicine; and the *formal sciences* that include theoretical mathematics, logic, philosophy, and theoretical linguistics. Before delving into where relational reasoning comes into play within each of these areas of study, we want to forward an important caveat. Specifically, the four categories of domains we identify herein and the variety of studies that aligned with each of those areas represent only one possible configuration that could be considered. There are innumerable classifications of academic domains that have been proposed and even the courses identified within those organizational schemes can vary. Our decision to focus on these four domain areas was driven, in part, by our desire to keep the comparisons and contrasts among the areas as simple as possible, while still allowing for the utility of relational reasoning to be adequately described.

Natural Sciences

As noted, natural sciences deal with the physical world and all that exists therein. Those studying or working in the natural sciences are generally concerned with carefully observing, accurately describing, systematically classifying, or predicting phenomena (Harris, 2014). The execution of these core processes necessitates that those studying or working in the natural sciences rely on empirical evidence derived through observation or direct experience that can be substantiated or disproven through experimentation that adheres to the scientific method. There are many ways in which the forms of relational reasoning are integral to learning and performance in the natural sciences and, thus, to tertiary students studying in these fields.

For example, in his treatise on analogical arguments, philosopher Bartha (2019) contends that analogies have long been a critical feature of scientific reasoning and a contributor to scientific discoveries. To support this contention, Bartha quotes Joseph Priestley, a renowned 18th century expert in chemistry and electricity who is created with the discovery of oxygen, who claimed that "analogy is our best guide in all philosophical investigations; and all discoveries, which were not made by mere accident, have been made by the help of it" (Priestley, 1767, p. 443–444). Similar claims for the power of analogical reasoning populate the history of science (Hofstadter, 1979; Dunbar and Blanchette, 2001; Gentner, 2002).

Despite the value that analogical reasoning holds in the natural sciences, it cannot stand alone to explain what must transpire when observing, describing, classifying, or predicting phenomena in this domain. As observations are made and empirical evidence gathered, scientists must attend to critical dissimilarities that emerge, as well. For one, anomalous reasoning is an essential tool for scientists or students in the natural sciences, because it results in the perception or identification of salient discrepancies or deviations from the expected or typical (Chinn and Brewer, 1993; Chinn and Malhotra, 2002). The presence of anomalies can be difficult to explain, and they can even bring tentative hypotheses or accepted theories into question. Under certain circumstances, anomalies can become the catalyst for alternative hypotheses or theories. Whether anomalies are treated as noise in a data set or give rise to new hypotheses or theories, may be dependent on scientists' ability to construct a cogent and compelling explanation for their existence (Lightman and Gingerich, 1992).

The act of labeling, classifying, or categorizing natural phenomena likewise demands more than noticing and cataloging similarities. Without analysis of meaningful differences within and between phenomena, their true nature would not be captured. In some instances, those differences take the form of scales that capture levels of a particular feature such as scales for water hardness or softness, wind speed, or soil types. Although the data on which they are based are continuous, these scales are often banded and labeled, which involves antinomous reasoning to determine distinct groupings. In other instances, the categories are meant to capture ontological distinctions that are considered discrete, as with the Hertzsprung–Russell system for classifying stars or the five biological kingdoms into which all living things are sorted. The precise categories that are formed in these instances arise through antinomous reasoning.

Social Sciences

The primary distinction between the natural sciences and social sciences—the study of nature versus humans—translates into significant differences in what is studied and how studies are undertaken (Nowotny, 2005). Consequently, the forms of relational reasoning can manifest in varied ways and to varying degrees in each domain. For one, much of what concerns social scientists are not directly measurable as is the case in natural sciences (Borsboom and Mellenbergh, 2002). There are certainly explicit behaviors of individuals and groups that can be documented and analyzed. Nonetheless, the questions often posed by social scientists are about the unseen or underlying forces, factors, or conditions that give rise to those behaviors. As a result, there is potentially much more that must be inferred from gathered data. The techniques and measures that must be created as well as the data analytic approaches can be quantitative, qualitative, or some combination of both (Mertens, 2019).

In light of this characterization of social sciences, how do the forms of relational reasoning come into play? We can look back at the opening discussion about the nature of relational reasoning for guidance on this matter. For instance, all manner of human reasoning lies buried in the mind of an individual or some societal group (James, 1893). Therefore, discerning patterns of similarity or dissimilarity requires social scientists to be attentive to any external markers that can suggest what is transpiring consciously or unconsciously within the individual, group, or society (Harris, 2014). The words individuals or groups utter, the behaviors they display, the decisions or choices they make, and even biophysiological indicators can prove invaluable to recognizing meaningful patterns (West et al., 2008; Kaplan and Berman, 2010).

When it comes to the forms of relational reasoning, analogical reasoning allows those in these fields to recognize important consistencies or commonalities across individuals, societies, or cultures. Developmental theories, for example, are predicated on assumed shared characteristics among individuals of similar age (Halford, 1992). Similarly, socioeconomic models look for predictable outcomes based on the level of wealth or poverty experienced at the level of the individual, group, or society (Hackman and Farah, 2009). Thus, what analogical reasoning can offer in such instances is a framework or starting point that captures common or typical conditions, as we discussed in terms of neurophysiological maturation and relational reasoning capability. As in the natural sciences, analogical reasoning is an important process for tertiary students pursuing careers in the social sciences to hone.

Of course, unearthing similarities in people's thoughts, actions, or experiences is only a portion of what social scientists seek to investigate. They are also invested in understanding how individuals, groups, or societies differ from one another and the nature of those differences (Fischer and Silvern, 1985; Beattie, 2002). Fields like special education, clinical psychology, criminology, expertise, and many others are focused on those who exhibit ways of thinking and acting that deviate in non-trivial ways from what is regarded as the norm (Samuel and Widiger, 2008; Sullivan and Bal, 2013). In effect, there is something that is perceived as anomalous about these individuals or groups that social scientists may set out to explain through their research and perhaps to ameliorate or amplify those differences through treatments or interventions.

Further, while classifying and categorizing occurs in the social sciences as in the natural sciences, there is a major distinction between these fields that must be appreciated. Specifically, because of the nature of the data or evidence that can be gathered in the natural sciences, there is the potential to uncover true dichotomies (Alexander, in press a). Through antinomous reasoning, compelling, and seemingly incontrovertible evidence, discrete categories such as living and non-living matter or animals and plants can emerge in the natural sciences. In the social sciences, however, ontological distinctions of this type are rare. For most of the social sciences, there are more designations that are generated on the basis of continuous or variable characterization (Lehtinen, 2012; Alexander, in press a). When sociologists consider political, socioeconomic, racial, and class distinctions, there are no unambiguous categories that result, and many distinctions between individuals, groups, or societies shift over time. For example, bigenderism that once prevailed has given way to more fluid gender distinctions—more antithetical than antinomous in form (Gilbert, 2009). Therefore, it seems critical that those engaged in describing and

classifying humans for the purpose of the study should recognize that antithetical reasoning is more central to their fields than antinomous reasoning.

Applied Sciences

Within tertiary education and professional practice, there are fields that are devoted to the application of knowledge and procedures garnered from research in the natural and social sciences to critical real-world problems. This focus on the use of existing knowledge and procedures to deal with pressing problems—be they structural, aesthetic, moral, physical, educational, social, or cultural—is why these domains are referred to as *applied sciences*. Fields that fall in this category include medicine and health, engineering, teaching, counseling and clinical psychology, computer science, and applied statistics/mathematics. As this litany suggests, domains that are regarded as applied can be extremely challenging and can require years of preparation and practice to master (Patel et al., 1999).

Among the essentials for performing well in any applied science are a breadth and depth of domain-specific knowledge and procedural capabilities (Alexander, 2003, in press b). For example, a breadth of engineering knowledge might include what one knows about the many fields of engineering, whereas a depth of civil engineering knowledge might relate to what one knows specifically about bridge construction. What goes hand in hand with such domain-specific competencies are more general competencies such as the ability to reason, think critically, make sound decisions, and collaborate with others (Alexander, 2004, in press b). Those in the applied sciences must also be able to recognize, analyze, and classify the nature of problems they will likely encounter, not only in terms of their surface features but also their underlying structure (Albanese and Dast, 2014). Moreover, those in these applied fields have to envision viable techniques or approaches to addressing those problems in order to be successful (Dumas et al., 2016). Finally, among the critical competencies associated with the applied sciences are strategies for monitoring the situation and evaluating the effectiveness of the actions being taken, which includes a judgment of one's own performance (McConnell et al., 2012).

With this general picture of the applied sciences in place, we will now turn our attention to the significance of relational reasoning in these fields. Within the applied sciences, in particular, there have been numerous empirical studies of relational reasoning's contributions to professional performance, especially in medicine, health, and engineering.

Medicine and Health

Research studies have shown that relational reasoning provides professionals in medicine and health with the necessary tools to solve complex problems in their field. For example, Dumas et al. (2014) captured many instances of relational reasoning that punctuated the exchanges between an expert attending physician and the residents he was mentoring. Those exchanges occurred as the residents were analyzing details of their patients' conditions in order to make accurate diagnoses. Dumas et al. (2014) found that while all four forms of relational reasoning were present in the doctors' real-time problem-solving, those forms unfolded in a patterned way. For instance, when the residents were first presenting their cases, they relied heavily on anomalous reasoning to delineate patients' atypical symptoms. This delineation eventually gave way to analogical reasoning, as the residents began to speculate on what conditions those symptoms may suggest. At this point in the process, antinomous reasoning was introduced as the attending physician or one of the residents noted that some essential features in the case made the proposed diagnosis untenable. This reasoning cycle repeated until an acceptable diagnosis was reached.

Critical thinking and decision-making within health professions are imperative as doctors and nurses engage in diagnosing and treating patients. A study of maternity nursing students and practicing nurses by Fountain (2016), for instance, revealed that relational reasoning was a significant contributor to their critical thinking, beyond domain knowledge, individual interest, and years of experience. Even though relational reasoning was a significant predictor of critical thinking skills, Fountain (2016) found that there were no significant differences in the four forms of relational reasoning between more or less experienced nurses. This finding was interpreted by the researcher as evidence that professional experience alone was not sufficient to advance relational reasoning abilities among the nurses.

Engineering

Engineering is a frequently studied domain by relational reasoning researchers (Dumas et al., 2016; Jablansky et al., 2020). What makes this domain appealing for these researchers are the nature of the problems and the fact that engineers often work in teams when designing and carrying out projects. Engineering provides unique opportunities to study relational reasoning, because students and practicing professionals are required to consider the feasibility of designs based on established principles in mathematics and physics. Further, there are frameworks such as the Theory of Inventive Problem Solving (TRIZ; Altshuller and Shapiro, 1956) that students and those practicing in the field can use to assess the creativity and viability of potential designs.

Interestingly, Dumas and Schmidt (2015) and Dumas et al. (2016) found that after engaging in the TRIZ intervention, students were likely to produce fewer but more innovative design ideas. As it pertains to the focus of this article, these researchers also determined that students who were the most creative both before and after the TRIZ intervention scored high on the TORR. Further, the students who were more receptive to the TRIZ intervention were those who were strong in antinomous reasoning. It would seem that for engineers, there needs to be a determination as to whether or not a project design will function or not. There is no room for error when calculating whether a bridge will stand or the foundation of a building will hold. Antinomous reasoning, along with analogical, anomalous, and antithetical reasoning, appears to be essential tools that engineers must apply when working in their field.

Finally, engineering like medicine is often a collaborative process and, thus, the relational reasoning that occurs among group members can greatly affect both the problem-solving

process and the resulting outcome. Jablansky (2020) followed several teams of senior mechanical engineering university students tasked with designing a creative but highly functional product for an existing problem. She videotaped these students' weekly meetings to document *how* members of the group contributed and *when*, including but not limited to how they applied relational reasoning to achieve their goals. Jablansky (2020) found evidence of relational reasoning patterns that differed not only by person but also by the task that they set out to complete each week. These reasoning patterns were mapped onto data representing the social and regulatory dynamics within teams that afforded a rich picture of the engineering design process. Studies of this nature within any collaborative undertaking among tertiary students in applied sciences can be invaluable to those overseeing course content or guiding the learning of these students.

Formal Sciences

The formal sciences may be a less familiar grouping of domains than the natural, social, or applied sciences, but are a particularly fast-growing area, especially because of technological advancements in the ability to create complex models and systems (Treur, 2021). Philosophy, logic, theoretical mathematics, systems theory, theoretical computer science, artificial intelligence, information theory or informatics, game theory, computational linguistics, and theoretical linguistics are among the academic domains that fit within the formal sciences. What distinguishes the formal sciences from the prior domain groups we overviewed are their aims and methods. In effect, the non-formal sciences involve gathering evidence about nature or about people to better describe what is observed or documented. Such evidence or data can then be subjected to analysis in order to support or reject researchers' hypotheses. For the applied sciences, the knowledge gained from studies in the natural and social sciences is put to work on real-world problems. In contrast, the aim of the formal sciences is to generate abstract macro-models or theoretical systems meant to explain the phenomena investigated by natural and social scientists or the outcomes observed by applied scientists (Löwe, 2002).

Rather than the tools of empiricism or the scientific method, students and professionals in formal sciences employ logic, reasoning, and symbolic systems of mathematics and language to formulate and test explanatory models or systems. According to de Laplante (2006), the formal sciences aid the natural and social sciences by providing information about the structures used to describe the physical world, and about what inferences may be made about these structures. For that reason, there is a certain domain-generality to the formal sciences since the models and systems formulated in this arena often are applicable to natural, social, and applied sciences.

Further, the methods through which the formal sciences support or disprove macro-models and can be quite varied and complex. Those methods can take the form of formal mathematical proofs, critical analysis of any underlying principles or axioms, or the explanatory power of the proposed models or systems. Cramer and Dauphin (2020) also suggested that structured argumentation, which is a mode of scientific argumentation, is an invaluable tool applied by those trying to evaluate the viability or credibility of proposed models and systems. These structured arguments are tantamount to high-level debates focused on premises, procedures, or principles represented in the proposed models or systems.

So, where do relational reasoning and its four manifestations fit within the formal sciences? Because of the weight placed on abstract and complex models or systems that are represented in mathematical or linguistic symbols, relational reasoning is critical. For instance, one crucial ability that requires analogical reasoning is envisioning how elements or components that may exist separately in nature or society display similarities that allow for their convergence into more macro-level theoretical models (Tsoukas, 1993). Similarly, conceiving of a theoretical model or system likely requires individuals to recognize how patterns that repeat in nature (e.g., snowflakes or leaves) or in human systems (e.g., circulatory system or brain cortex) at one level of generality may iterate at a higher or lower level.

Those who are familiar with Mandlebrot's (1982) theory on fractal geometry in nature will understand the aforementioned reference to iterations. An iteration is a recognizable version of a pattern without being exactly the same—not a perfect replication (Bringsjord et al., 2017). For instance, we can recognize a snowflake by its features, while understanding that no two snowflakes are exactly the same. This reminds us that perceiving differences in the formal sciences remains as important as discerning similarities. Those working in the formal sciences must be able to capture the theorized or modeled patterns in natural or human systems linguistically, mathematically, or in computer codes. This ability to capture phenomena symbolically again brings analogical reasoning into play since there is an essential association between the symbolic notions and the phenomena they are depicting (Hummel et al., 2014).

Of course, articulating a theoretical model or system is not the end point of the formal science. Once articulated, these hypothetical or abstracted models or systems must be tested, argued, or scrutinized. This process of proving or disproving a resulting model or theory will call upon analogical, antinomous, antithetical, and antinomous reasoning. Analogical reasoning is needed to map the underlying similarities of the individual model components, while anomalous reasoning is essential to demonstrate that some aberrant iteration still fits with the established system. Antithetical reasoning is used to mark the boundaries of the features that define the model or system, whereas antinomous reasoning is involved when some identified instance fails to fit within the parameters of a proposed system. If such a fundamental failure is identified, some variation of the proposed system must be derived, or an alternative must be considered.

To this point in the article, we have provided a general overview of relational reasoning and an explication of its specific forms. We also described how relational reasoning and its particular manifestations—analogical, anomalous, antinomous, and antithetical—undergird learning and performance in a range of academic domains and professional practices. To support that argument, we first organized the discussion around four clusters of fields found within institutions of higher education: the natural

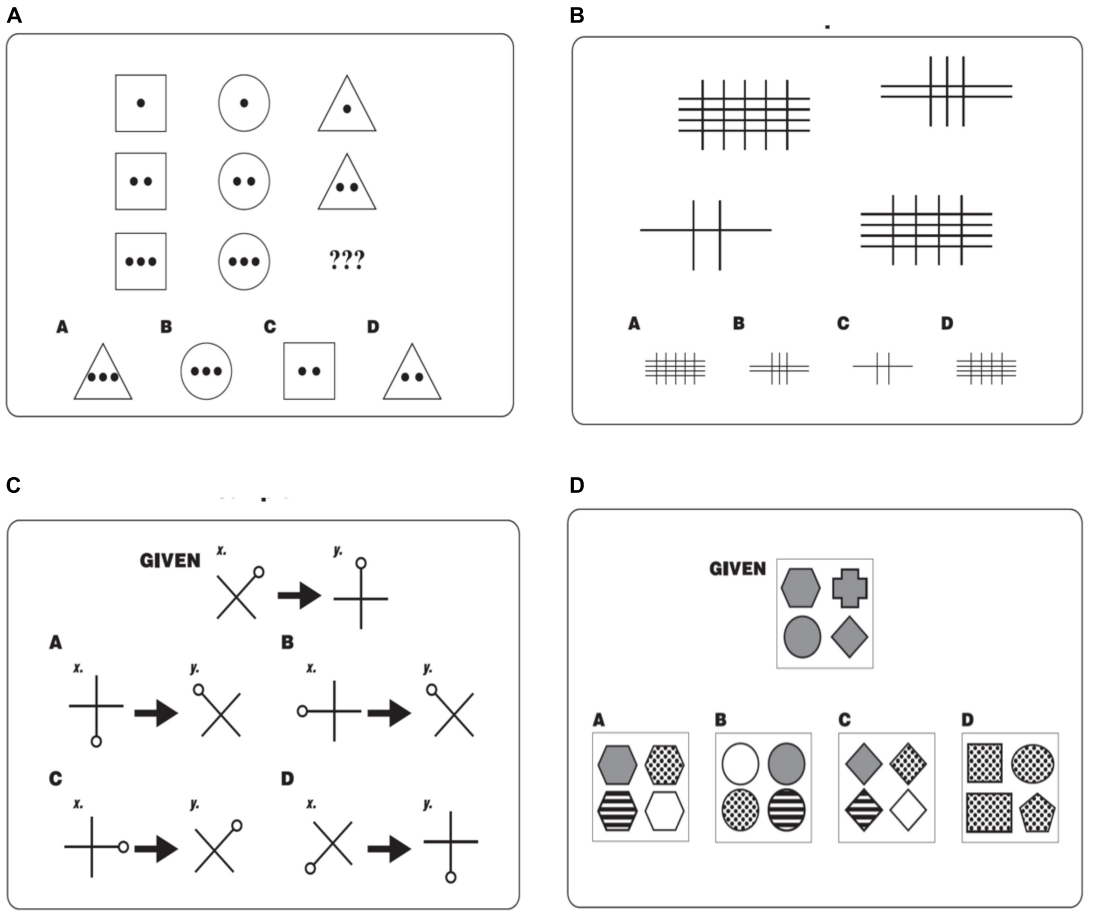


FIGURE 1 | Sample items from the analogy **(A)**, anomaly **(B)**, antithesis **(C)**, and antinomy **(D)** scales of the test of relational reasoning.

sciences, social sciences, applied sciences, and formal sciences. Then, we offered a brief explanation of what distinguishes each of these clusters and described the role that relational reasoning plays in each. In the remaining section, we want to move more abstracted discourse on relational reasoning into a real-world context—a university course in which relational reasoning it assessed and trained.

EMBEDDING RELATIONAL REASONING IN HIGHER EDUCATION: AN ILLUSTRATIVE CASE

Learning How to Learn (LHL) is a general education course designed and taught by the first author and where we have embedded relational reasoning. This course was expressly developed to (a) improve tertiary students' understanding of the complexity and processes of learning and (b) prepare them for the diverse professions they will enter upon graduation. In **Table 1**, we overview the three phases of our instructional procedure. Specifically, we describe how students' relational reasoning abilities are initially assessed (Phase 1); how they are then taught about relational reasoning and its underlying processes (Phase 2); and how this new conceptual and procedural understanding becomes an anchoring point for subsequent instruction in crucial learning topics such as transfer, critical reading, and quality discussion (Phase 3).

Phase 1: Assessment

The ability to forge meaningful relations within any information stream (i.e., relational reasoning) occurs in any medium in which information can be conveyed—words, pictures, sounds, figures, or numbers (Dumas et al., 2013). When individuals notice that a musical sequence in a composition reappears in a different key or at a different tempo; when certain themes in a painting can be identified in an alternative art form; or when researchers recognize outliers in their dataset, relational reasoning is demonstrated (Hofstadter, 1979; Loughlin et al., 2015). Over the past decade, Alexander and colleagues have created valid and reliable measures of relational reasoning that not only consist of figural representations but are also composed of sentences (*Verbal Test of Relational Reasoning* or vTORR; Alexander et al., 2016c) and single words (*Relational Reasoning with Words* or R^2W^2; Zhao and Alexander, 2022).

TABLE 1 | Embedding relational reasoning within a higher-education course.

Relational reasoning training procedure	
Instructional events	**Details**
Phase 1: assessment	
Students' relational reasoning capabilities assessed	TORR is administered to gauge students' ability to reason analogically, anomalously, antinomously, and antithetically Fluid measure used to limit influence of students' background knowledge or experiences
Students receive TORR results	Students automatically receive performance data upon test completion Their total score is reported by raw score and by relational reasoning quotient (M = 100, SD = 15) Raw scores for each of the scales (M = 4) are also reported to form students' reasoning profiles No additional feedback on correctness or incorrectness of responses by individual items is given
Phase 2: explication	
Nature and importance of relational reasoning explained	Relational reasoning is defined and its essential nature to learning and performance is overviewed
Four forms are differentiated and illustrated	Analogy, anomaly, antinomy, and antithesis are compared and contrasted using various examples
Underlying component processes are introduced and practiced	Componential processes of encoding, inferring, mapping, and applying are explained and used to solve analogy problems These componential processes then used with anomaly, antinomy, and antithesis sample problems
Phase 3: extension and transfer	
Domain-specific exploration	Students meet in major or disciplinary groups to identify the role of analogical, anomalous, antinomous, and antithetical reasoning in their fields of study
Multiple-document research	Students carry out research on a controversial topic using multiple documents and prepare an argumentative essay that integrates content across the identified sources After completing this task, students engage in quality discussion on the controversial topic
Transfer activity	Students are tasked with finding and categorizing instances of transfer that they can find in 24 h The key similarities and differences that triggered transfer are then analyzed vis-à-vis forms of relational reasoning

However, when measuring the relational reasoning capability of tertiary students, the decision was made to employ the TORR over other measures like the vTORR or R^2W^2 because it is a fluid rather than a crystallized measure. Cattell (1940, 1963) is credited with drawing the distinction between fluid and crystallized mental assessments (Carpenter et al., 1990; Schipolowski et al., 2014). What characterizes a *fluid ability measure* like the TORR or the Raven's (1941) is the presumption that respondents have access to all that is needed to complete the problem or task within the problem itself. In effect, there is no specific body of conceptual knowledge or procedural skills that the students in LHL would need to have acquired in order to perform well on the TORR over and above their ability to reason relationally. On more crystallized measures, like vTORR and R^2W^2, students would have to be familiar with words and their subtle meanings to demonstrate their reasoning abilities. Those differences between fluid and crystallized measures can be seen by comparing the sample TORR items in **Figure 1** to sample items from the vTORR and R^2W^2 displayed in **Figure 2**.

Another reason to assess relational reasoning with a generic measure like the TORR is because students enrolled in LHL are pursuing majors in the natural, social, applied, and formal sciences (Alexander, 2019). Such domain diversity means that these students' knowledge and skills are expected to vary. The use of a generic measure, therefore, creates a more level playing field when making judgments about these students' relational reasoning capabilities. Further, the TORR remains a strong predictor of achievement in varied academic fields and professional practices (Dumas and Schmidt, 2015; Dumas et al., 2016; Fountain, 2016; Baggetta, 2019).

After completing the 32 items on the TORR, the students enrolled in LHL receive their results reported as a standardized relational reasoning quotient (RRQ) with a mean of 100 and a standard deviation of 15 (Dumas and Alexander, 2016). The students also receive performance data for each of the four 8-item scales that represent the four reasoning forms. The mean for each scale is 4. Students receive no additional feedback on the specific items nor are they given any explanation of how items in the scales should have been analyzed.

Phase 2: Explicit Instruction in Relational Reasoning

Once students have taken the TORR and have their profile recorded, we share with them a definition of relational reasoning and explain briefly what each form captures in terms of its underlying pattern: analogy (similarity), anomaly (aberrance), antinomy (exclusivity/binary), and antinomy (opposition/continuous). The gist of this relational reasoning overview was presented in the first section of this article. In sharing this general information about relational reasoning with the students, we never refer to any of the problems on the TORR or deal with figural problems similar to those items. Not only would that be unacceptable, since the TORR items except for the sample problems, are proprietary information, but also because we want students to see how relational reasoning permeates all academic domains. One or two examples of each form are provided representing different domains. For instance, to exemplify analogies we might display a ratio problem such as "3:9 is equivalent to 4:?" or a classic verbal comparison like "ocean: bay: continent: _**X**_." In this initial phase of instruction, we also establish the predictive power of relational reasoning as demonstrated in the research and describe the role that this higher-order cognitive ability plays in the students' academic domains.

The second phase of explicit instruction involves introducing and then practicing the four essential componential processes critical to any form of relational reasoning (Sternberg, 1977): encoding, inferring, mapping, and applying. Alexander and colleagues have conducted explicit training in analogical reasoning with very young children through young adults using the componential processes as the framework (White and Alexander, 1986; Alexander et al., 1987a,b; Pate et al., 1989). Simply defined, *encoding* entails examining elements of a problem or problem space to ensure understanding of any givens, whereas *inferring* requires finding connections between individual elements based on whatever meaning was derived from encoding. These two processes thus result in a meaningful association essential for any pattern, even a simple linear sequence. However, unlike simple linear patterns, relational reasoning forms involve "relations among relations." *Mapping* is the componential process that is required to link the initial pattern just inferred to another set of associated elements that represent a related pattern. The final componential process, *applying*, involves completing the problem and recognizing the underlying structure, which could then be iterated.

To illustrate these componential processes, we let students see how they function in a simple analogy problem like "ocean: bay: continent: _**X**_." Students begin by encoding the term *ocean* and identifying its salient attributes (largest body of water) and then encoding the term *bay* (inlet of water connected to a larger body of water like an ocean and surrounded by land on three sides). Next students must infer a relation between ocean and bay (both are bodies of water, but bays are smaller and open to oceans on one side and surrounded by land on the remaining sides), To complete the mapping, students must form a meaningful association between *ocean* and *continent* (both are the largest geographical bodies of water and land, respectively). Finally, applying means that students must identify the critical attributes of the missing element that would parallel the relation of ocean to bay. Specifically, they need to recognize that they are looking for a small body of land, connected to a continent on one side and water on the remaining three sides (answer: peninsula).

While the componential processes of encoding, inferring, mapping, and applying were conceptualized with only analogical reasoning in mind (Sternberg, 1977), Grossnickle et al. (2016) found that those same processes were core to anomalous, antinomous, and antithetical reasoning, as well. Moreover, these researchers found that lower performing students exhibited difficulties inferring and mapping on problems representing all relational reasoning forms. Thus, once students had practiced using the componential processes on verbal analogy problems, they were introduced to verbal problems tapping anomalous, antinomous, and antithetical reasoning. As with the analogy problems, they were directed to encode, infer, map, and apply while solving these problems and received feedback on their performance. When this training and practice phase concluded, our goal was to demonstrate the importance of relational reasoning to other key facets of academic learning and performance beginning with transfer.

A Directions: The sentence below describes a situation. Select the sentences from the answer choices below that describe the ***most similar*** situation.

The man breathed a sigh of disappointment when he opened his wife's gift.

A. The boy felt saddened as he packed up his old clothes to hand down to his younger brother.
B. The mother couldn't hide her sense of defeat when she received her child's report card.
C. The girl felt an immediate sense of relief after getting an invitation to the dance.
D. The child's bafflement showed when his friend didn't share his dessert with him.

B **Directions**: In the item below, there are ***two distinct sets*** of words (Set 1 and Set 2). Within each set, the words fit together according to their meaning. Select one of the four given words that fits ***only*** in Set 1, but ***not*** Set 2.

Set 1	Set 2
TAIL	BASEMENT
BEAK	HALL
PECK	BUILD
FLY	RENOVATE

A. DOME
B. **CLAW**
C. PAVE
D. WING

FIGURE 2 | Sample items from the verbal test of relational reasoning analogy scale **(A)** and the relational reasoning with words antinomy scale **(B)**.

Phase 3: Expansion and Transfer

Transfer, "the process of using knowledge or skills acquired in one context in a new or varied context" (Alexander and Murphy, 1999, p. 561), has the well-earned reputation of being one of the most challenging cognitive abilities for students to master (Gick and Holyoak, 1980; Detterman and Sternberg, 1993). Indeed, the literature is replete with evidence that students are typically poor at transferring knowledge and skills from one formal learning environment to another or to situations in the world outside the classroom (Perkins and Salomon, 2012). This indictment notwithstanding, transfer remains a foundational ability for all students, especially tertiary students, to develop and hone if they are to be successful both in their university studies and their chosen professions. Many factors or conditions have been proposed as barriers of transfer such as the contention that the initial learning was not substantive enough to foster transfer (Dinsmore et al., 2013). It could also be that the new context appears quite dissimilar to the context in which the knowledge and skills were acquired. Learners' individual characteristics can also foster or frustrate transfer including their perceptiveness, motivations, or metacognitive and strategic abilities (Corkill and Fager, 1995; Billing, 2007; Dinsmore et al., 2013).

Because of the importance of transfer for learning in higher education and for future career success, this area is stressed in

LHL. We concur with Billing (2007) that transfer is more likely to occur when principles of reasoning are taught in conjunction with academic content. Thus, when we introduce this topic to students, we stress that learning to transfer will "bootstrap" the knowledge and skills they are working hard to acquire. Yet, we acknowledge the difficulty of developing a habit of transfer. A first step in forming this habit of mind is for them to become perceptive and alert to transfer opportunities—a process that can be aided by their use of relational reasoning. We share evidence with them that shows that analogical reasoning ability can be a key to transfer (Reeves and Weisberg, 1994; Richland and McDonough, 2010).

As Alexander and Murphy (1999) argued, transfer and analogical reasoning are related processes. Unless students can perceive similarities between a specific task they learned with another task encountered in a different context, they will not be primed to engage in transfer. Further, the more these students can look beyond the surface features of those tasks and contexts and find underlying similarities, the better they can make use of what they already know and can do, which is where their relational reasoning, and especially analogical reasoning can be most helpful (Richland and McDonough, 2010). We also alert students to be aware of the dissimilarities between the initial tasks and contexts and these transfer opportunities, so they can iterate or modify their problem-solving processes appropriately. When this instruction on transfer coupled with relational reasoning has been completed, students in LHL are given the task of documenting as many instances of transfer as they can within a 24-h period. The cases of transfer the students record are then discussed in class for added reinforcement.

Critical Reading is another basic skill for students in tertiary education, as they are required to read about what they are studying. For students majoring in certain fields, such as history, philosophy, psychology, and sociology, the reading required can be quite extensive. In LHL, for instance, students are reading, summarizing, and comparing articles and chapters routinely. They also are required to carry out a multiple-source use (MSU) project for which they conduct an online search on a controversial topic (e.g., The effects of overuse of social media on students' academic, social, physical, and emotional well-being). The students then select and summarize appropriate sources to use as the basis for an argumentative essay. The students' prior training in relational reasoning becomes relevant to these tasks in several ways.

For one, we recognized early in the rollout of this course that many of these tertiary students did not have effective strategies for dealing with course readings in an integrative manner. Thus, when they were tasked with writing a comparison of two readings that offered different perspectives on an issue (e.g., expertise), many had no clue how to begin. Drawing on their relational reasoning training, we suggested that they could chart key similarities and dissimilarities between the readings and then use that relational analysis to organize their written comparisons. This deep analysis technique was also advantageous when the students had to integrate information across multiple documents. Again, by thinking relationally about the documents in terms of core similarities and dissimilarities, the students were better prepared to integrate the positions, arguments, and evidence presented in their selected sources. The resulting analysis could help them formulate their own position on the controversial issue and provide them with evidence to support their position when composing their argumentative essay.

As with their writing, we were somewhat surprised to find that a good number of the tertiary students enrolled in LHL were stymied in their ability to carry out ***quality discussion***. Frequently what we witnessed were students voicing points that were unconnected to what others had already been said or they did not make it clear whether their statement was meant to support or counter what others had previously expressed. In effect, what should have been a discussion became a string of separate statements. Thus, what we did to improve the quality of class discussions was to share the research illustrating how professionals like medical doctors diagnosing their patients or engineering students working collaboratively on a project design would use relational reasoning in their discourse to work toward a shared outcome. We also had students read an excellent piece by Murphy et al. (2017) that describes how instances of relational reasoning in students' discourse served to reinforce, conditionalize, or counter others' comments. Tertiary students in LHL then tried their hands at carrying on a discussion based on their MSU research, and the ideas expressed in their argumentative essays. We saw a marked improvement in discussion quality following this brief intervention with many more uptakes. Murphy et al. define *uptakes* as direct acknowledgments of what others have contributed and an explicit indication of how one's response is related.

In this final section of the article, we have attempted to show how theory and research in relational reasoning can easily and effectively be embedded in instruction within a university classroom. Of course, we do not know if those students who were part of LHL carried the lessons from this unique course into the rest of their tertiary education or into their continued professional development. That remains our hope and the subject for future empirical research.

FINAL THOUGHTS AND FUTURE DIRECTIONS

An overarching goal we set for this treatise on relational reasoning was to establish its value as a higher-order cognitive ability for tertiary students' current learning and performance as well as their future success in their chosen professions. To support that claim, we not only shared what is known about the nature, forms, and development of relational reasoning, but also its contributions to a range of human activities and academic domains. Finally, we looked at how relational reasoning was woven through the content of one university course—from its assessment and training to its expansion into other topics of importance such as transfer, critical reading, and writing abilities. There is no question that higher education has a mission of equipping students with the knowledge and skills they will need to function in their lives, including in their chosen professions. It is also presumed that tertiary education will contribute to

students' ability to reason deeply and effectively and to manifest habits of mind and habits of action that are indicative of a well-educated mind. But how is this ability to reason deeply and effectively explicitly developed within tertiary education? Our contention is that the assessment, training, and expansion of relational reasoning is one significant step in the right direction.

Of course, what we have shared in this article is one humble case of what could be done within tertiary education. There is much more that must be done before the potential value of relational reasoning in tertiary education can be more fully assessed and more fully realized. Nonetheless, we would like this overview to be an opportunity for others to explore relational reasoning as an essential component within higher education; one that can foster the habits of mind and habits of action that we seek to instill in students who currently populate universities and will become tomorrow's doctors, scientists, teachers, counselors, and engineers.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the IRB University of Maryland. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

REFERENCES

Albanese, M. A., and Dast, L. (2014). Problem-based learning: outcomes evidence from the health professions. *J. Excell. Coll. Teach.* 25, 239–252.

Alexander, P. A. (2003). The development of expertise: the journey from acclimation to proficiency. *Educ. Res.* 32, 10–14. doi: 10.3102/0013189X032008010

Alexander, P. A. (2004). "A model of domain learning: reinterpreting expertise as a multidimensional, multistage process," in *Interactive Models*, eds D. Y. Dai and R. J. Sternberg (Mahwah, NJ: Lawrence Erlbaum Associates), 271–298.

Alexander, P. A. (2006). *Psychology in Learning and Instruction*. Upper Saddle River, NJ: Pearson.

Alexander, P. A. (2016). Relational thinking and relational reasoning: harnessing the power of patterning. *NPJ Sci. Learn.* 1:16004. doi: 10.1038/npjscilearn.2016.4

Alexander, P. A. (2019). Individual differences in college-age learners: the importance of relational reasoning for learning and assessment in higher education. *Br. J. Educ. Psychol.* 89, 416–428. doi: 10.1111/bjep.12264

Alexander, P. A. (in press a). "Good versus bad motivation? Avoiding the lure of false dichotomies," in *Motivation Science: Controversies and Insights*, eds M. Bong, S.-I. Kim, and J. Reeve (Oxford: Oxford University Press).

Alexander, P. A. (in press b). "The interplay of knowledge, strategies, and interest in the development of expertise within professions," in *Professions and Proficiency (Knowledge and Space)*, Vol. 18, eds J. Glückler, C. Winch, and A. M. Punstein (Berlin: Springer).

Alexander, P. A., and Baggetta, P. (2014). "Percept-concept coupling and human error," in *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences*, eds D. N. Rapp and J. L. G. Baasch (Boston, MA: MIT Press), 297–327.

Alexander, P. A., Dumas, D., Grossnickle, E. M., List, A., and Firetto, C. M. (2016a). Measuring relational reasoning. *J. Exp. Educ.* 84, 119–151. doi: 10.1080/00220973.2014.963216

Alexander, P. A., Jablansky, S., Singer, L. M., and Dumas, D. (2016b). Relational reasoning: what we know and why it matters. [Special issue on education]. *Policy Insights Behav. Brain Sci.* 3, 36–44. doi: 10.1177/2372732215622029

Alexander, P. A., Singer, L., Jablansky, S., and Hattan, C. (2016c). Relational reasoning in word and in figure. *J. Educ. Psychol.* 108, 1140–1152. doi: 10.1037/edu0000110

Alexander, P. A., and Murphy, P. K. (1999). Nurturing the seeds of transfer: a domain-specific perspective. *Int. J. Educ. Res.* 31, 561–567. doi: 10.1016/S0883-0355(99)00024-5

Alexander, P. A., and The Disciplined Reading and Learning Research Laboratory [DRLRL] (2012). Reading into the future: competence for the 21st century. *Educ. Psychol.* 47, 1–22. doi: 10.1080/00461520.2012.722511

Alexander, P.A., and The Disciplined Reading and Learning Research Laboratory [DRLRL] (2014). *Test of Relational Reasoning*. College Park, MD: University of Maryland.

Alexander, P.A., and The Disciplined Reading and Learning Research Laboratory [DRLRL] (2018). *Test of Relational Reasoning-Junior*. College Park, MD: University of Maryland.

Alexander, P. A., White, C. S., Haensly, P. A., and Crimmins-Jeanes, M. (1987a). Training in analogical reasoning. *Am. Educ. Res. J.* 24, 387–404.

Alexander, P. A., Wilson, A. F., White, C. S., Willson, V. L., Tallent, M. K., and Shutes, R. E. (1987b). Effects of teacher training on children's analogical reasoning performance. *Teach. Teach. Educ.* 3, 275–285. doi: 10.1016/0742-051X(87)90020-5

Altshuller, G. S., and Shapiro, R. B. (1956). Psychology of inventive creativity. *Vopr. Psikhol.* 6, 37–49.

Baggetta, P. (2019). The Contributions of Crystallized Cross-Domain Knowledge and Fluid Relational Reasoning Ability to Ninth- and Twelfth-Grade Students' Performance on Scholastic Aptitude and Content-Specific Achievement Measures (Order No. 27735583). ProQuest Dissertations & Theses Global. Availabe online at: https://www.proquest.com/dissertations-theses/contributions-crystallized-cross-domain-knowledge/docview/2434104455/se-2?accountid=14696 (accessed March 30, 2022).

Baggetta, P., and Alexander, P. A. (2016). Conceptualization and operationalization of executive function. *Mind Brain Educ.* 10, 10–33. doi: 10.1111/mbe.12100

Bartha, P. (2019). "Analogy and analogical reasoning," in *The Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta. Available online at: https://plato.stanford.edu/archives/spr2019/entries/reasoning-analogy/ (accessed February 07, 2022).

Beattie, I. R. (2002). Are all "adolescent econometricians" created equal? Racial, class, and gender differences in college enrollment. *Sociol. Educ.* 75, 19–43. doi: 10.2307/3090252

Billing, D. (2007). Teaching for transfer of core/key skills in higher education: cognitive skills. *High. Educ.* 53, 483–516. doi: 10.1007/s10734-005-5628-5

Borsboom, D., and Mellenbergh, G. J. (2002). True scores, latent variables, and constructs: a comment on Schmidt and Hunter. *Intelligence* 30, 505–514. doi: 10.1016/S0160-2896(02)00082-X

Bringsjord, S., Hummel, J., and Licato, J. (2017). *Great Computational Intelligence in the Formal Sciences via Analogical Reasoning*. Arlington, VA: Air Force Research Laboratory/AF Office of Scientific Research.

Bunge, S. A., and Leib, E. R. (2020). How does education hone reasoning ability? *Curr. Dir. Psychol. Sci.* 29, 167–173. doi: 10.1177/0963721419898818

Carlson, S. M. (2009). Social origins of executive function development. *New Dir. Child Adolesc. Dev.* 2009, 87–97. doi: 10.1002/cd.237

Carpenter, P. A., Just, M. A., and Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychol. Rev.* 97, 404–431. doi: 10.1037/0033-295X.97.3.404

Cattell, R. B. (1940). A culture-free intelligence test. I. *J. Educ. Psychol.* 31, 161–179. doi: 10.1037/h0059043

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: a critical experiment. *J. Educ. Psychol.* 4, 1–22. doi: 10.1037/h0046743

Chae, S.-E., and Alexander, P. A. (2021a). Exploring potential educational and social contributors to relational reasoning development: a study of typically-and atypically-schooled adolescents and adults in South Korea. *Mind Brain Educ.* doi: 10.1002/MBE.12311 [Epub ahead of print].

Chae, S.-E., and Alexander, P. A. (2021b). The development of relational reasoning in South Korean elementary and middle-school students: a cross-sectional investigation. *Front. Psychol. Cogn.* 12:630609. doi: 10.3389/fpsyg.2021.630609

Chinn, C. A., and Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science instruction. *Rev. Educ. Res.* 63, 1–49. doi: 10.2307/1170558

Chinn, C. A., and Malhotra, B. A. (2002). Children's responses to anomalous scientific data: how is conceptual change impeded? *J. Educ. Psychol.* 94, 327–343. doi: 10.1037/0022-0663.94.2.327

Corkill, A. J., and Fager, J. J. (1995). Individual differences in transfer *via* analogy. *Learn. Individ. Differ.* 7, 163–187. doi: 10.1016/1041-6080(95)90009-8

Cramer, M., and Dauphin, J. (2020). A structured argumentation framework for modeling debates in the formal sciences. *J. Gen. Philos. Sci.* 51, 219–241. doi: 10.1007/s10838-019-09443-z

de Laplante, K. (2006). "Sources of domain-independence in the formal sciences," in *Foundations of the Formal Sciences IV: History of the Concept of the Formal Sciences*, eds B. Löwe, V. Peckhaus, and T. Räsch (Norcross, GA: College Publications), 1–16.

Detterman, D. K., and Sternberg, R. J. (1993). *Transfer on Trial: Intelligence, Cognition, and Instruction*. Norwood, NJ: Ablex Publishing.

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurevpsych-113011-143750

Dinsmore, D. L., Baggett, P., Doyle, S., and Loughlin, S. M. (2013). The role of initial learning, problem features, prior knowledge, and pattern recognition on transfer success. *J. Exp. Educ.* 82, 121–141. doi: 10.1080/00220973.2013.835299

Dumas, D., and Alexander, P. A. (2016). Calibration of the test of relational reasoning. *Psychol. Assess.* 28, 1303–1318. doi: 10.1037/pas0000267

Dumas, D., and Alexander, P. A. (2018). Assessing differential item functioning on the test of relational reasoning. *Front. Educ.* 3:14. doi: 10.3389/feduc.2018.00014

Dumas, D., Alexander, P. A., Baker, L. M., Jablansky, S., and Dunbar, K. N. (2014). Relational reasoning in medical education: patterns in discourse and diagnosis. *J. Educ. Psychol.* 106, 1021–1035. doi: 10.1037/a0036777

Dumas, D., Alexander, P. A., and Grossnickle, E. M. (2013). Relational reasoning and its manifestations in the educational context: a systematic review of the literature. *Educ. Psychol. Rev.* 25, 391–427. doi: 10.1007/s10648-013-9224-4

Dumas, D., and Schmidt, L. (2015). Relational reasoning as predictor for engineering ideation success using TRIZ. *J. Eng. Des.* 26, 74–88. doi: 10.1080/09544828.2015.1020287

Dumas, D., Schmidt, L. C., and Alexander, P. A. (2016). Predicting creative problem solving in engineering design. *Think. Skills Creat.* 21, 50–66. doi: 10.1016/j.tsc.2016.05.002

Dumontheil, I., Houlton, R., Christoff, K., and Blakemore, S. (2010). Development of relational reasoning during adolescence. *Dev. Sci.* 13, F15–F24. doi: 10.1111/j.1467-7687.2010.01014.x

Dunbar, K., and Blanchette, I. (2001). The *in vivo*/*in vitro* approach to cognition: the case of analogy. *Trends Cogn. Sci.* 5, 334–339. doi: 10.1016/S1364-6613(00)01698-3

Fischer, K. W., and Silvern, L. (1985). Stages and individual differences in cognitive development. *Annu. Rev. Psychol.* 36, 613–648. doi: 10.1146/annurev.ps.36.020185.003145

Fountain, L. (2016). *Relations Among Topic Knowledge, Individual Interest, and Relational Reasoning, and Critical Thinking in Maternity Nursing* (Order No. 10159037). ProQuest Dissertations & Theses Global. Available online at: https://www.proquest.com/dissertations-theses/relations-among-topic-knowledge-individual/docview/1824362967/se-2 (accessed March 30, 2022).

Gentner, D. (2002). "Analogy in scientific discovery: the case of Johannes Kepler," in *Model-Based Reasoning: Science, Technology, Values*, eds L. Magnani and N. J. Nersessian (New York, NY: Kluwer Academic/ Plenum Publisher), 21–a39.

Gentner, D., and Gentner, D. R. (1983). *Flowing Waters or Teeming Crowds: Mental Models of Electricity*. Hillsdale, NJ: Erlbaum.

Gick, M. L., and Holyoak, K. J. (1980). Analogical problem solving. *Cogn. Psychol.* 12, 306–355. doi: 10.1016/0010-0285(80)90013-4

Gilbert, M. A. (2009). Defeating bigenderism: changing gender assumptions in the twenty-first century. *Hypatia* 24, 93–112. doi: 10.1111/j.1527-2001.2009.01047.x

Grossnickle, E. M., Dumas, D., Alexander, P. A., and Baggetta, P. (2016). Individual differences in the process of relational reasoning. *Learn. Instr.* 42, 141–159. doi: 10.1016/j.learninstruc.2016.01.013

Hackman, D. A., and Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends Cogn. Sci.* 13, 65–73. doi: 10.1016/j.tics.2008.11.003

Halford, G. S. (1992). Analogical reasoning and conceptual complexity in cognitive development. *Hum. Dev.* 35, 193–217. doi: 10.1159/000277167

Harris, E. E. (2014). *Hypothesis and Perception: The Roots of Scientific Method*. London: Routledge.

Hofstadter, D. R. (1979). *Gödel, Escher, Bach: An Eternal Golden Braid*. New York, NY: Basic Books.

Hsu, N. S., Novick, J. M., and Jaeggi, S. M. (2014). The development and malleability of executive control abilities. *Front. Behav. Neurosci.* 8:221. doi: 10.3389/fnbeh.2014.00221

Hummel, J. E., Licato, J., and Bringsjord, S. (2014). Analogy, explanation, and proof. *Front. Hum. Neurosci.* 8:867. doi: 10.3389/fnhum.2014.00867

Jablansky, S. (2020). *Relational Reasoning and Socially Shared Regulation of Learning in Collaborative Problem Solving* (Order No. 28151748). ProQuest Dissertations & Theses Global. Available online at: https://www.proquest.com/dissertations-theses/relational-reasoning-socially-shared-regulation/docview/2496359895/se-2?accountid=14696 (accessed March 30, 2022).

Jablansky, S., Alexander, P. A., Dumas, D., and Compton, V. (2016). Developmental differences in relational reasoning among primary and secondary school students. *J. Educ. Psychol.* 108, 592–608. doi: 10.1037/edu0000070

Jablansky, S., Alexander, P. A., Dumas, D., and Compton, V. (2020). The development of relational reasoning in primary and secondary school students: a longitudinal investigation in technology education. *Int. J. Technol. Des. Educ.* 30, 973–993. doi: 10.1007/s10798-019-09529-1

James, W. (1893). *Psychology: A Briefer Course*. New York, NY: Henry Holt & Co.

Kaplan, S., and Berman, M. G. (2010). Directed attention as a common resource for executive functioning and self-regulation. *Perspect. Psychol. Sci.* 5, 43–57. doi: 10.1177/1745691609356784

Kavousi, S., Miller, P., and Alexander, P. A. (2020). The role of metacognition in the first-year design lab. *Educ. Technol. Res. Dev.* 68, 3471–3494. doi: 10.1007/s11423-020-09848-4

Krawczyk, D. C. (2012). The cognition and neuroscience of relational reasoning. *Brain Res.* 1428, 13–23. doi: 10.1016/j.brainres.2010.11.080

Krawczyk, D. C., McClelland, M. M., and Donovan, C. M. (2011). A hierarchy for relational reasoning in the prefrontal cortex. *Cortex* 47, 588–597. doi: 10.1016/j.cortex.2010.04.008

Lehtinen, E. (2012). "Learning of complex competences: on the need to coordinate multiple theoretical perspectives," in *Language: Competencies—Contact—Change*, eds A. Koskensalo, J. Smeds, A. Huguet, and R. de Cillia (Berlin: LIT Verlag), 13–27. doi: 10.1186/s12913-016-1423-5

Lightman, A., and Gingerich, O. (1992). When do anomalies begin? *Am. Assoc. Adv. Sci.* 255, 690–695. doi: 10.1126/science.255.5045.690

Loughlin, S. M., Grossnickle, E. M., Dinsmore, D. L., and Alexander, P. A. (2015). "Reading" paintings: evidence for trans-symbolic and symbol-specific comprehension processes. *Cogn. Instr.* 33, 257–293. doi: 10.1080/07370008.2015.1076822

Löwe, B. (2002). The formal sciences: their scope, their foundations, and their unity. *Synthese* 133, 5–11. doi: 10.1023/a:1020887832028

Mandlebrot, B. B. (1982). *The Fractal Geometry of Nature*. San Francisco, CA: W. H. Freeman & Co.

McConnell, M. M., Regehr, G., Wood, T. J., and Eva, K. W. (2012). Self-monitoring and its relationship to medical knowledge. *Adv. Health Sci. Educ.* 17, 311–323. doi: 10.1007/s10459-011-9305-4

Mertens, D. M. (2019). *Research and Evaluation in Education and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods*. Thousand Oaks, CA: Sage.

Murphy, P. K., Firetto, C. M., and Greene, J. A. (2017). Enriching students' scientific thinking through relational reasoning: seeking evidence in texts, tasks, and talk. *Educ. Psychol. Rev.* 29, 105–117. doi: 10.1007/s10648-016-9387-x

Nowotny, H. (2005). The increase of complexity and its reduction: emergent interfaces between the natural sciences, humanities and social sciences. *Theory Cult. Soc.* 22, 15–31. doi: 10.1177/0263276405057189

Orrill, R., and Shapiro, L. (2005). From bold beginnings to an uncertain future: the discipline of history and history education. *Am. Hist. Rev.* 110, 727–751. doi: 10.1086/ahr.110.3.727

Pate, P. E., Alexander, P. A., and Kulikowich, J. M. (1989). "Assessing the effects of training social studies content and analogical reasoning processes on sixth-graders' domain-specific and strategic knowledge," in *Middle School Research: Selected Studies 1989*, ed. D. B. Strahan (Columbus, OH: Research Committee of the National Middle School Association), 19–29. doi: 10.1080/08851700.1989.11670298

Patel, V. L., Arocha, J. F., and Kaufman, D. R. (1999). "Expertise and tacit knowledge in medicine," in *Tacit Knowledge in Professional Practice: Researcher and Practitioner Perspectives*, eds R. J. Sternberg and J. A. Horvath (Hove: Psychology Press), 75–99.

Perkins, D. N., and Salomon, G. (2012). Knowledge to go: a motivational and dispositional view of transfer. *Educ. Psychol.* 47, 248–258. doi: 10.1080/00461520.2012.693354

Priestley, J. (1767). *The History and Present State of Electricity: With Original Experiments*. London: J. Dodsley, J. Johnson, B. Davenport, and T. Cadell.

Raven, J. C. (1941). Standardization of progressive matrices, 1938. *Br. J. Med. Psychol.* 19, 137–150. doi: 10.1111/j.2044-8341.1941.tb00316.x

Reed, S. K. (2016). The structure of ill-structured (and well-structured) problems revisited. *Educ. Psychol. Rev.* 28, 691–716. doi: 10.1007/s10648-015-9343-1

Reeves, L. M., and Weisberg, R. W. (1994). The role of content and abstract information in analogical transfer. *Psychol. Bull.* 115, 381–400. doi: 10.1037/0033-2909.115.3.381

Richland, L. E., and McDonough, I. M. (2010). Learning by analogy: discriminating between potential analogs. *Contemp. Educ. Psychol.* 35, 28–43. doi: 10.1016/j.cedpsych.2009.09.001

Russell, B., and Lackey, D. (1973). *Essays in Analysis*. London: Allen & Unwin.

Samuel, D. B., and Widiger, T. A. (2008). A meta-analytic review of the relationships between the five-factor model and DSM-IV-TR personality disorders: a facet level analysis. *Clin. Psychol. Rev.* 28, 1326–1342. doi: 10.1016/j.cpr.2008.07.002

Schipolowski, S., Wilhelm, O., and Schroeders, U. (2014). On the nature of crystallized intelligence: the relationship between verbal ability and factual knowledge. *Intelligence* 46, 156–168. doi: 10.1016/j.intell.2014.05.014

Schmidt, W. H., Wang, H. C., and McKnight, C. C. (2005). Curriculum coherence: an examination of US mathematics and science content standards from an international perspective. *J. Curric. Stud.* 37, 525–559. doi: 10.1080/0022027042000294682

Singley, A. T. M., and Bunge, S. A. (2014). Neurodevelopment of relational reasoning: implications for mathematical pedagogy. *Trends Neurosci. Educ.* 3, 33–37. doi: 10.1016/j.tine.2014.03.001

Spearman, C. (1927). *The Abilities of Man: Their Nature and Measurement*. New York, NY: Macmillan.

Stanovich, K. E. (2010). *Rationality and the Reflective Mind*. New York, NY: Oxford University Press.

Sternberg, R. J. (1977). *Intelligence, Information Processing and Analogical Reasoning: The Componential Analysis of Human Abilities*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sullivan, A. L., and Bal, A. (2013). Disproportionality in special education: effects of individual and school variables on disability risk. *Except. Child.* 79, 475–494. doi: 10.1177/001440291307900406

Treur, J. (2021). "Adaptive networks at the crossroad of artificial intelligence and formal, biological, medical, and social sciences," in *Integrated Science*, ed. N. Rezaei (Cham: Springer), 335–375. doi: 10.1007/978-3-030-65273-9_17

Tsoukas, H. (1993). Analogical reasoning and knowledge generation in organization theory. *Organ. Stud.* 14, 323–346. doi: 10.1177/017084069301400301

West, R. F., Toplak, M. E., and Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: associations with cognitive ability and thinking dispositions. *J. Educ. Psychol.* 100, 930–941. doi: 10.1037/a0012842

White, C. S., and Alexander, P. A. (1986). Effects of training on four-year-olds' ability to solve geometric analogy problems. *Cogn. Instr.* 3, 261–268. doi: 10.1207/s1532690xci0303_6

Zhao, H., and Alexander, P. A. (2022). *Uncovering Word Knowledge Quality Through Reasoning Relationally with Words*. [Manuscript submitted for publication]. College Park, MD: Department of Human Development and Quantitative Methodology, University of Maryland.

Zhao, H., Alexander, P. A., and Sun, Y. (2021). Relational reasoning's contributions to mathematical thinking and performance in Chinese elementary and middle-school students. *J. Educ. Psychol.* 113, 279–303. doi: 10.1037/edu0000595

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Education

SYSTEMATIC REVIEW
published: 26 May 2022
doi: 10.3389/feduc.2022.885917

Systematic Review of Learning Generic Skills in Higher Education—Enhancing and Impeding Factors

Tarja Tuononen[1], Heidi Hyytinen[1], Katri Kleemola[1], Telle Hailikari[1], Iina Männikkö[2] and Auli Toom[1]*

[1] Centre for University Teaching and Learning (HYPE), Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland, [2] Department of Teacher Education, University of Turku, Turku, Finland

The research field on generic skills in higher education has expanded rapidly. In addition, the importance of generic skills has been highlighted both in educational policy discourses and in practice of higher education. The present study reviews theoretical, methodological, and empirical viewpoints on learning generic skills and synthesizes the empirical evidence about the factors that enhance and impede student learning of generic skills. Altogether 116 articles were included in the analysis. The systematic analysis revealed remarkable variation in concepts, research methods, and operationalization of generic skills. These findings suggest that research in this field is still incoherent. According to the results, contextual factors that enhance or impede higher education students' learning of generic skills were investigated more often than individual factors. Furthermore, the articles included in this review emphasized learning of work-oriented professional skills over higher-order thinking skills. To ensure the development of research on generic skills, it is important to focus on more coherent theorization and operationalization of the various generic skills. More longitudinal studies with methods that genuinely capture actual skills and their development are also needed to advance the field. The results can be used for future discussions on theorization, empirical research, and practical development of student learning of generic skills.

Keywords: generic skills, learning, higher education, systematic (literature) review, enhancing and impeding factors

OPEN ACCESS

Edited by:
Dominik E. Froehlich, University of Vienna, Austria

Reviewed by:
Tzu-Chiang Lin, National Kaohsiung University of Science and Technology, Taiwan
Edith Braun, Justus-Liebig Universität, Germany

****Correspondence:***
Tarja Tuononen
tarja.tuononen@helsinki.fi

Specialty section:
This article was submitted to Higher Education, a section of the journal Frontiers in Education

Received: *28 February 2022*
Accepted: *03 May 2022*
Published: *26 May 2022*

Citation:
Tuononen T, Hyytinen H, Kleemola K, Hailikari T, Männikkö I and Toom A (2022) Systematic Review of Learning Generic Skills in Higher Education—Enhancing and Impeding Factors. Front. Educ. 7:885917. doi: 10.3389/feduc.2022.885917

INTRODUCTION

Generic skills, such as critical thinking, collaboration, communication, argumentation, and problem-solving skills, usually refer to cognitive skills and higher order thinking skills, as well as twenty-first century competence and future citizens' literacy. Learning generic skills is widely singled out as the key aim of higher education in addition to domain-specific knowledge and skills (e.g., Arum and Roksa, 2011; Hyytinen et al., 2019; Shavelson et al., 2019). The importance of generic skills has been also highlighted in the transition phase to work and later in working life (Tuononen et al., 2019). Similarly, generic skills are considered essential for citizens of the twenty-first century in various policy papers and reports (Strijbos et al., 2015; OECD, 2019). As part of a discussion on educational policy, several lists of the key generic skills of higher education have been compiled (European Parliament Council, 2008; OECD, 2019). For example, the European Parliament Council (2008) has determined the key generic skills that should be included in higher

education degrees. Consequently, generic skills are found as learning objectives in almost all higher education curricula today. Naturally, the aim is to organize teaching so as to enhance student learning in the best possible way. Therefore, it is not surprising that higher education students' generic skills have also attracted remarkable interest from researchers, and become an expanding field of research.

Unfortunately, this broad interest in generic skills and proliferation of studies involves some disadvantages. The interests, intentions, and perspectives of various stakeholders have influenced the research on generic skills and especially the development of research instruments (Strijbos et al., 2015; Muukkonen et al., 2019; Toom et al., 2021). Thus, the research field is at risk of fragmentation. Recent evidence suggests that there is conceptual incoherence in the research field of generic skills as well as a lack of clear theoretical frameworks and robust instruments (e.g., Barrie, 2006; Braun et al., 2012; El Soufi and See, 2019). Another disadvantage is related to research designs and methods. It seems that previous research has relied mainly on indirect methods and materials, such as self-reports of learning, in the investigations of generic skills, and only a limited number of studies have applied performance-based methods and focused on learning generic skills in authentic situations (Braun et al., 2012; Zlatkin-Troitschanskaia et al., 2015). In addition to the scattered research on student learning of generic skills, systematic research on the characteristics of the learning environment or other factors contributing to student learning of generic skills is scarce. This may be related to the laborious research designs that the studies would require, or the lack of robust and valid research instruments to measure generic skills and characteristics of the learning environment. In order to obtain a more coherent picture of the status of generic skills research, there is a need for the systematic analysis of the methods and concepts utilized in the studies.

Through a systematic review, this study aims to contribute to existing theoretical, methodological, and empirical viewpoints on learning of generic skills. This study reviews and synthesizes the empirical evidence about higher education students' generic skills and the factors that enhance and impede their learning of generic skills. Moreover, this study explores methods that are used in the empirical studies and elaborates on concepts related to learning generic skills. The research questions are as follows:

1) From the perspective of student learning in higher education, which generic skills are explored in empirical research, and how are they explored?
2) How do higher education students learn generic skills during their studies?
3) Which factors have been identified to enhance or impede student learning of generic skills?

MATERIALS AND METHODS

A combined literature search in the electronic databases of EBSCOHost, Scopus, and Eric was carried out to identify peer-reviewed journal articles in English. The three main keywords utilized in the search were "student learning," "generic skills," and "higher education," but the searching of databases included the combination of words and phrases such as learning or "student learning" and "generic skills" and "higher education" or "university." The search included all disciplines. We searched online and empirical research articles from 2014 to 2019, resulting in over 907 articles. After that, the first and second authors went through the titles and abstracts, and selected those studies that specifically addressed higher education student learning of generic skills. Therefore, the articles focusing solely on teachers' or employers' perspectives on learning generic skills were excluded. In addition, educational policy articles related to generic skills, quality assurance, curriculum analysis, theoretical, and review articles were excluded. Finally, 393 articles were selected for the first phase of review. During this phase, the first and second author read the articles and ensured that the articles met the inclusion criteria. The following inclusion criteria were used: (1) the study was retrieved from a peer-reviewed journal, (2) it was written in English, (3) the study was conducted in the context of higher education, and (4) the study reported empirical evidence on students' learning of generic skills. In addition, duplicates were removed in this phase. After that, a total of 273 articles were included in the analysis. In the second phase of review, the first and sixth author went through the articles and re-checked that they met the criteria. Especially the fourth criterion was at the focus in this phase of article selection. After these thorough reading rounds, 116 articles were finally included in the analysis. In the **Figure 1**, flow selection process is presented.

Analysis

Qualitative content analysis was adopted for the analysis of the articles (Elo and Kyngäs, 2007). First, all articles were read through to gain familiarity with the data and to identify the concepts that were utilized in the studies focusing on the higher education students' learning of generic skills. Each article was analyzed separately and systematically. We found extensive conceptual variation. The articles were categorized based on the concepts utilized in the articles. We identified two types of articles based on the focus of the articles (see **Table 1**). The first type of article focused on sets of generic skills while the others concentrated on specific generic skills. In total, analysis revealed six different specific generic skills. Below, we consider these two types of articles in greater detail.

In the second phase of analysis, the first author further analyzed which generic skills were measured in the first type of article, namely those that focused on sets of generic skills. The measured skills were categorized into 17 main categories based on the analysis. These categories (see **Table 2**) were subsequently reviewed and refined through discussion between all authors. In the third phase, the articles of both types were further analyzed in terms of the research methods used in the studies. In addition, during this phase, learning of generic skills as well as enhancing and impeding factors in learning generic skills were identified from the results sections of the articles. Descriptions of the qualities analyzed were written for each article and collected in Excel worksheets. The fourth and final phase consisted of final interpretations discussed by all the authors. All authors participated in all the phases of analysis, except the second phase which was conducted by the first author.

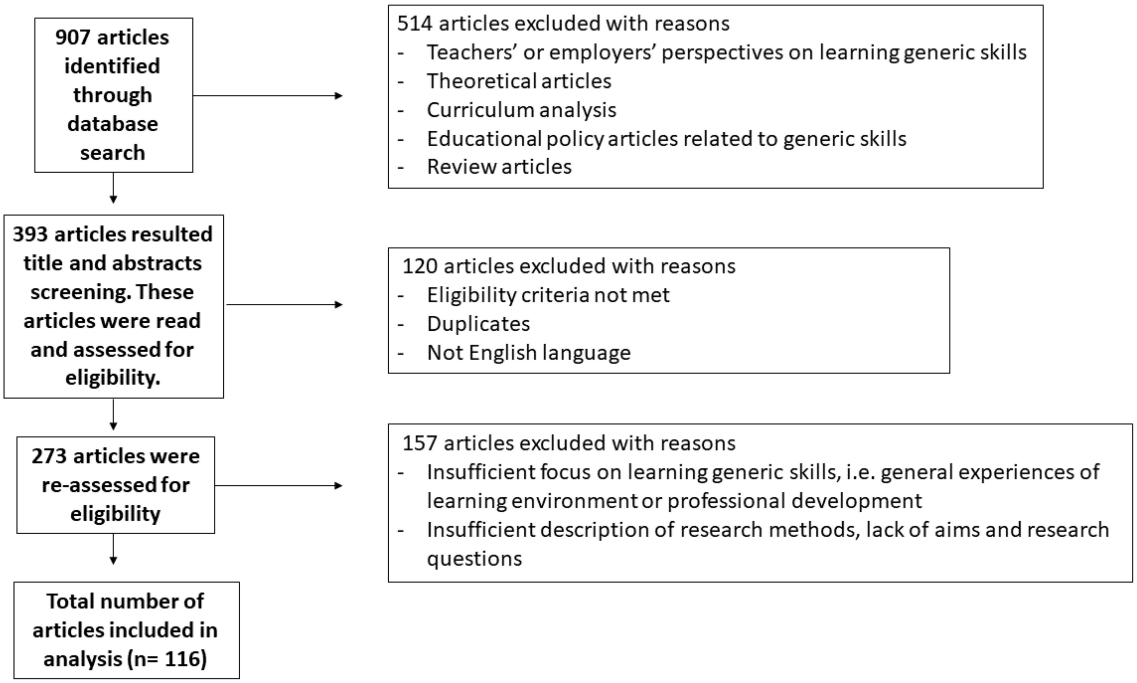


FIGURE 1 | Flow diagram of selection process.

TABLE 1 | Phases of analysis.

Phase of analysis		
Familiarizing oneself with the data	**116 generic skills articles**	
First phase	*Set of generic skills articles* ($n = 70$)	*Specific generic skills articles* ($n = 46$) - Critical thinking skills - Communication skills - Collaboration skills - Creativity and problem-solving skills - Self-regulation skills - Ethical skills
Second phase	Operationalization of measured skills → 17 categories	
Third phase	Research methods, learning of generic skills, enhancing and impeding factors	Research methods, learning of generic skills, enhancing and impeding factors

RESULTS

Measured Skills and Methods in the Reviewed Articles

Our first aim was to explore, from the perspective of student learning in higher education, which generic skills are explored in empirical research, and how they are explored. The final sample included studies that had various objectives, and that were conducted using a wide variety of research methods. There was great variation in the number of the participants in the studies reviewed, from six students to 74,687 students. As mentioned above, there was remarkable variation in the generic skills investigated in the articles (see **Table 1**). There were two types of articles, namely, those focusing on a set of generic skills and those focusing on a specific generic skill at a time. Most of the articles (60%, $n = 70$) focused broadly on sets of generic skills. These studies described their focus as generic skills, or a similar concept, such as employability skills, transferable skills, soft skills, graduate attributes, generic competencies, learning outcomes, academic competencies, core competencies, and non-technical skills. In addition, the rest of the studies framed their research with generic skills but focused on more specific generic skills ($n = 46$), namely critical thinking skills, communication skills, collaboration skills, creativity and problem-solving skills, self-regulation skills, or ethical skills. Due to the difference in the approach the studies adopted, in the following section we report separately the studies that focused on *sets of skills* and the studies that focused on *a specific generic skill*. Hence, section Sets of Generic Skills reveals the variation identified in studies focusing on a set of generic skills and, respectively, section Specific Generic Skills concentrates on articles that focus on specific generic skills by describing the identified skills and the methods used to investigate these skills.

Sets of Generic Skills

In the articles that focused broadly on sets of generic skills, the definitions of generic skills and the methods that were used to measure those skills were varied. Based on the information

TABLE 2 | Sets of generic skills: the main categories and subcategories of measured generic skills.

Main category of generic skills (f)	Subcategories (f)
Professional skills (93)	Professionalism (32), leadership (24), managerial (16), entrepreneurial skills (11), information management (6), project skills (4)
Analytical skills (66)	Critical thinking (23), analytical thinking (20), creative thinking/innovation (18), systems thinking (5)
Applying knowledge (59)	Problem-solving (34), decision making (16), applying theory to practice (9)
Communication skills (59)	Communication (43), writing skills (16)
Collaboration skills (51)	Collaboration skills, teamwork (51)
Time-management skills (29)	Time-management, planning (29)
Study skills (29)	Lifelong learning (10), information searching (8), study skills (5), ability to understand theories (4), knowledge building (2)
Self-knowledge (25)	Self-knowledge including confidence, self-regulation skills, ability to manage emotions and stress, reflection; knowing what study methods are suitable for me; sense of worth and world view; self-criticism (25)
Information technology skills (23)	ICT, computer skills and social media skills (23)
Ethics and responsibility (18)	Research ethics including work ethics, professional moral quality, social responsibility, treat customers data confidentially (18)
Globalization (17)	Globalization, community and citizenship (13), multidisciplinary (4)
Research skills (15)	Research skills, analyze and use numbers and data accurately (15)
Adaptability (10)	Adaptability including adapting to new situations; ability to understand and adapt environment, ability to make changes, flexibility (10)
Personal attributes (10)	Dedication, right personality, perseverance, frankness, open-mindedness, curiosity, resilience, persistence (10)
Foreign language skills (10)	Foreign language (10)
Career skills (5)	Career skills including career planning CV; job applications, interviews, grant applications (5)
Feedback (2)	Utilizing and providing feedback (2)

available about the surveys used, we analyzed which skills were measured as a part of the sets of generic skills (see **Table 2**). The number of skills measured varied from three to 89. These skills were categorized into 17 main categories. The skills most often measured were professional skills (f = 93), including professionalism, leadership, project skills, and entrepreneurial skills. Next, analytical skills (f = 66), applying knowledge (f = 59), communication skills (f = 59), and collaboration skills (f = 51). After these, time-management (f = 29), study skills (f = 29), self-knowledge (f = 25), and ICT skills (f = 23) were included in the instruments. Ethics (n = 18), globalization (f = 17), research skills (f = 15), adaptability (f = 10), foreign language skills (f = 10), and personal attributes (f = 10) were also measured in numerous studies. Additionally, career skills (f = 5) and giving and receiving feedback (f = 2) were measured in a few studies. However, it is important to note that not all articles reported the survey instrument used at all, or the instrument was not reported accurately. In **Table 2**, categories and subcategories of the measured generic skills are presented in greater detail.

Most of the these studies that measured sets of generic skills utilized surveys (e.g., Jackson, 2014a, 2015; Pita et al., 2015; Prokofieva et al., 2015; Abayadeera and Watty, 2016; Joseph et al., 2016; Monteiro et al., 2016; He et al., 2017; Burch et al., 2018; Akhmetshin et al., 2019; López et al., 2019). Control and experimental groups were also used in study designs (Guo, 2019; Tomasson Goodwill et al., 2019). In addition, five articles used qualitative methods (Viviers, 2016; Kridiotis and Swart, 2017; Sonnenschein et al., 2017; Nastiti et al., 2018; Lee et al., 2019), and mixed methods (Bellew and Gabaudan, 2017; Dinning, 2017; Sarkar et al., 2017; Ssegawa and Kasule, 2017; Tran, 2017; Tomasson Goodwill et al., 2019), and one was a mixed-method study using performance-based assessment and interviews (Feldon et al., 2016). In these articles various generic skills were measured using scales including several items or one-item measures (Yin et al., 2014, 2016; Abayadeera and Watty, 2016; Jackson, 2016a; Liu et al., 2017; Yin and Ke, 2017; Guo, 2018; Tuononen et al., 2019).

Specific Generic Skills

The articles that focused on specific generic skills explored critical thinking (10), communication skills (10), collaboration skills (9), creativity and problem-solving skills (8), and self-regulation skills (6). Furthermore, there were a few articles that studied ethical skills (3). These studies utilized various research methods that are presented in greater detail in the following.

Critical Thinking

Studies measuring critical thinking used a variety of methods. In some studies, performance-based assessments were used. These included multiple-choice tests and a few open-ended tasks. Some of the performance assessments used standardized tests (Al-Thani et al., 2016; Ding et al., 2016; Nedelova and Šukolova, 2017; Stone et al., 2017), and in some studies, researchers had created their own performance tasks or used regular examination tasks or course assignments (Sotiriadou and Hill, 2015; Calma, 2017; Utriainen et al., 2017; Lespiau and Tricot, 2018). Many of the studies investigated used self-report surveys to investigate experiences and opinions (Kim, 2015; Sotiriadou and Hill, 2015; Danczak et al., 2017; Ibrahim and Jaaffar, 2017a). One study used an interview as a method (Kim, 2015). In two studies, mixed methods were used, combining two of the above-mentioned methods (Kim, 2015; Sotiriadou and Hill, 2015). In investigating the development of critical thinking, various designs were used. A cross-sectional design was used to compare junior and senior students (Al-Thani et al., 2016), and students in different groups or study fields (Ding et al., 2016; Lespiau and Tricot, 2018). In a few studies, a longitudinal design was used, comparing pre-course and post-course measurements (Kim, 2015; Sotiriadou and Hill, 2015; Stone et al., 2017).

Communication Skills

Many of the studies that focused on students' communication skills also used self-report surveys (Jackson, 2014b, 2016b; Tun Lee-Foo et al., 2015; Mercer-Mapstone and Matthews, 2017; Ibrahim and Jaaffar, 2017a). Typically the studies on students' communications skills utilized multi-method designs, for example combining a survey with written reports (Drury and Muirb, 2014), and writing assignments (Rayner et al., 2016), or multiple-choice tests with long answer questions (Hryciw and Dantas, 2016) and performance assessments (Van Ginkel et al., 2015). In addition, some studies utilized even more complex designs, for example, including dialogue circles, videoing, and team performance measures (Pöysä-Tarhonen et al., 2016), or student surveys, teacher interviews, and student performance in communication tasks (Mercer-Mapstone and Kuchel, 2016).

Creativity and Problem-Solving

Most of the studies that explored creativity and problem-solving used self-report surveys (Wood and Bilsborow, 2014; Techanamurthy et al., 2018; Keinänen and Kairisto-Mertanen, 2019; Mareque et al., 2019). However, there were exceptions as well, especially regarding problem-solving skills. For example, an online game-based assessment tool (Seow et al., 2019), problem-solving tests (Klegeris et al., 2017) and evaluation rubrics were used. Furthermore, many studies explored the influence of some specific factor on the development of the skills, such as innovation pedagogy (Keinänen and Kairisto-Mertanen, 2019), experiential learning pedagogy (Seow et al., 2019), participating in leisure activities (Mareque et al., 2019), or engaging students in complex learning activities (in this case, design-based research) (Wood and Bilsborow, 2014). Mostly the studies focused on exploring students' own perceptions of the level of their skills during studies or upon graduation (Tahir et al., 2017; Techanamurthy et al., 2018) or after a specific pedagogical intervention (Keinänen and Kairisto-Mertanen, 2019; Mareque et al., 2019). Seow et al. (2019) used a quasi-experimental design with a control-group and pre-post test design to explore differences in performance after a specific intervention between the groups. Klegeris et al. (2017) used a cross-sectional design to compare the problem-solving abilities of first- and upper-year students.

Collaboration Skills

Studies measuring collaboration skills utilized surveys (Bravo et al., 2016; Ibrahim and Jaaffar, 2017b; Sridharan et al., 2018; Christensen et al., 2019). Some studies used pre- and post-design to explore students' collaboration skills (Christensen et al., 2019). In addition, evaluation rubrics were used to assess teamwork competencies, including identity, communication, implementation, and regulation (Cela-Ranilla et al., 2014b). Collaboration skills were also explored qualitatively through students' reflection about teamwork.

Self-Regulation Skills

Self-regulation skills were often explored using self-assessments, such as surveys and learning diaries (Ibrahim and Jaaffar, 2017b; Tseng et al., 2019). In addition, evaluation rubrics were used to evaluate self-management skills including planning, organization, development, and assessment (Cela-Ranilla et al., 2014b).

Ethical Skills

Studies investigating ethical skills utilized surveys and students' written reflections as research methods (Howells et al., 2016; Steur et al., 2016; Taplin et al., 2018).

Higher Education Students' Learning of Generic Skills During Their Studies

Our second aim was to explore whether students learn generic skills in higher education. First, we present the results of the studies that focused on sets of generic skills and then the studies that focused on specific generic skills.

Sets of Generic Skills

Most of the articles that investigated sets of generic skills explored students' perceptions of learning of generic skills. The results showed that the students had learnt the generic skills under investigation well (Bonesso et al., 2015; Joseph et al., 2016; Pirog, 2016; Yin et al., 2016; Guo et al., 2017; Larraz et al., 2017; Sarkar et al., 2017; Tahir et al., 2017; Rozlin et al., 2018; López et al., 2019; Skaniakos et al., 2019). A study of Spanish university students showed that students reported to have learnt best the basic general knowledge in the field of study, learning, information management, problem solving, teamwork, concern for quality and motivation to achieve objectives (López et al., 2019). Martínez-Clares and González-Morga (2018) found that students evaluated that they had developed the most in teamwork as well as in ethical and social commitment. Dinning (2017) showed that 60% or more of the students reported improvements in creativity, problem-solving, persuading and influencing, team work, project management, verbal communication, developing new ideas and making things happen, time management, and flexibility. Similarly, Ssegawa and Kasule (2017) found that students reported having learnt skills well, especially adapting to new environments and willingness to learn new ideas. Another study found that students had learned the ability to articulate employability skills (Tomasson Goodwill et al., 2019). Sarkar et al. (2017) found that students' awareness of employability and underpinning skills increased. Students perceived themselves as capable of working independently (Pop and Khampirat, 2019).

Some of the studies that focused broadly on sets of generic skills found that students had not learned generic skills very well (e.g., Perdigones et al., 2014; Monteiro et al., 2016) or learned only a few of them (Abayadeera and Watty, 2016). Some articles listed generic skills which students experienced that they had learnt the least. These skills included time management, oral communication, negotiation, coping with stress, creating viable solutions, and meeting deadlines, ability to use computers, and teamwork (Perdigones et al., 2014; Jackson, 2016a; Ssegawa and Kasule, 2017). In addition, the generic skills that students perceived having had least learning in included entrepreneurial cooperation, leadership skills, IT skills, and cooperation with people from different cultures (Pirog, 2016), speaking and writing in a foreign language (Conchado et al., 2015; Pirog, 2016;

Martínez-Clares and González-Morga, 2018), as well as conflict management (Bonesso et al., 2015). Chan and Fong (2018) found that students generally rated their current competency level lower than the perceived importance of the generic skills to their future career.

Some articles also found differences in generic skills between the students. For example, students' perceptions of generic skills were the highest for students who were satisfied with the guidance and who had progressed well in their studies (Skaniakos et al., 2019). Disciplinary differences were also found, showing that students from the Faculty of Education had the highest scores, while the lowest means were from students in the Faculty of Mathematics and Science and in the Faculty of Social Sciences (Skaniakos et al., 2019). In addition, it was revealed that students with different motivations as well as students from different university types, disciplines, and university years engaged differently with developing generic skills (Tran, 2017). Students in the flipped group reported higher scores for generic skills than students in traditional lecture courses (Guo, 2019). Kirstein et al. (2019) found that students from poorer quality schools perceived that the education program developed their generic skills more than students from better quality schools. Furthermore, they found that male and African students had lower perceptions of the development of generic skills than female and white students. However, no statistically significant differences were found between students with different home languages (Kirstein et al., 2019).

Specific Generic Skills

Critical Thinking

The findings relating to learning and development of critical thinking skills were contradictory, depending on the study design, methods and sample size. For example, Al-Thani et al. (2016) found that senior students performed better in a thinking test than junior students. However, Ding et al. (2016) did not find differences across different study years, across fields, or across university tiers. Kim (2015) reported in her case study that both graduate and doctoral students tended to show low critical thinking under minimal and enhanced scaffolds. Sotiriadou and Hill (2015) found that students reported some improvement in their critical thinking. However, at the same time, the most versatile levels of critical thinking were challenging to develop (Sotiriadou and Hill, 2015). Danczak et al. (2017) found some development of critical thinking during a course, but it seems that their findings could be explained by the time that the students used in completing their test. In sum, based on the studies covered here, it seems that the development of critical thinking is uncertain or limited (Kim, 2015; Ding et al., 2016; Danczak et al., 2017).

Communication Skills

Studies on communication skills focused on both oral and written communication. Students were found to manage oral communication skills better than their counterparts in working life (Tun Lee-Foo et al., 2015). It was also found that third-year students perceived significantly higher levels of improvement of oral communication skills than students in the first or second year of studies (Mercer-Mapstone and Matthews, 2017).

Several articles reported improvement in students' scientific writing skills during their studies (Drury and Muirb, 2014; Hryciw and Dantas, 2016; Pöysä-Tarhonen et al., 2016; Rayner et al., 2016). Physiology students were found to improve their performance especially in writing laboratory reports, comparing information from different sources, proposing further experiments, constructing logical arguments, interpreting results, as well as writing hypotheses, introductions, discussions, and conclusions (Drury and Muirb, 2014).

Collaboration Skills

The studies that focused on collaboration skills emphasized the importance of collaboration (Chydenius and Gaisch, 2016; Salleh et al., 2016, 2017) and teamwork skills (García et al., 2016). Many studies found that the students in higher education developed a good level of performance with regard to teamwork skills (Cela-Ranilla et al., 2014b; Tynjälä et al., 2016; Sridharan et al., 2018; Christensen et al., 2019). Bravo et al. (2016) found that students perceived improvement in their understanding of how teams work.

Creativity and Problem-Solving Skills

The studies that focused on exploring students' learning and level of the skills showed contradictory results. Some studies showed that the students had learnt problem-solving and creativity skills well during their degrees (Klegeris et al., 2017; Tahir et al., 2017; López et al., 2019), whereas in some studies this was true only to a certain extent (Calma, 2017; Techanamurthy et al., 2018). Some studies explored whether the learning of these skills could be enhanced with various pedagogical approaches. Most of the studies indicated that the learning of problem-solving and creativity skills can be positively enhanced (Wood and Bilsborow, 2014; Mareque et al., 2019; Seow et al., 2019). An exception to this was a study where only some of the students felt that their skills had improved, whereas others did not (Keinänen and Kairisto-Mertanen, 2019).

Self-Regulation Skills

Studies focusing on self-regulation, self-management, and self-monitoring showed that students were learning these skills. First-year students reported learning time management, learning skills, and self-monitoring skills (Mah and Ifenthaler, 2018). It was also found that senior students report higher performance in self-management skills compared to freshmen (Cela-Ranilla et al., 2014a; Tseng et al., 2019). The students developed a good level of performance with regard to self-management (Cela-Ranilla et al., 2014a).

Ethical Skills

There were some studies that investigated student learning of ethical skills during higher education, and a variety of concepts were utilized. Students in teacher education studies were found to develop in terms of their social responsibility skills (Howells et al., 2016) as well as scholarship and moral citizenship (Steur et al., 2016).

Factors Enhancing or Impeding Student Learning of Generic Skills

The third aim of this review study was to identify factors that enhance or impede student learning of generic skills. First, the enhancing and impeding factors of the studies focusing on sets of generic skills are presented, followed by the results of the studies focusing on specific generic skills.

Sets of Generic Skills

Both enhancing and impeding factors were identified in the studies that focused broadly on sets of generic skills. Most of the studies highlighted that good and well-organized teaching (Boahin and Hofman, 2014; Guo et al., 2017) and various active learning methods, such as project-based learning (Dinning, 2017; Lee et al., 2019), problem-based learning (Bautista, 2016; Joseph et al., 2016; Martínez-Clares and González-Morga, 2018; Adriaensen et al., 2019; Deep et al., 2019), cooperative learning (El Tantawi et al., 2014; Canelas et al., 2017; Kridiotis and Swart, 2017; Larraz et al., 2017; Martínez-Clares and González-Morga, 2018), flipped classroom (Ng, 2016; Canelas et al., 2017; Guo, 2019), and workshops (Krassadaki et al., 2014; Sarkar et al., 2017) enhanced the learning of generic skills. It was found that students' generic skills developed in disciplinary courses that intentionally integrated the learning of generic skills (Windsor et al., 2014; Rocha, 2015). Additionally, satisfaction with the guidance (Skaniakos et al., 2019), group work (Prokofieva et al., 2015), peer interaction (Guo, 2018), interaction with tutor, and defining the teamwork rules (Carvalho, 2016) were positively related to generic skills learning. Positive course experiences, including appropriate workload, good teaching, clear goals and standards, and emphasis on independence were related to positive evaluations of generic skills development (Liu et al., 2017). In addition, constructively aligned and continuous assessment was found to be positively related to the learning of generic skills (Murdoch-Eaton et al., 2016; Ruge and McCormack, 2017). Peer assessment, feedback, general study guidance, and portfolio (Adriaensen et al., 2019) or other reflection tasks (Tomasson Goodwill et al., 2019) also enhanced the learning of generic skills.

Some studies found that games (Fitó-Bertran et al., 2015; Hermnandez-Lara et al., 2018), role playing (El Tantawi et al., 2014), business simulations (Kelton and Kingsmill, 2016; Levant et al., 2016; Buil et al., 2018), and online tools or competitions (Viviers, 2016; Abdulwahed and Hasna, 2017) enhanced the learning of generic skills. Some studies showed that different kinds of work-integrated learning environments enhanced the learning of generic skills (e.g., Jackson, 2015). For example, work-integrated learning curricula (Jackson, 2015; Smith and Worsfold, 2015; Rambe, 2018), work experience and internships (Levant et al., 2016; Bellew and Gabaudan, 2017; Sonnenschein et al., 2017), service learning (Kao et al., 2014), and workplace simulations (Bautista-Mesa et al., 2018) were perceived to enhance student learning of generic skills. The importance of a mentor during work-integrated learning was highlighted in a few studies (Jackson, 2015; Bellew and Gabaudan, 2017). Furthermore, social media use for employment purposes was positively related to generic skills and internship served as a mediating mechanism through which social media use affects generic skills (He et al., 2017).

Students' own personal activities also contributed to the learning of generic skills (Ssegawa and Kasule, 2017). Student engagement (Guo, 2018), deep approach to learning, interest, and flow experiences (Buil et al., 2018) were mentioned as promoting factors. A few studies also found that higher initial skills levels was a promoting factor for learning more new skills during the academic year compared to those whose initial skills levels were lower (Feldon et al., 2016) and for students' entrepreneurial intentions (Bonesso et al., 2018).

Some of the studies that we reviewed identified factors that impede or challenge student learning of generic skills. Most of the impeding factors were associated with the learning environment. More precisely, teacher-focused instruction (Guo, 2018), students' passive role in teaching (Guo, 2018), lack of teacher-student interaction (Guo et al., 2017), and overly rapid pace of teaching impeded the learning of generic skills (Viviers, 2016). Poor working life and practice experiences as well as mismatches between employers' and students' expectations were also found to be challenging factors for generic skills development. Tran (2017) found five inhibiting factors: students' working part-time, a lack of information about extra-curricular activities, students' beliefs about participating bringing no benefits, competition with curriculum-based activities, and unprofessional organization of these activities. Additionally, it was shown that students' surface approach to learning (Guo et al., 2017), surface motives and poor study strategies (Yin et al., 2016) were related to their poor learning of generic skills.

Specific Generic Skills

Critical Thinking

Asking students about their experiences on factors that enhance their learning of critical thinking, one study found that inquiry-based learning methods were helpful (Danczak et al., 2017). It has also been suggested that instruction that takes critical thinking into account could be a powerful tool for enhancing students' level of critical thinking (Al-Thani et al., 2016). For example, scaffolding and sequential assignments have been found to improve students' critical thinking skills in some studies (Sotiriadou and Hill, 2015) but not always (Kim, 2015). Research on performance-based assessment has shown that students' primary knowledge enhances performance and motivation in reasoning (Lespiau and Tricot, 2018).

Communication Skills

In several studies, various e-learning resources were found to enhance students' written communication skills. A specific e-learning environment that provides resources for learning discipline-specific content and writing was found to improve both students' written communication skills and content understanding (Drury and Muirb, 2014). A scaffolded learning approach including both online writing tasks and active-learning lectures, small-group discussions, and collaborative workshops improved students' scientific literacy skills (Hryciw and Dantas, 2016). Additionally, role models in terms of communication skills, feedback on performance (Van Ginkel et al., 2015),

mentoring, and peer collaboration were found to be influential factors for student learning (Jackson, 2014b, 2016b). Also, explicit teaching of science communication skills embedded in courses was found to be influential (Mercer-Mapstone and Kuchel, 2016).

Collaboration Skills

Studies focusing on collaboration skills indicated that factors related to teaching and learning environments were found to enhance the learning of generic skills. Team-based learning in accounting courses enhanced student perceptions of their ability to work effectively in diverse teams, as well as other teamwork abilities such as cultural diversity, leadership and planning, and implementation (Christensen et al., 2019). Students were found to learn collaboratively when working on their study task in a culturally mixed small group (Daly et al., 2015). Bravo et al. (2016) showed that teamwork processes have significant effects on improvements in teamwork skills, and thus teachers should use assignments that require managing these teamwork processes rather than focusing solely on the success of the assignment. Students perceived six factors that contribute to positive student teamwork experiences: shared team goals; cultural diversity; adaptable work skills; challenging task context; collaborative research; cross-functional teams (Volkov and Volkov, 2015). Sridharan et al. (2018) found that peer assessment improved collaboration skills. Digital games provide an excellent online learning environment for students to work in and improve their teamwork skills (Cela-Ranilla et al., 2014b). Online learning environments utilizing problem-based learning, and providing versatile support and encouragement for continuous assessment, were reported to enhance students' teamwork skills (García et al., 2016). Work-integrated learning helped undergraduates to develop their interpersonal skills (Ibrahim and Jaaffar, 2017a).

Creativity and Problem-Solving Skills

Most of the studies explored the effect of implementing different pedagogical approaches or interventions to enhance students' learning of problem-solving, innovation, and creativity skills. For example, a design-based research approach was found to improve students' creativity skills (Wood and Bilsborow, 2014). Innovation pedagogy enhanced students' learning of different innovation competences, and introducing an experiential learning pedagogy was found to improve students' problem-solving skills (Seow et al., 2019). Various arts-related leisure activities were found to be positively related to creativity (Mareque et al., 2019). Incorporating generic skills (including creativity and problem-solving) within curricula and academic courses was found to be correlated with students' satisfaction in learning those skills (Tahir et al., 2017). Some studies also indicate that students' problem-solving skills evolve along with university experience, further suggesting that some instructional methods might be especially beneficial in enhancing the learning of those skills (such as problem-based learning, case studies, team-based learning) as opposed to traditional lecture-style courses (Klegeris et al., 2017).

Self-Regulation Skills

Several enhancing factors of self-regulation skills were indicated by the studies. The use of teaching and learning materials improved the attitude of learners to the development of self- and social competencies (Edeling and Pilz, 2016). In addition, work-related factors such as work-integrated programs (Ibrahim and Jaaffar, 2017b) and the training company enhance students' learning of self-management skills (Edeling and Pilz, 2016). Furthermore, it was also found that a 3D simulation learning environment and digital games (Cela-Ranilla et al., 2014a,b) enhanced student learning of self-management skills.

Ethical Skills

It was found that reflective writing tasks as well as other learning and assessment experiences provided during the course enhanced student teachers' learning of social responsibility skills (Howells et al., 2016). Taplin et al. (2018) found that use of role-play enhanced student learning of ethical skills.

DISCUSSION

Our systematic review study contributes to the research on generic skills by structuring the current research in the field, elaborating the concepts and theories related to learning generic skills, and clarifying the methods utilized in the empirical studies. The study revealed the remarkable variation in concepts and their definitions, research methods, and the way generic skills were measured. The conceptual variation manifested itself in many different ways. Most of the reviewed studies investigated sets of generic skills and used the term generic skills or other similar concept, such as employability skills, transferable skills, soft skills, graduate attributes, or generic competencies. These results reflect those of Lizzio et al. (2002) and Barrie (2006), who also found that generic skills are known by several other terms. The number of generic skills explored ranged from one or two skills to several dozen. The present study thus clearly shows that "generic skills" is used as an umbrella term, which can include various wide-ranging skills. Some of the articles framed their research with generic skills but focused more specifically on individual specific generic skills. The studies exploring specific skills had their focus on one of six generic skills. These skills were critical thinking, communication skills, collaboration skills, creativity and problem-solving skills, self-regulation skills, and ethical skills. Similar skills have been found in a previous review study that explored generic competences and found that the most frequently appearing generic competences were a set of conceptual skills, people skills, and personal skills (Strijbos et al., 2015).

It was somewhat surprising that the studies that focused on sets of generic skills most often measured professional skills such as professionalism and leadership skills. These skills are not higher-order thinking skills, which are outlined as the key skills and aims of higher education (Strijbos et al., 2015; OECD, 2019). The high amount of professional skills in the articles studied may be due to the emphasis on working life. Consequently, in many studies the learning of generic skills was justified by the need for these skills in working life. After the professional skills, analytical skills, applying knowledge, communication, and collaboration were most often operationalized as generic skills in the surveys. These skills can be considered higher-order thinking skills and important for professionals in various fields. There is surprisingly

little research on generic skills and their relation to learning processes, although these skills are needed in quality learning and studying (Badcock et al., 2010; Arum and Roksa, 2011; Tuononen and Parpala, 2021). In addition, these skills are important for lifelong learning.

A more accurate analysis of the articles focusing on sets of generic skills showed inconsistency in the instruments used. Almost every study introduced its own survey instrument to measure generic skills. In these studies, the operationalization of the measured skills was often incoherent, and they failed to give an explicit definition of generic skills. The present review study confirmed the previous findings, which have demonstrated several problems in surveys in the research field of generic skills. For example, abstract or vague expressions and double-barreled items in the questionnaires have been found (Braun et al., 2012).

Most of the studies in this review used self-report methods with a cross-sectional study design. The studies with a longitudinal design focused mostly on a short period of time, e.g., one course or one semester. While the methods chosen may reflect a lack of long-term research resources, more thought should be put into methods to capture actual skills and their development. Self-report measures only capture students' perceptions and experiences, while performance-based assessments would enable a deeper understanding of students' competency (Zlatkin-Troitschanskaia et al., 2015). Furthermore, while cross-sectional studies do not inform us about the development of generic skills, even longitudinal designs that focus on short periods of time can provide inaccurate information about actual development of skills. The learning of generic skills takes time (Arum and Roksa, 2011; Hyytinen et al., 2019; Muukkonen et al., 2019), and such designs may not be able to capture the development, or the development that they capture may not be lasting. Additionally, only a few studies used performance-based assessments to explore generic skills. This was somewhat surprising, since investigating skills with performance-based assessment would be ontologically and methodologically reasonable (McClelland, 1973; Ercikan and Oliveri, 2016; Hyytinen et al., 2021). Additionally, more performance-based assessments of generic skills with larger data sets are needed (Al-Thani et al., 2016). One another methodological aspect relates to the level of analysis. Most of the articles utilize group-level analysis, which may not reveal individual variation in perceptions of learning generic skills.

Conceptual and methodological shortcomings make it difficult to compare studies, and build a cumulative understanding about the status of generic skills in higher education. Additionally, each of the studies focuses on different sets of skills, which complicates the matter. The studies that focus on specific individual skills such as critical thinking, problem-solving or collaboration skills are often more advanced in their theoretical and conceptual framework, as well as methods, compared with studies that focus on a varied set of skills. This is probably due to the conceptual clarity in the field of the respective skill, e.g., critical thinking. In order to contribute to higher education research, studies of generic skills need to strive for increased clarity and coherence.

Our aim was also to explore how higher education students have learned generic skills during their studies. A coherent picture of the learned generic skills is relatively challenging to capture because the articles have focused on different sets of generic skills with various surveys. Although students in many studies perceived that they had learned generic skills well, some studies indicated that their learning of generic skills was limited. Additionally, some studies indicated that there were differences between the students in learning generic skills, for example, regarding their discipline. Furthermore, it is also important to remember that most of the studies reviewed mainly explored students' own experiences in learning generic skills, not their actual level of generic skills (cf. Braun et al., 2012; Zlatkin-Troitschanskaia et al., 2015). However, it is noteworthy that there was also contradictory evidence about students' generic skills learning based on the studies that used performance-based assessment and whether these skills develop during studies.

The present review study also identified enhancing and impeding factors that were found to be associated with learning generic skills in the studies. The results indicated that most of these factors were contextual, relating to the teaching and learning environment, rather than focusing on individual factors. Active learning methods that emphasize students' activity and role in the learning process were most often found to be enhancing factors. In addition, the role of different digital learning environments such as games and online tools in the learning of generic skills was investigated, and they were usually perceived as useful. Work-based learning and work-related projects were also perceived as useful. It was interesting that previous knowledge and initial skill level were related to the learning of new generic skills. This finding supports evidence from previous studies (e.g., Richardson et al., 2012). In addition, students' own personal activities, such as student engagement, deep approach to learning, and interest were individual factors that were found to enhance the learning of generic skills (cf. Arum and Roksa, 2011). The impeding factors were also mainly associated with the learning environment. For example, teacher-focused instruction, lack of interaction, and poor working life and practice experiences were found to be negatively related to the learning of generic skills.

LIMITATIONS

Some limitations of this review study have been identified. First, these articles were searched for using the term generic skills, not specific skills. Therefore, the sample of critical thinking studies, and certain other skills, does not represent all research in the field. Second, we only included articles for 5 years in the analysis, and thus the sample does not comprehensively describe the research of generic skills and how generic skills research has developed during the 2010s. The number of studies published was so large that we were not able to include more years in this review study. Third, the enhancing and impeding factors found in the studies were based mainly on self-reports. Thus, how much they actually enhance or impede the learning of generic skills has not been explored in these studies.

CONCLUSIONS

The present study shows that there is a lot of research activity in this area, indicating the importance and relevance of generic skills research. To ensure the development of research on generic skills, it is essential to enhance the dialogue between theoretical, methodological, and empirical perspectives to extend previous work in the field. The results of the present study demonstrate that the challenges in exploring generic skills are both methodological and theoretical in nature (cf. Barrie, 2006; Braun et al., 2012; El Soufi and See, 2019). The problem is that the results do not accumulate because so many different theoretical frameworks, concepts, definitions, and instruments are used. Therefore, we suggest that existing valid instruments should be utilized when new studies are constructed. In this way, the definition of concepts will become clearer and valid instruments will evolve. Generic skills can be explored using self-reports if valid instruments are used. In addition, self-reports can be used to develop students' reflection skills and help students to recognize and evaluate their generic skills (Kyndt et al., 2014). However, this review study showed that intervention and longitudinal studies are needed but such study designs are demanding and require greater resources. In the future, it would be interesting to explore how the learning of generic skills progresses during studies and how a high level of certain skills can promote the learning of other skills. This review study advances new research on higher education student learning of generic skills and also contributes to the practical development of teaching and learning in higher education by indicating the enhancing and impeding factors.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

TT, HH, and AT contributed to conception and design of the study. TT and HH conducted literature search. TT, HH, AT, and KK wrote other sections of the manuscript. All authors analyzed and wrote the results. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported by the Finnish Ministry of Education and Culture (Grant Number: OKM/280/523/2017). University of Helsinki, Finland has covered the article processing fee.

REFERENCES

Abayadeera, N., and Watty, K. (2016). Generic skills in accounting education in a developing country: exploratory evidence from Sri Lanka. *Asian Rev. Account.* 24, 1–30. doi: 10.1108/ARA-03-2014-0039

Abdulwahed, M., and Hasna, M. O. (2017). The role of engineering design in technological and 21st century competencies capacity building: comparative case study in the Middle East, Asia, and Europe. *Sustainability* 9, 520. doi: 10.3390/su9040520

Adriaensen, J., Bijsmans, P., and Groen, A. (2019). Monitoring generic skills development in a bachelor european studies. *J. Contemp. Eur. Res.* 15, 110–127. doi: 10.30950/jcer.v15i1.1018

Akhmetshin, E. M., Larionova, G. N., Lukiyanchina, E. V., Savitskaya, Y. P., Aleks, R., and Aleynikova, O. S. (2019). The influence of educational environment on the development of entrepreneurial skills and competencies in students. *J. Entrepreneurship Educ.* 22, 1S.

Al-Thani, S., Abdelmoneim, A., Cherif, A., Moukarzel, D., and Daoud, K. (2016). Assessing general education learning outcomes at Qatar University. *J. Appl. Res. Higher Educ.* 8, 159–176. doi: 10.1108/JARHE-03-2015-0016

Arum, R., and Roksa, J. (2011). *Academically Adrift: Limited Learning on College Campuses*. Chicago: University of Chicago Press.

Badcock, P. B. T., Pattison, P. E., and Harris, K.-L. (2010). Developing generic skills through university study: a study of arts, science and engineering in Australia. *Higher Educ.* 60, 441–458. doi: 10.1007/s10734-010-9308-8

Barrie, S. C. (2006). Understanding what we mean by the generic attributes of graduates. *Higher Educ.* 51, 215–241. doi: 10.1007/s10734-004-6384-7

Bautista, I. (2016). Generic competences acquisition through classroom activities in first-year agricultural engineering students. *Int. J. Educ. Technol. Higher Educ.* 13, 29. doi: 10.1186/s41239-016-0028-8

Bautista-Mesa, R., Molina Sánchez, H., and Ramírez Sobrino, J. N. (2018). Audit workplace simulations as a methodology to increase undergraduates' awareness of competences. *Account. Educ.* 27, 234–258. doi: 10.1080/09639284.2018.1476895

Bellew, L., and Gabaudan, O. (2017). An investigation into the development and progressive adaptation of graduate attributes in tourism programmes. *J. Teach. Travel Tourism* 17, 139–158. doi: 10.1080/15313220.2017.1318104

Boahin, P., and Hofman, W. H. A. (2014). Perceived effects of competency-based training on the acquisition of professional skills. *Int. J. Educ. Dev.* 36, 81–89. doi: 10.1016/j.ijedudev.2013.11.003

Bonesso, S., Gerli, F., and Pizzi, C. (2015). The interplay between experiential and traditional learning for competency development. *Front. Psychol.* 6, 1305. doi: 10.3389/fpsyg.2015.01305

Bonesso, S., Gerli, F., Pizzi, C., and Cortellazzo, L. (2018). Students' entrepreneurial intentions: the role of prior learning experiences and emotional, social, and cognitive competencies. *J. Small Business Manage.* 56, 215–242. doi: 10.1111/jsbm.12399

Braun, E., Woodley, A., Richardson, J. T. E., and Leidner, B. (2012). Self-rated competences questionnaires from a design perspective. *Educ. Res. Rev.* 7, 1–18. doi: 10.1016/j.edurev.2011.11.005

Bravo, R., Lucia-Palacios, L., and Martin, M. J. (2016). Processes and outcomes in student teamwork. An empirical study in a marketing subject. *Stud. Higher Educ.* 41, 302–320. doi: 10.1080/03075079.2014.926319

Buil, I., Catalán, S., and Martínez, E. (2018). Exploring students' flow experiences in business simulation games. *J. Comput. Assisted Learn.* 34, 183–192. doi: 10.1111/jcal.12237

Burch, V. C., Sikakana, C. N. T., Gunston, G. D., and Murdoch-Eaton, D. (2018). Self-reported generic learning skills proficiency: another measure of medical school preparedness. *Afr. J. Health Professions Educ.* 10, 114–123. doi: 10.7196/AJHPE.2018.v10i2.971

Calma, A. (2017). The long and winding road: problems in developing capabilities in an undergraduate commerce degree. *Int. J. Educ. Manage.* 31, 418–429. doi: 10.1108/IJEM-09-2015-0122

Canelas, D., Hill, J., and Novicki, A. (2017). Cooperative learning in organic chemistry increases student assessment of learning gains in key transferable skills. *Chem. Educ. Res. Prac.* 18, 441–456. doi: 10.1039/C7RP00014F

Carvalho, A. (2016). The impact of PBL on transferable skills development in management education. *Innov. Educ. Teach. Int.* 53, 35–47. doi: 10.1080/14703297.2015.1020327

Cela-Ranilla, J. M., Esteve-Gonzalez, V., Esteve-Mon, F., and Gisbert-Cervera, M. (2014a). 3D simulation as a learning environment for acquiring the skill of self-management: an experience involving spanish university students of education. *J. Educ. Comput. Res.* 51, 295–309. doi: 10.2190/EC.51.3.b

Cela-Ranilla, J. M., Esteve-Mon, F. M., Esteve-González, V., and Gisbert-Cervera, M. (2014b). Developing self-management and teamwork using digital games in 3D simulations. *Austral. J. Educ. Technol.* 30, 634–651. doi: 10.14742/ajet.754

Chan, C. K., and Fong, E. T. (2018). Disciplinary differences and implications for the development of generic skills: a study of engineering and business students' perceptions of generic skills. *Euro. J. Eng. Educ.* 43, 927–949. doi: 10.1080/03043797.2018.1462766

Christensen, J., Harrison, J. L., Hollindale, J., and Wood, K. (2019). Implementing team-based learning (TBL) in accounting courses. *Accoun. Educ.* 28, 195–219. doi: 10.1080/09639284.2018.1535986

Chydenius, T., and Gaisch M. (2016). Work-life Interaction Skills: An exploration of definitional and functional perspectives within the austrian and finnish ICT industry. *Bus. Perspect. Res.* 4, 169–181. doi: 10.1177/2278533716642654

Conchado, A., Carot, J. M., and Bas, M. C. (2015). Competencies for knowledge management: development and validation of a scale. *J. Knowl. Manage.* 19, 836–855. doi: 10.1108/JKM-10-2014-0447

Daly, A., Hoy, S., Hughes, M., Islam, J., and Mak, A. S. (2015). Using group work to develop intercultural skills in the accounting curriculum in Australia. *Accoun. Educ.* 24, 27–40. doi: 10.1080/09639284.2014.996909

Danczak, S. M., Thompson, C. D., and Overton, T. L. (2017). 'What does the term Critical Thinking mean to you?'A qualitative analysis of chemistry undergraduate, teaching staff and employers views of critical thinking. *Chem. Educ. Res. Prac.* 18, 420–434. doi: 10.1039/C6RP00249H

Deep, S., Salleh, B. M., and Othman, H. (2019). Study on problem-based learning towards improving soft skills of students in effective communication class. *Int. J. Innov. Learn.* 25, 17–34. doi: 10.1504/IJIL.2019.096512

Ding, L., Wei, X., and Mollohan, K. (2016). Does higher education improve student scientific reasoning skills?. *Int. J. Sci. Math. Educ.* 14, 619–634. doi: 10.1007/s10763-014-9597-y

Dinning, T. (2017). Embedding employability and enterprise skills in sport degrees through a focused work-based project; a student and employer viewpoint. *Cogent Educ.* 4, 1387085. doi: 10.1080/2331186X.2017.1387085

Drury, H., and Muirb, M. (2014). Using an E-learning environment for developing science students' written communication: the case of writing laboratory reports in physiology. *Int. J. Innov. Sci. Math. Educ.* 22, 79–93.

Edeling, S., and Pilz, M. (2016). Teaching self- and social competencies in the retail sector findings from vocational schools in Germany, Italy and Poland. *Educ. Train.* 58, 1041–1059. doi: 10.1108/ET-07-2015-0060

El Soufi, N., and See, B. H. (2019). Does explicit teaching of critical thinking improve critical thinking skills of English language learners in higher education? A critical review of causal evidence. *Stud. Educ. Eval.* 60, 140–162. doi: 10.1016/j.stueduc.2018.12.006

El Tantawi, M. M., Abdelaziz, H., AbdelRaheem, A. S., and Mahrous, A. A. (2014). Using peer-assisted learning and role-playing to teach generic skills to dental students: the health care simulation model. *J. Dent. Educ.* 78, 85–97. doi: 10.1002/j.0022-0337.2014.78.1.tb05660.x

Elo, S., and Kyngäs, H. (2007). The qualitative content analysis process. *J. Adv. Nurs.* 62, 107–115. doi: 10.1111/j.1365-2648.2007.04569.x

Ercikan, K., and Oliveri, M. E. (2016). In search of validity evidence in support of the interpretation and use of assessments of complex constructs: discussion of research on assessing 21st century skills. *Appl. Measure. Educ.* 29, 310–318. doi: 10.1080/08957347.2016.1209210

European Parliament Council (2008). *The Establishment of the European Qualifications Framework for Lifelong Learning*. Official Journal of European Union. Available online at: https://www.cedefop.europa.eu/en/projects/european-qualifications-framework-eqf

Feldon, D. F., Maher, M. A., Roksa, J., and Peugh, J. (2016). Cumulative advantage in the skill development of STEM graduate students: a mixed-methods study. *Am. Educ. Res. J.* 53, 132–161. doi: 10.3102/0002831215619942

Fitó-Bertran, A., Hernández-Lara, A. B., and López, E. S. (2015). The effect of competences on learning results in an educational experience with a business simulator. *Comput. Human Behav.* 51, 910–914. doi: 10.1016/j.chb.2014.11.003

García, M. G., López, C. B., Molina, E. C., Casas, E. E., and Morales, Y. A. R. (2016). Development and evaluation of the team work skill in university contexts. Are virtual environments effective? *Int. J. Educ. Technol. Higher Educ.* 13, 1–11. doi: 10.1186/s41239-016-0014-1

Guo, J. (2018). Building bridges to student learning: perceptions of the learning environment, engagement, and learning outcomes among Chinese undergraduates. *Stud. Educ. Eval.* 59, 195–208. doi: 10.1016/j.stueduc.2018.08.002

Guo, J. (2019). The use of an extended flipped classroom model in improving students' learning in an undergraduate course. *J. Comput. Higher Educ.* 31, 362–390. doi: 10.1007/s12528-019-09224-z

Guo, J., Yang, L., and Shi, Q. (2017). Effects of perceptions of the learning environment and approaches to learning on Chinese undergraduates' learning. *Stud. Educ. Eval.* 55, 125–134. doi: 10.1016/j.stueduc.2017.09.002

He, C., Gu, J., Wu, W., Zhai, X., and Song, J. (2017). Social media use in the career development of graduate students: the mediating role of internship effectiveness and the moderating role of Zhongyong. *Higher Educ.* 74, 1033–1051.

Hermnandez-Lara, A. B. M., Serradell-Lopez, E., and Fito-Bertran, A. (2018). Do business games foster skills? A cross-cultural study from learners' views. *Intangible Capital* 14, 315–331. doi: 10.3926/ic.1066

Howells, K., Fitzallen, N., and Adams, C. (2016). Using assessment to develop social responsibility as a graduate attribute in teacher education. *Austral. J. Teacher Educ.* 41, 52–67. doi: 10.14221/ajte.2016v41n6.4

Hryciw, D. H., and Dantas, A. M. (2016). Scaffolded research-based learning for the development of scientific communication in undergraduate physiology students. *Int. J. Innov. Sci. Math. Educ.* 24, 1–11.

Hyytinen, H., Toom, A., and Shavelson, R. (2019). "Enhancing scientific thinking through the development of critical thinking in higher education," in *Redefining Scientific Thinking for Higher Education*, eds M. Murtonen and K. Balloo (Cham: Palgrave Macmillan), 59–78.

Hyytinen, H., Ursin, J., Silvennoinen, K., Kleemola, K., and Toom, A. (2021). The dynamic relationship between response processes and self-regulation in critical thinking assessments. *Stud. Educ. Eval.* 1–12. doi: 10.1016/j.stueduc.2021.101090

Ibrahim, H. I., and Jaaffar, A. H. (2017a). Investigating post-work integrated learning (WIL) effects on motivation for learning: an empirical evidence from Malaysian public universities. *Int. J. Business Soc.* 18, 13–32. doi: 10.33736/ijbs.487.2017

Ibrahim, H. I., and Jaaffar, A. H. (2017b). The outcomes of work-integrated learning programmes: the role of self-confidence as mediator between interpersonal and self-management skills and motivation to learn. *Soc. Sci. Humanities* 25, 931–348.

Jackson, D. (2014a). Testing a model of undergraduate competence in employability skills and its implications for stakeholders. *J. Educ. Work* 27, 220–242. doi: 10.1080/13639080.2012.718750

Jackson, D. (2014b). Business graduate performance in oral communication skills and strategies for improvement. *Int. J. Manage. Educ.* 12, 22–34. doi: 10.1016/j.ijme.2013.08.001

Jackson, D. (2015). Employability skill development in work-integrated learning: barriers and best practice. *Stud. Higher Educ.* 40, 350–367. doi: 10.1080/03075079.2013.842221

Jackson, D. (2016a). Skill mastery and the formation of graduate identity in Bachelor graduates: evidence from Australia. *Stud. Higher Educ.* 41, 1313–1332. doi: 10.1080/03075079.2014.981515

Jackson, D. (2016b). Modelling graduate skill transfer from university to the workplace. *J. Educ. Work* 29, 199–231. doi: 10.1080/13639080.2014.907486

Joseph, N., Rai, S., Madi, D., Bhat, K., Kotian, S. M., and Kantharaju, S. (2016). Problem-based learning as an effective learning tool in community medicine: initiative in a private medical college of a developing country. *Indian J. Commun. Med.* 41, 133–140. doi: 10.4103/0970-0218.177535

Kao, H.-Y., Wang, Y.-T., Huang, C.-H., Lai, P.-L., and Chen, J.-Y. (2014). Assessment and classification of service learning: a case study of CS/EE students. *Sci. World J.* 1–8. doi: 10.1155/2014/183732

Keinänen, M. M., and Kairisto-Mertanen, L. (2019). Researching learning environments and students' innovation competences. *Educ. Train.* 61, 17–30. doi: 10.1108/ET-03-2018-0064

Kelton, M., and Kingsmill, V. (2016). Simulations for the discipline specific and professional education of foreign policy graduates. *J. Univ. Teach. Learn. Pract.* 13, 1–15. doi: 10.53761/1.13.5.7

Kim, N. (2015). Critical thinking in wikibook creation with enhanced and minimal scaffolds. *Educ. Technol. Res. Dev.* 63, 5–33. doi: 10.1007/s11423-014-9361-6

Kirstein, M., Coetzee, S., and Schmulian, A. (2019). Differences in accounting students' perceptions of their development of professional skills: a South African case. *Higher Educ. Skills Work Based Learn.* 9, 41–59, doi: 10.1108/HESWBL-04-2018-0051

Klegeris, A., McKeown, S. B., Hurren, H., Spielman, L. J., Stuart, M., and Bahniwal, M. (2017). Dynamics of undergraduate student generic problem-solving skills captured by a campus-wide study. *Higher Educ.* 74, 877–896. doi: 10.1007/s10734-016-0082-0

Krassadaki, E., Matsatsinis, N., and Lakiotaki, K. (2014). Adopting a strategy for enhancing generic skills in engineering education. *Eng. Educ. Res.* 28, 185–192. doi: 10.5367/ihe.2014.0206

Kridiotis, C. A., and Swart, S. (2017). A learning development module to support academically unsuccessful 1st-year medical students. *Afr. J. Health Professions Educ.* 9, 62–66. doi: 10.7196/AJHPE.2017.v9i2.694

Kyndt, E., Janssens, I., Coertjens, L., Gijbels, D., Donche, V., and Van Petegem, P. (2014). Vocational education students' generic working life competencies: developing self-assessment instrument. *Vocations Learn.* 7, 365–392. doi: 10.1007/s12186-014-9119-7

Larraz, N., Vázquez, S., and Liesa, M. (2017). Transversal skills development through cooperative learning. Training teachers for the future. *On Horizon* 25, 85–95. doi: 10.1108/OTH-02-2016-0004

Lee, H., Shimotakahara, R., Fukada, A., Shinbashi, S., and Ogata, S. (2019). Impact of differences in clinical training methods on generic skills development of nursing students: a text mining analysis study. *Heliyon* 5, 1–21. doi: 10.1016/j.heliyon.2019.e01285

Lespiau, F., and Tricot, A. (2018). Primary knowledge enhances performance and motivation in reasoning. *Learn. Instruct.* 56, 10–19. doi: 10.1016/j.learninstruc.2018.02.007

Levant, Y., Coulmont, M., and Raluca, S. (2016). Business simulation as an active learning activity for developing soft skills. *Accoun. Educ.* 25, 368–395. doi: 10.1080/09639284.2016.1191272

Liu, J., St. John, K., and Courtier, A. (2017). Development and validation of an assessment instrument for course experience in a general education integrated science course. *J. Geosci. Educ.* 65, 435–454. doi: 10.5408/16-204.1

Lizzio, A., Wilson, K., and Simons, R. (2002). University students' perceptions of the learning environment and academic outcomes: implications for theory and practice. *Stud. Higher Educ.* 27, 27–52. doi: 10.1080/03075070120099359

López, A. R., Souto, J. E., and Noblejas, M. L. A. (2019). Improving teaching capacity to increase student achievement: the key role of communication competences in higher education. *Stud. Educ. Eval.* 60, 205–213. doi: 10.1016/j.stueduc.2018.10.002

Mah, D. K., and Ifenthaler, D. (2018). Students' perceptions toward academic competencies: the case of German first-year students. *Issues Educ. Res.* 28, 120–137.

Mareque, M., de Prada Creo, E., and Gonzalez-Sanchez, M. B. (2019). Fostering creativity and communicative soft skills through leisure activities in management studies. *Educ. Train.* 61, 94–107, doi: 10.1108/ET-07-2018-0149

Martínez-Clares, P., and González-Morga, N. (2018). Teaching methodologies at university and their relationship with the development of transversal competences. *Cultura Educ.* 30, 233–275. doi: 10.1080/11356405.2018.1457610

McClelland, D. C. (1973). Testing for competence rather than for 'intelligence.' *Am. Psychol.* 28, 1–14. doi: 10.1037/h0034092

Mercer-Mapstone, L. D., and Kuchel, L. J. (2016). Integrating communication skills into undergraduate science degrees: a practical and evidence-based approach. *Teach. Learn. Inquiry* 4, 1–14. doi: 10.20343/teachlearninqu.4.2.11

Mercer-Mapstone, L. D., and Matthews, K. E. (2017). Student perceptions of communication skills in undergraduate science at an Australian research-intensive university. *Assessment Eval. Higher Educ.* 42, 98–114. doi: 10.1080/02602938.2015.1084492

Monteiro, S., Almeida, L., and Aracil, A. (2016). Graduates' perceptions of competencies and preparation for labour market transition: the effect of gender and work experience during higher education. *Higher Educ. Skills Work Based Learn.* 6, 208–220. doi: 10.1108/HESWBL-09-2015-0048

Murdoch-Eaton, D., Louw, A. J. N., and Bezuidenhout, J. (2016). Effect of curriculum changes to enhance generic skills proficiency of 1st-year medical students. *Afr. J. Health Professions Educ.* 8, 15–19. doi: 10.7196/AJHPE.2016.v8i1.414

Muukkonen, H., Lakkala, M., Lahti-Nuuttila, P., Ilomäki, L., Karlgren, K., and Toom, A. (2019). Assessing the development of collaborative knowledge work competence: scales for higher education course contexts. *Scand. J. Educ. Res.* 64, 1071–1089. doi: 10.1080/00313831.2019.1647284

Nastiti, D., Rahardio, S. B., Elfi Susanti, V. H., and Perdana, R. (2018). The need analysis of module development based on search, solve, create, and share to increase generic science skills in chemistry. *Indonesian J. Sci. Educ.* 7, 428–434. doi: 10.15294/jpii.v7i4.12393

Nedelova, M., and Šukolova, D. (2017). Critical thinking in initial teacher education: secondary data analysis from Ahelo GS feasibility study in Slovakia. *New Educ. Rev.* 49, 19–29. doi: 10.15804/tner.2017.49.3.01

Ng, E. M. W. (2016). The flipped classroom: two learning modes that foster two learning outcomes. *Issues Informing Sci. Information Technol.* 13, 15–23. doi: 10.28945/3462

OECD (2019). *Education at a Glance 2019. OECD Indicators*. Paris: OECD.

Perdigones, A., Valera, D. L., Moreda, G. P., and García, J. L. (2014). Competences in demand within the Spanish agricultural engineering sector. *Euro. J. Eng. Educ.* 39, 527–538. doi: 10.1080/03043797.2013.766673

Pirog, D. (2016). The role of competencies for geography higher education in university-to-work transition. *Geogr. Pol.* 89, 221–236. doi: 10.7163/GPol.0055

Pita, C., Eleftheriou, M., Fernández-Borrás, J., Gonçalves, S., Mente, E., Santos, M. B., et al. (2015). Generic skills needs for graduate employment in the aquaculture, fisheries and related sectors in Europe. *Aquaculture Int.* 23, 767–786. doi: 10.1007/s10499-014-9843-x

Pop, C., and Khampirat, B. (2019). Self-assessment instrument to measure the competencies of Namibian graduates: testing of validity and reliability. *Stud. Educ. Eval.* 60, 130–139. doi: 10.1016/j.stueduc.2018.12.004

Pöysä-Tarhonen, J., Elen, J., and Tarhonen, P. (2016). Student teams' development over time: tracing the relationship between the quality of communication and teams' performance. *Higher Educ. Res. Dev.* 35, 787–799. doi: 10.1080/07294360.2015.1137887

Prokofieva, M., Jackling, B., and Natoli, R. (2015). A tale of two cohorts: identifying differences in group work perceptions. *Asian Rev. Account.* 23, 68–85. doi: 10.1108/ARA-10-2013-0063

Rambe, P. (2018). Using work integrated learning programmes as a strategy to broaden academic and workplace competencies. *J. Hum. Resour. Manag.* 16, 1–16. doi: 10.4102/sajhrm.v16i0.999

Rayner, G., Papakonstantinou, T., and Gleadow, R. (2016). Comparing the self-efficacy and writing-related abilities of native and non-native English-speaking students. *Cogent Educ.* 3, 1–11. doi: 10.1080/2331186X.2016.1179164

Richardson, M., Abraham, C., and Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychol. Bull.* 138, 353–387. doi: 10.1037/a0026838

Rocha, M. (2015). Predictors of the acquisition and portability of transferable skills: a longitudinal Portuguese case study on education. *Higher Educ.* 69, 607–624. doi: 10.1007/s10734-014-9793-2

Rozlin, R., Ismail, F., Idris, N., Mustaffa, N. E., Saat, M. M., Jamal, N. M., et al. (2018). Generic skills of the undergraduates: a case study of faculty of built environment in Universiti Teknologi Malaysia. *Int. J. Eng. Technol.* 7, 297–302. doi: 10.14419/ijet.v7i2.29.13642

Ruge, G., and McCormack, C. (2017). Building and construction students' skills development for employability – reframing assessment for learning in discipline-specific contexts. *Arch. Eng. Design Manage.* 13, 35–383. doi: 10.1080/17452007.2017.1328351

Salleh, K. M., Subhi, N. I., Sulaiman, N. L., and Latif, A. A. (2016). Generic skills of technical undergraduates and industrial employers perceptions in Malaysia. *Int. J. Appl. Business Econ. Res.* 14, 907–919.

Salleh, K. M., Sulaiman, N. L., Mohamad, M. M., and Sera, L. S. (2017). Assessing soft skills components in science and technology programs within Malaysian. *Songklanakarin J. Sci. Technol.* 39, 399–405. doi: 10.14456/sjst-psu.2017.43

Sarkar, M., Overton, T. L., Thompson, C., and Rayner, G. (2017). Undergraduate science students' perceptions of employability: efficacy of an intervention. *Int. J. Innov. Sci. Math. Educ.* 25, 21–37.

Seow, P. S., Pan, G., and Koh, G. (2019). Examining an experiential learning approach to prepare students for the volatile, uncertain, complex and ambiguous (VUCA) work environment. *Int. J. Manage. Educ.* 17, 62–76. doi: 10.1016/j.ijme.2018.12.001

Shavelson, R. J., Zlatkin-Troitschanskaia, O., Beck, K., Schmidt, S., and Marino, J. P. (2019). Assessment of university students' critical thinking: Next generation performance assessment. *Int. J. Test.* 19, 337–362. doi: 10.1080/15305058.2018.1543309

Skaniakos, T., Honkimäki, S., Kallio, E., Nissinen, K., and Tynjälä, P. (2019). Study guidance experiences, study progress, and perceived learning outcomes of Finnish university students. *Euro. J. Higher Educ.* 9, 203–218. doi: 10.1080/21568235.2018.1475247

Smith, C., and Worsfold, K. (2015). Unpacking the learning–work nexus: 'priming' as lever for high-quality learning outcomes in work-integrated learning curricula. *Stud. Higher Educ.* 40, 22–42. doi: 10.1080/03075079.2013.806456

Sonnenschein, K., Barker, M., and Hibbins, R. (2017). Chinese international students' perceptions of and reflections on graduate attributes needed in entry-level positions in the Chinese hotel industry. *J. Hospitality Tourism Manage.* 30, 39–46. doi: 10.1016/j.jhtm.2017.01.008

Sotiriadou, P., and Hill, B. (2015). Using scaffolding to promote sport management graduates' critical thinking. *Ann. Leisure Res.* 18, 105–122. doi: 10.1080/11745398.2014.925406

Sridharan, B., Muttakin, M. B., and Mihret, D. G. (2018). Students' perceptions of peer assessment effectiveness: an explorative study. *Accoun. Educ.* 27, 259–285. doi: 10.1080/09639284.2018.1476894

Ssegawa, J. K., and Kasule, D. (2017). A self-assessment of the propensity to obtain future employment: a case of final-year engineering students at the University of Botswana. *Euro. J. Eng. Educ.* 42, 513–532. doi: 10.1080/03043797.2016.1193124

Steur, J., Jansen, E., and Hofman, A. (2016). Towards graduateness: Exploring academic intellectual development in university master's students. *Educ. Res. Evaluat*. 22, 6–22. doi: 10.1080/13803611.2016.1165708

Stone, G. A., Duffy, L. N., Pinckney, H. P., and Templeton-Bradley, R. (2017). Teaching for critical thinking: preparing hospitality and tourism students for careers in the twenty-first century. *J. Teach. Travel Tourism* 17, 67–84. doi: 10.1080/15313220.2017.1279036

Strijbos, J., Engels, N., and Struyven, K. (2015). Criteria and standards of generic competences at bachelor degree level: a review study. *Educ. Res. Rev.* 14, 18–32. doi: 10.1016/j.edurev.2015.01.001

Tahir, L. M., Yusof, S. M., Abdul Ghafar, M. D., Omar, W., Samah, N. A., Mohamad, S., et al. (2017). Employability skills policy in heis: are Malaysian graduates from a public technical and engineering-based university contended? *Man India* 97, 1–21.

Taplin, R., Singh, A., Kerr, R., and Lee, A. (2018). The use of short role-plays for an ethics intervention in university auditing courses. *Accoun. Educ.* 27, 383–402. doi: 10.1080/09639284.2018.1475244

Techanamurthy, U., Alias, N., and Dewitt, D. (2018). Problem-solving strategies among culinary arts students in community colleges. *J. Tech. Educ. Train.* 10, 56–70. doi: 10.30880/jtet.2018.10.01.005

Tomasson Goodwill, J., Goh, J., Verkoeyen, S., and Lithgow, K. (2019). Can students be taught to articulate employability skills? *Educ. Train.* 61, 445–460. doi: 10.1108/ET-08-2018-0186

Toom, A., Pyhältö, K., Pietarinen, J., and Soini, T. (2021). Professional agency for learning as a key for developing teachers' competencies? *Educ. Sci.* 11, 1–10. doi: 10.3390/educsci11070324

Tran, L. H. N. (2017). Developing employability skills via extra-curricular activities in Vietnamese universities: student engagement and inhibitors of their engagement. *J. Educ. Work* 30, 854–867. doi: 10.1080/13639080.2017.1349880

Tseng, H., Yi, X., and Yeh, H.-T. (2019). Learning-related soft skills among online business students in higher education: grade level and managerial role differences in self-regulation, motivation, and social skill. *Comput. Human Behav.* 95, 179–186. doi: 10.1016/j.chb.2018.11.035

Tun Lee-Foo, A., Gnanaselvam, P., and Poaw Sim, C. (2015). Communication apprehension among Chinese accounting and business students: a demographic exploration. *Int. J. Manage. Educ.* 9, 161. doi: 10.1504/IJMIE.2015.068760

Tuononen, T., and Parpala, A. (2021). The role of academic competences and learning processes in predicting Bachelor's and Master's thesis grades. *Stud. Educ. Eval.* 70, 101001. doi: 10.1016/j.stueduc.2021.101001

Tuononen, T., Parpala, A., and Lindblom-Ylänne, S. (2019). Graduates' evaluations of usefulness of university education, and early career success - a longitudinal study of the transition to working life. *Assessment Eval. Higher Educ.* 44, 581–595. doi: 10.1080/02602938.2018.1524000

Tynjälä, P., Virtanen, A., Klemola, U., Kostiainen, E., and Rasku-Puttonen, H. (2016). Developing social competence and other generic skills in teacher education: applying the model of integrative pedagogy. *Euro. J. Teacher Educ.* 39, 368–387. doi: 10.1080/02619768.2016.1171314

Utriainen, J., Marttunen, M., Kallio, E., and Tynjälä, P. (2017). University applicants' critical thinking skills: the case of the Finnish educational sciences. *Scand. J. Educ. Res.* 61, 629–649. doi: 10.1080/00313831.2016.1173092

Van Ginkel, S., Gulikers, J., Biemans, H., and Mulder, M. (2015). The impact of the feedback source on developing oral presentation competence. *Stud. Higher Educ.* 1–15. doi: 10.1080/03075079.2015.1117064

Viviers, H. (2016). Qualitative evaluation of the design variables of a teaching intervention to expose accounting students to pervasive skills. *Indus. Higher Educ.* 30, 402–412. doi: 10.1177/0950422216664244

Volkov, A., and Volkov, M. (2015). Teamwork benefits in tertiary education: student perceptions that lead to best practice assessment design. *Educ. Train.* 57, 262–278. doi: 10.1108/ET-02-2013-0025

Windsor, S. A. M., Rutter, K., McKay, D. V., and Meyers, N. (2014). Embedding graduate attributes at the inception of a chemistry major in a bachelor of science. *J. Chem. Educ.* 91, 2078–2083. doi: 10.1021/ed5001526

Wood, D., and Bilsborow, C. (2014). "I am not a person with a creative mind": facilitating creativity in the undergraduate curriculum through a design-based research approach. *Electronic J. e-Learn.* 12, 111–125.

Yin, H., and Ke, Z. (2017). Students' course experience and engagement: an attempt to bridge two lines of research on the quality of undergraduate education. *Assessment Eval. Higher Educ.* 42, 1145–1158. doi: 10.1080/02602938.2016.1235679

Yin, H., Lu, G., and Wang, W. (2014). Unmasking the teaching quality of higher education: students' course experience and approaches to learning in China. *Assessment Eval. Higher Educ.* 39, 949–970. doi: 10.1080/02602938.2014.880107

Yin, H., Wang, W., and Han, J. (2016). Chinese undergraduates' perceptions of teaching quality and the effects on approaches to studying and course satisfaction. *Higher Educ.* 71, 39–57. doi: 10.1007/s10734-015-9887-5

Zlatkin-Troitschanskaia, O., Shavelson, R. J., and Kuhn, C. (2015). The international state of research on measurement of competency in higher education. *Stud. Higher Educ.* 40, 393–411. doi: 10.1080/03075079.2015.1004241

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Education

ORIGINAL RESEARCH
published: 27 May 2022
doi: 10.3389/feduc.2022.886052

What Factors of the Teaching and Learning Environment Support the Learning of Generic Skills? First-Year Students' Perceptions in Medicine, Dentistry and Psychology

Milla Räisänen, Eeva Pyörälä and Tarja Tuononen*

Faculty of Educational Sciences, Department of Education, University of Helsinki, Helsinki, Finland

Future health professions need generic skills in their working lives, such as knowledge analysis, collaboration, communication and problem-solving skills. The teaching and learning environment is crucial in the development of generic skills when studying at university. The aim of this research was to examine students' perceptions of learning generic skills during their first study year and how the teaching and learning environment related to their learning perceptions. The data were collected from first-year students (medicine n = 215, dentistry n = 70 and psychology n = 89) who completed a questionnaire at the end of their first study year. Two cohorts of first-year students from 2020 and 2021 were combined. The teaching and learning environments in medicine, dentistry and psychology differed from each other. The results showed that learning of problem-solving, communication and collaboration skills were emphasized more among medical and dental students, whereas analytical skills more among psychology students. There were no statistically significant differences in perceptions of the teaching and learning environment. Perceptions of generic skills and the teaching and learning environment were positively related to each other. In medicine, the strongest predictors of generic skills were peer support and feedback and in dentistry, peer support, interest and relevance. In psychology, the strongest predictors were interest and relevance. The results emphasize the relevance of the teaching and learning environment in learning generic skills.

Keywords: generic skills, teaching and learning environment, university students, health professions education, higher education, first-year experience

OPEN ACCESS

Edited by:
Mohammed Saqr,
University of Eastern Finland, Finland

Reviewed by:
Jouhanna Menegaz,
Santa Catarina State University, Brazil
Mohammad Khami,
Tehran University of Medical Sciences, Iran

****Correspondence:***
Milla Räisänen
milla.raisanen@helsinki.fi

Specialty section:
This article was submitted to Higher Education, a section of the journal Frontiers in Education

Received: *28 February 2022*
Accepted: *09 May 2022*
Published: *27 May 2022*

Citation:
Räisänen M, Pyörälä E and Tuononen T (2022) What Factors of the Teaching and Learning Environment Support the Learning of Generic Skills? First-Year Students' Perceptions in Medicine, Dentistry and Psychology. Front. Educ. 7:886052. doi: 10.3389/feduc.2022.886052

INTRODUCTION

Learning Generic Skills

Learning academic generic skills, such as analyzing skills, problem-solving, and collaboration and communication skills, in addition to domain specific knowledge and skills, is a key goal of higher education (Tuononen, 2019). Generic skills refer to skills that are general and important in any discipline, although the skills that are required and emphasized in different disciplines vary

(Barrie, 2006). These skills are crucial both for students and their learning in higher education as well as their learning and development throughout their careers. Generic skills are referred to by using several terms, such as key skills, transferable skills, working life skills, core skills, academic competencies, meta-competencies and general characteristics (Barrie, 2006; Tuononen, 2019; Girotto et al., 2021).

There is no coherent definition of generic skills (Barrie, 2006; Chan et al., 2017). The present study focuses on most frequently cited core generic skills: analytical skills, problem-solving skills, collaboration skills, and communication skills (Piróg, 2016; Chan et al., 2017; Liu et al., 2017) which all graduates should achieve and that are important for healthcare professionals (Batalden et al., 2002; Breen et al., 2003; Winston et al., 2012; Joseph et al., 2016; Hamilton et al., 2018; Schot et al., 2020).

University students perceive that collaboration skills develop the least during their studies, even though they are important working life skills (García-Aracil and Van der Velden, 2008; Tuononen et al., 2019a), whereas analyzing skills are well learned at university (Murdoch-Eaton et al., 2016; Tuononen et al., 2019a). Generic skills are also needed in learning, and medical students need them for their academic success (Murdoch-Eaton and Whittle, 2012; Winston et al., 2012; Burch et al., 2018). Therefore, it is important to examine how students perceive the learning of generic skills in health professions education.

Previous research has shown that students learn generic skills better when intertwined with disciplinary content and contexts rather than in separate courses (Bath et al., 2004; Star and Hammer, 2008; Murdoch-Eaton and Whittle, 2012; Virtanen and Tynjälä, 2018). Analytical skills refer to the ability to analyze and critically appraise information, present arguments, and examine things from different perspectives (Winston et al., 2012). In health professions education, analytical and problem-solving skills are needed in clinical reasoning and making decisions about patient care (Monteiro and Norman, 2013; Young et al., 2019; Cooper et al., 2021).

Collaboration and communication skills are crucial for all students, but they are key competences for students graduating from healthcare professions (Cuyvers et al., 2015; Reeves et al., 2015, 2016). Collaboration skills, and in particular, interprofessional collaboration skills are pivotal and should be systematically enhanced during the study years (D'Amour et al., 2005; Bridges et al., 2011; Haddara and Lingard, 2013; Reeves et al., 2017). At the heart of the healthcare professionals' work is good communication with patients, their significant others and the interprofessional team. Indeed, communication skills studies have become an established part of medical education worldwide (Berkhof et al., 2011; Deveugele, 2015; Moura et al., 2021). Specific learning methods have been developed to improve students' communication skills, such as simulated patients followed by reflective feedback discussions (Lane and Rollnick, 2007; Bokken et al., 2009; Cleland et al., 2009; Koponen et al., 2012; El Tantawi et al., 2014; So et al., 2019).

Active and collaborative learning strategies such as problem-based learning (Hmelo-Silver, 2004; Dolmans et al., 2005; Dolmans and Schmidt, 2006; Trullàs et al., 2022), team-based learning (Parmelee and Michaelsen, 2010; Parmelee et al., 2012), case-based learning (Krupat et al., 2016; McLean, 2016) and the flipped classroom (Prober and Khan, 2013; McLaughlin et al., 2014; Hew and Lo, 2018) have been developed to involve students in interactive learning processes in which they collaboratively seek solutions to problems arising from the real-world. What all these learning methods have in common is collaborative learning in small groups, the activation of the learners' existing knowledge, and the application of newly acquired information to the solution of relevant problems and cases. Studies show that problem-based learning (PBL) contributes in many ways to the development of generic skills (Joseph et al., 2016), improves students' communication and collaboration skills, problem-solving skills and development into self-directed learners (Trullàs et al., 2022). However, students need to practice collaboration and teamwork skills to become an active participant in PBL tutorials (Aarnio et al., 2010). One important way to motivate students to learn collaboration skills is to show the connection between collaboration skills and their future working life. Problem-solving skills develop in case-based and problem-based learning (Razzaq and Ahsin, 2011; Gade and Chari, 2012; Karantzas et al., 2013). Critical thinking skills, knowledge acquisition skills (e.g., Joseph et al., 2016; Knipprath, 2017) and the ability to relate knowledge to a range of subjects (basic sciences with preclinical and clinical subjects) improve in active teaching and learning environments (Gade and Chari, 2012). In addition, students perceive that they have learned collaboration and communication skills, the ability to apply clinical reasoning skills, and presentation skills in problem-based learning (Schwartz et al., 1997; Khan and Fareed, 2001). Research among psychology students have shown that generic skills have been learned in working life (Golding et al., 2019). However, generic skills have been examined less among psychology students.

The Interaction Between the Teaching and Learning Environment and Learning Generic Skills

The concept of the teaching and learning environment is used to describe the various elements of the academic environment that support students' quality learning (Entwistle et al., 2003). The teaching and learning environment has been examined through: (1) interest in and relevance of study programs, (2) alignment in teaching, (3) support from other students, and (4) feedback from teachers (Parpala et al., 2010). These elements are related to generic skills and to facilitate students' engagement in deep learning (Entwistle et al., 2002; Parpala et al., 2010; Karagiannopoulou and Milienos, 2018; Utriainen et al., 2018).

A range of elements of the teaching and learning environment support generic skills learning (e.g., Tynjälä et al., 2016; Virtanen and Tynjälä, 2018), such as the use of authentic and complex tasks, combining theory and practice, teacher-student and student-student interaction, peer collaboration, and feedback (Kember and Leung, 2005; Murdoch-Eaton and Whittle, 2012; Virtanen and Tynjälä, 2018). Studies of students' experiences of the teaching and learning environment have shown that peer support encourages their learning the most and constructive feedback and alignment in teaching the least

(Asikainen et al., 2014; Herrmann et al., 2017). The interest and relevance of the subject matter affect students' time and effort management (Parpala et al., 2017). In addition, positive experiences of feedback and support from teachers and other students increase their interest (Hidi and Renninger, 2006) and a stimulating teaching and learning environment leads to high quality learning outcomes (Mikkonen et al., 2009).

The term constructive alignment in teaching and learning refers to instruction in which teachers clearly articulate the intended learning outcomes for students and design learning activities and assessment in a way that directs students toward the learning outcomes (Biggs, 2003). Clear goals and standards (Hongbiao and Zheng, 2017; Liu et al., 2017; Ruge and McCormack, 2017), peer support (El Tantawi et al., 2014; Kridiotis and Swart, 2017), and systematic feedback all foster the development of students' generic skills and improve their interprofessional collaboration skills (Chesluk et al., 2015; McGinness et al., 2019). Unfortunately, medical students have expressed that they get too little feedback about their generic skills (Mubuuke et al., 2016).

Aims of the Study

Still today, little is known about the students' perceptions of the various elements of teaching and learning environment and about the way in which they are related to their perceptions of learning generic skills in health professions education. At the university we are studying, our target groups are students in medicine, dentistry and psychology who study on the same campus. Our research provides keys to how these trainings could be further developed in the health care teaching and learning community.

The aim of this study is to explore how students in medicine, dentistry, and psychology perceive generic skills learning during their first year of study and what are the different dimensions of the teaching and learning environment that are related to generic skills learning. Our specific research questions are as follows:

1) What perceptions do students in medicine, dentistry and psychology have about learning generic skills and their teaching-learning environment at the end of their first study year?
2) What are the differences in perceptions of learning generic skills and teaching-learning environment among students in medicine, dentistry and psychology?
3) What is the relationship between the perceptions of the teaching-learning environment and the learning of generic skills among students in medicine, dentistry and psychology?

MATERIALS AND METHODS

Context

In the current research-intensive university, the Faculty of Medicine educates healthcare professionals, such as physicians, dentists, and psychologists. The degree programs prepare them to become licensed health care professionals. Students in medicine and dentistry pursue a licentiate degree of medicine and dentistry. Studying medicine lasts for 6 years and five and a half years for dentistry. The degrees consist of a biomedical preclinical phase (1st and 2nd years of study) and a clinical phase (from 3rd to 6th years of study). Psychology students complete a bachelor's degree (3 years) and a master's degree (2 years) in about 5 years. Graduates of the Master of Psychology degree receive a license to operate as a healthcare professional (psychologists). Students of psychology also have an opportunity to complete a master's degree that does not lead to the profession of a psychologist.

The Faculty of Medicine implements student-centered, mutually supportive learning methods, strives to reconcile theory and practice, and uses practical professional situations as the basis for teaching and learning. The preclinical phase is largely the same for both medical and dental students, and teaching is largely based on problem-based learning (Norman and Schmidt, 1992; Dolmans et al., 2005; Dolmans and Schmidt, 2006). About 200 medical and dental students are divided into groups of ten, with one or two dental students in each group. In addition to PBL tutorials, students have complementary lectures, assignments and laboratory work. When studying psychology, the first academic year includes basic and intermediate level material as well as material related to research methods in psychology. More detailed descriptions of the used teaching and learning methods are presented in **Table 1**. Because of the COVID-19 pandemic, teaching and learning in medicine, dentistry and psychology was to a great extent organized remotely in 2020 and 2021.

Participants

The study involved 374 first-year students of medicine, dentistry and psychology. The students completed the HowULearn questionnaire (Parpala and Lindblom-Ylänne, 2012) at the end of their first year of study. The questionnaire was sent to the students electronically using the Unihow system which is a digital reflection tool and feedback system. Students fill in the questionnaire as a part of their studying, and are provided with individual feedback for their learning and studying. The response rate was 91.4%. Students were asked for informed consent to use their responses for research purposes. Overall, 76.8% of the students agreed, and only their responses were used in this study. The study combined data from two academic years and cohorts: medical students in 2020 (n = 108) and in 2021 (n = 107), dental students in 2020 (n = 35) and in 2021 (n = 35), and psychology students in 2020 (n = 41) and in 2021 (n = 48).

TABLE 1 | Disciplines and used teaching and learning methods.

Discipline	Teaching and learning methods
Medicine	Problem-based learning, lectures, laboratory work, assignments
Dentistry	Problem-based learning, lectures, laboratory work, assignments
Psychology	Activating lectures, small group teaching, group work, case and observation tasks, experimental tasks and tasks which require application of knowledge

Materials

The HowULearn survey (Parpala and Lindblom-Ylänne, 2012) is used by the university we undertook this study at to provide feedback to students about their learning and to provide the university with information on how to support student learning (i.e., searching for evidence, relating ideas, understanding and systematic learning, workload, and paid employment while studying) throughout their studies until graduation. In this study, the HowULearn questionnaire (Parpala and Lindblom-Ylänne, 2012) was used to measure students' perceptions of generic skills and teaching-learning environment. The measures are presented in **Table 2**. The students were asked how they had developed generic skills, such as analyzing skills, problem-solving skills and collaboration and communication skills. The instrument included eight items. The items originated partly from a review of the literature and partly from the investigation of previous inventories [Course Experiences Questionnaire (CEQ); Wilson et al., 1997; Tuononen et al., 2019b]. The items used in the present study have not been used, and thus detailed analysis was conducted (see section "Analysis").

Students' perceptions of their teaching-learning environment were examined using fourteen items. The scale of the teaching-learning environment included four dimensions: perceived interest and relevance, alignment in teaching, peer support and constructive feedback. The scale originated from the Experiences of Teaching and Learning Environment Questionnaire (ETLQ; Entwistle et al., 2003). The instrument has been validated in other research and found to be robust across contexts (Parpala and Lindblom-Ylänne, 2012; Parpala et al., 2013; Karagiannopoulou and Milienos, 2018; Parra-González et al., 2021). The students responded to all items in this study on a 5-point Likert-type scale ranging from 1 = totally disagree to 5 = totally agree. Items and scales measuring generic skills and different elements of the teaching and learning environment are presented in more detail in **Supplementary Appendix A**.

Analysis

Firstly, Exploratory (EFA) and Confirmatory Factor Analysis (CFA) were performed to investigate the factorial structure of the measures of generic skills. EFA was conducted using SPSS (version 28.0) and CFA was conducted using MPlus (version 8.6; Muthén and Muthén, 1998–2012). EFA using principal axis factoring and promax rotation was conducted to explore the structure of the items measuring generic skills because the instrument was new and it had not been tested and validated in this context. Based on the factor analysis, a three-factor solution was the clearest. All loadings were above the desired 0.32 mark (Tabachnick and Fidell, 2014). Communalities varied from moderate to low and one item remained below the desired 0.40 (Costello and Osborne, 2005). This result was further supported by the results of the testing of a three-factor CFA model (CFI = 0.951, SRMR = 0.044, RMSEA = 0.070). The Comparative Fit Index (CFI), the Standardized Root Mean Square Residual (SRMR) and the Root Mean Square Error of Approximation (RMSEA) were used to assess the overall quality of the model. Cronbach's alphas were 0.77 for knowledge analyzing skills, 0.61 for problem-solving skills and 0.81 for collaboration and communication skills.

The four scales which measured perceptions of the teaching and learning environment were examined with the CFA. The Comparative Fit Index (CFI), the Standardized Root Mean Square Residual (SRMR) and the Root Mean Square Error of Approximation (RMSEA) were used to assess the overall quality of the model. The fit model with the scales was good (CFI = 0.961, SRMR = 0.055, RMSEA = 0.062). Cronbach's alphas were 0.79 for interest and relevance, 0.77 for alignment, 0.63 for peer support and 0.70 for constructive feedback.

A one-way ANOVA was used to analyze differences in generic skills and perceptions of the teaching and learning environment among students in medicine, dentistry and psychology. All the generic skills, except collaboration and communication skills and problem-solving skills were normally distributed. In order to ensure the reliability of the findings, we conducted both parametric and non-parametric tests. These tests yielded similar results. All dimensions of the teaching and learning environment were normally distributed. The relationship between generic skills and teaching and learning environment was explored using Pearson's correlations. A linear regression model (stepwise method) was used to analyze which teaching and learning environment factors had the strongest relationship to generic skills. Separate analyses were conducted for each generic skill, using generic skill as a dependent variable and factors of teaching and learning environment as independent factors. SPSS version 28 was used to conduct the analyses.

TABLE 2 | The scales, factors, number of items and an example item for each factor.

Scale	Factor	Number of items	Example item
Generic skills	Knowledge analyzing skills	3	I have learnt to see things from different points of view.
	Problem-solving skills	2	My studies have developed my problem-solving skills in practical situations.
	Collaboration and communication skills	3	My studies have developed my collaboration skills.
Teaching- learning environment	Interest and relevance	3	I can see the relevance of what we are taught.
	Alignment	4	It is clear to me what I am expected to learn in courses.
	Peer support	3	Talking with other students helps me to develop my understanding.
	Feedback	4	The feedback given helps me to improve my ways of learning and studying.

RESULTS

Perceptions of Generic Skills and the Teaching and Learning Environment Among Students in Medicine, Dentistry and Psychology at the End of the First Study Year

Regarding generic skills, the results showed that all the students, the students in medicine, dentistry and psychology, received the highest scores for skills related to analyzing knowledge. Medical and dental students also scored highly on collaboration and communication and problem-solving skills.

In terms of perceptions of the teaching and learning environment, the students in medicine, dentistry and psychology received the highest scores on peer support. Interest and relevance also received high scores, whereas alignment in teaching and feedback were scored the lowest by the students. **Table 3** presents the means and standard deviations of students' perceptions of learning generic skills and perceptions of the teaching and learning environment among students in medicine, dentistry and psychology.

Differences in Perceptions of Learning Generic Skills and the Teaching-Learning Environment Among Students in Medicine, Dentistry and Psychology

The results of ANOVA showed statistically significant differences in students' perceptions of generic skills among students in medicine, dentistry and psychology (**Table 3**). The results showed statistically significant differences in knowledge analyzing skills [$F(2, 371) = 1.143$, $p = 0.030$, $\eta 2 = 0.02$], problem-solving skills [$F(2, 371) = 7.541$, $p = 0.000$, $\eta 2 = 0.07$] and collaboration and communication skills [$F(2, 371) = 8.697$, $p = 0.000$, $\eta 2 = 0.04$]. The effect size in knowledge analyzing skills as well as collaboration and communication skills can be considered small and the effect size in problem-solving skills can be considered medium (Cohen, 1988). The Bonferroni *post hoc* test was used. The pairwise comparisons showed that students in psychology scored higher than students in dentistry on skills related to analyzing knowledge ($p < 0.05$). Both medical and dental students scored higher than psychology students on problem-solving skills ($p < 0.001$), and medical students scored higher than psychology students on collaboration and communication skills ($p < 0.001$). There were no statistically significant differences in students' perceptions of the teaching and learning environment.

The Relationship Between Perceptions of the Teaching and Learning Environment and Learning Generic Skills Among Students in Medicine, Dentistry and Psychology

The analysis of the data showed that students' perceptions of the teaching and learning environment and learning generic skills were mostly statistically significantly positively related to each other (**Table 4**). The relationship between interest and relevance and all generic skills as well as feedback and all generic skills was statistically significantly positively related among students in medicine, dentistry and psychology.

TABLE 3 | Means and standard deviations of generic skills and elements of the teaching and learning and differences among Finnish students in medicine, dentistry and psychology.

	1. Medicine (n = 215) M (SD)	2. Dentistry (n = 70) M (SD)	3. Psychology (n = 89) M (SD)
Perceptions of generic skills			
Knowledge analyzing skills	3.78 (0.54)	3.71 (0.63)	3.94 (0.59)
Problem-solving skills	3.48 (0.67)	3.48 (0.63)	3.01 (0.89)
Collaboration and communication skills	3.76 (0.66)	3.64 (0.82)	3.38 (0.77)
Perceptions of teaching and learning environment			
Interest and relevance	3.98 (0.64)	3.91 (0.66)	3.88 (0.77)
Alignment	3.63 (0.61)	3.62 (0.56)	3.59 (0.67)
Peer support	4.12 (0.64)	4.00 (0.78)	4.16 (0.67)
Feedback	3.27 (0.72)	3.23 (0.62)	3.12 (0.75)

Bonferroni's test. Knowledge analyzing skills 3 > 2 Problem-solving skills 1 > 3**, 2 > 3**. Collaboration and communication skills 1 > 3**. *p < 0.05. **p < 0.001.*

TABLE 4 | Correlations between perceived generic skills and perceptions of the teaching and learning environment among Finnish students in medicine (n = 215), dentistry (n = 70) and psychology (n = 89).

	Knowledge analyzing skills			Problem-solving skills			Collaboration and communication skills		
	Medicine	Dentistry	Psychology	Medicine	Dentistry	Psychology	Medicine	Dentistry	Psychology
Interest and relevance	0.30**	0.54**	0.54**	0.21**	0.52**	0.24*	0.27**	0.49**	0.30**
Alignment	0.29**	0.19	0.35**	0.22**	0.21	0.29**	0.11	0.30*	0.25*
Peer support	0.30**	0.50**	0.30**	0.21**	0.52**	−0.01	0.43**	0.42**	0.24*
Feedback	0.32**	0.43**	0.34**	0.36**	0.36**	0.32**	0.14*	0.31*	0.32**

***Correlation is significant at the 0.01 level (2-tailed). *Correlation is significant at the 0.05 level (2-tailed).*

TABLE 5 | Summary of the regression analysis.

	Knowledge analyzing skills			Problem-solving skills			Collaboration and communication skills		
	Medicine[a] β	Dentistry[b] β	Psychology[c] β	Medicine[d] β	Dentistry[e] β	Psychology[f] β	Medicine[g] β	Dentistry[h] β	Psychology[i] β
Interest and relevance		0.37**	0.32**		0.31**	0.39**		0.48**	
Alignment	0.14*								
Peer support	0.21**	0.18*		0.15*	0.26*		0.44**	0.20*	0.28*
Feedback	0.24**	0.18*		0.30**			0.11*		0.31**

*$p < 0.05$, **$p < 0.001$.
[a] $R = 0.462$, adjusted $R^2 = 0.206$ [$F(3, 330) = 29.78$, $p < 0.001$].
[b] $R = 0.609$, adjusted $R^2 = 0.353$ [$F(3, 107) = 21.00$, $p < 0.001$].
[c] $R = 0.431$, adjusted $R^2 = 0.179$ [$F(1, 127) = 28.99$, $p < 0.001$].
[d] $R = 0.367$, adjusted $R^2 = 0.129$ [$F(2, 331) = 25.76$, $p < 0.001$].
[e] $R = 0.551$, adjusted $R^2 = 0.291$ [$F(2, 108) = 23.53$, $p < 0.001$].
[f] $R = 0.337$, adjusted $R^2 = 0.107$ [$F(1, 127) = 16.28$, $p < 0.001$].
[g] $R = 0.452$, adjusted $R^2 = 0.200$ [$F(2, 331) = 42.59$, $p < 0.001$].
[h] $R = 0.531$, adjusted $R^2 = 0.269$ [$F(2, 108) = 21.22$, $p < 0.001$].
[i] $R = 0.413$, adjusted $R^2 = 0.158$ [$F(2, 126) = 12.97$, $p < 0.001$].

Peer support was significantly positively related to all generic skills among students in medicine, dentistry and psychology except to problem-solving skills among students in psychology. Alignment was significantly positively related to all generic skills among students in psychology but not among students in medicine and dentistry.

The linear regression analysis showed that students' perceptions of the teaching and learning environment were differently related to perceptions of generic skills depending on the skill among students in medicine, dentistry and psychology. Knowledge analyzing skills were positively related to alignment ($p < 0.05$), peer support ($p < 0.001$) and feedback ($p < 0.001$) in medicine, interest and relevance ($p < 0.001$), peer support ($p < 0.05$) and feedback ($p < 0.05$) in dentistry, and interest and relevance ($p < 0.001$) in psychology. Problem-solving skills were positively related to peer support ($p < 0.05$) and feedback ($p < 0.001$) in medicine, interest and relevance ($p < 0.001$) and peer support ($p < 0.05$) in dentistry, and interest and relevance ($p < 0.001$) in psychology. Collaboration and communication skills were positively related to peer support ($p < 0.001$) and feedback ($p < 0.05$) in medicine, interest and relevance ($p < 0.001$) and peer support ($p < 0.001$) in dentistry and peer support ($p < 0.05$) and feedback ($p < 0.001$) in psychology. In **Table 5**, the relationships between generic skills and perceptions of the various elements of the teaching and learning environments among students in medicine, dentistry and psychology are presented in more detail.

DISCUSSION

This study focused, firstly, on the perceptions that students in medicine, dentistry and psychology had about learning generic skills and their teaching-learning environment at the end of their first study year, and secondly, how their perceptions about generic skills and teaching and learning environment differed, and thirdly, how the perceptions of the teaching and learning environment were related to perceptions of learning generic skills. Our results showed that the students in medicine, dentistry and psychology scored the highest on knowledge analyzing skills, although other skills were also relatively highly evaluated. The results are in line with previous studies that have shown that university students feel that analyzing skills are well learned at the university (Murdoch-Eaton et al., 2016; Tuononen et al., 2019a).

Our research showed that for students in medicine, dentistry and psychology, the most important element perceived in the teaching and learning environment was peer support. A similar observation has been made in previous studies, which show that peer support was reported the highest of the various elements of the teaching and learning environment among university students (Asikainen et al., 2014; Herrmann et al., 2017). Among the respondents of this study, interest and relevance also received high scores. However, students rated alignment in teaching and feedback the least, meaning we obtained similar results as in previous studies (Asikainen et al., 2014; Herrmann et al., 2017).

There were differences in the perceptions of learning generic skills between the students in medicine, dentistry and psychology. Learning problem-solving, communication and collaboration skills were more emphasized among medical and dental students, while learning analyzing skills was more pronounced among psychology students. One explanation for the results may be that the teaching and learning environments in medicine and dentistry differ from the learning environment in psychology, and that different environments support the development of different generic skills. Studies in psychology include more lectures, i.e., traditional instruction during the first study year, whereas studies in medicine and dentistry include problem-based learning. It should be noted that the differences among the students were rather small although they were statistically significant.

There is evidence that problem-based learning promotes the development of generic skills better than traditional teaching among medical students (Joseph et al., 2016). More precisely, research has shown that problem-based learning fostered the learning of collaboration and problem-solving skills (Razzaq and Ahsin, 2011; Karantzas et al., 2013). It seems that the active learning methods, such as problem-based learning in health professions education, enable students to learn generic skills as they include a variety of active and collaborative learning activities. Such integration of generic skills into disciplinary courses has proven to be a better way to learn these skills than separate courses (Bath et al., 2004; Star and Hammer, 2008; Virtanen and Tynjälä, 2018).

In addition to problem-based learning (PBL), other active learning methods, such as team-based learning (TBL), case-based learning (CBL) and Flipped Classroom, have been developed and implemented in education for health professionals in recent decades (Parmelee et al., 2012; Prober and Khan, 2013; McLean, 2016). What all of these learning methods have in common is the use of authentic problems as stimuli for learning, collaboration and communication in small groups to solve the problems given. They provide a well-designed and student-centered approach to learning, take advantage of small-group work in large groups, and thus provide resource efficiency for teaching (Burgess et al., 2020). In addition to these learning methods, health professions education involves communication skills studies to enhance the future professionals' communication with patients (Berkhof et al., 2011). Surprisingly, there has been little research on the active learning methods except for problem-based learning and generic skills learning, so we cannot answer the question of whether team-based learning, case-based learning, flipped classroom, and communication skills studies promote students' generic skills learning, and if so, how. Therefore, we suggest further studies be undertaken on how different active learning methods foster generic skills learning.

A surprising finding in our study was that there were no statistically significant differences in students' perceptions of the different dimensions of the teaching and learning environment. In other words, despite the differences in their learning environments, they perceived different elements in the teaching and learning environment (interest and relevance of the programs, alignment in teaching, support from other students and feedback from teachers) in the same way. As in a previous study (Tynjälä et al., 2016), students in different courses perceived pedagogical elements similarly. In this study, this may be due to the fact that these elements could also be included in traditional lecture courses.

According to our study, the different dimensions of the teaching and learning environment were positively related to students' perceptions of learning generic skills. In medicine, the strongest predictors of generic skills were peer support and feedback, whereas in dentistry, the strongest predictors were peer support and interest and relevance, and in psychology, interest and relevance. These results indicated that interest and relevance, peer support, and feedback were related to perceptions of learning all generic skills, whereas alignment was only related to analyzing skills. Similarly in previous research, peer support was found to be related to generic skills (Myllykoski-Laine et al, 2022).

Limitations

There are some methodological limitations that should be considered when interpreting the results of the present study. First, the results of this study are based entirely on students' self-reports regarding their perceptions of learning generic skills and perceptions of their teaching and learning environment. Although self-reports have been widely used in assessing students' generic skills, through self-reports it is difficult to ascertain the students' actual performance in real-life environments (Braun and Mishra, 2016). The assessment of students' actual level of generic skills would require performance-based assessment (Zlatkin-Troitschanskaia et al., 2015; Hyytinen et al., 2021). However, self-reports enable students to recognize and evaluate their generic skills and thus develop their reflection skills (Kyndt et al., 2014). Second, it was not possible to draw conclusions about the impact of the specific learning method, i.e., problem-based learning or different elements of teaching and learning environment, on the learning of generic skills. This would require the use of different research design and data, for example, observational data. Third, the factor analysis showed acceptable fit for generic skills instruments, although the Cronbach's alpha for the problem-solving skills scale was relatively low. This could be because the scale included only two items. In addition, one factor measured both collaboration and communication skills, which could be problematic if students perceived learning these two skills differently (Tuononen, 2019). Therefore, a coherent and valid instrument for measuring generic skills should be developed. Fourth, the participants of this study represented students in health care professions. Therefore, the findings cannot be directly generalized to students' perceptions in other disciplines. Finally, the study was conducted at one university in one country.

Conclusion

To conclude, this study contributes to the research into generic skills as well as the interaction between the teaching and learning environment and generic skills learning by highlighting the

importance of different elements of the teaching and learning environment in the learning of generic skills. Learning problem-solving, as well as collaboration and communication skills were more emphasized among medical and dental students, whereas analyzing skills were more pronounced among psychology students. Students' perceptions of learning generic skills were related to perceptions of teaching and learning environment. These differences could be explained by the different teaching and learning environments.

Practical Implications and Future Research

It is important that students have an opportunity to study in active and collaborative learning environments and practice diverse skills and that they are provided with an opportunity for peer support. The study showed that students' perceptions of different elements of the teaching and learning environment, were related to perceptions of learning generic skills. Therefore, they should be considered in teaching and developing curricula. The teachers should pay greater attention to the relevance of the subject matter, as it supports the development of generic skills (Parpala et al., 2017). For example, highlighting the clinical relevance of collaboration skills, motivates students to learn these skills (Aarnio et al., 2010). For students, it is crucial to understand the relevance of generic skills to their studying and future working life so that they are motivated to learn them (Chan and Fong, 2018; Tuononen et al., 2019b). It is also essential that students are given feedback on their development during the courses.

The students, who participated in the study, were first-year students. It is important that students' generic skills are developed from the early phases of university studies. Therefore, it is essential to examine students' perceptions of their generic skills and the relationship between students' perceptions of the teaching and learning environment and generic skills already during the first study year. When interpreting the results it is important to take into account that the results could be different among higher grade students. In the future, it would be important to examine students' perceptions in different phases of their studies and to conduct longitudinal studies on how perceptions of learning generic skills change during the study program. In addition, the actual level of generic skills among students in medicine, dentistry and psychology could be examined in more detail through performance-based assessment, through simulations (Boursicot et al., 2021) and authentic patient encounters (Norcini et al., 2003). Furthermore, it would be intriguing to study the relationship among generic skills, teaching and learning environment and academic achievement.

Finally, it is important to consider what different degree programs could learn from each other about students' generic skills learning. In the design and implementation of curricula it could be considered whether generic skills could be learned through interdisciplinary courses that would enhance students' interprofessional communication and collaboration, skills that are essential to them in the working life (Reeves et al., 2015, 2016, 2017).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MR and TT contributed to conception, design of the study, organized the database, and performed the statistical analysis. MR, TT, and EP wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

ACKNOWLEDGMENTS

We would like to thank Saara Repo for participating in the data collection of the present study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2022.886052/full#supplementary-material

REFERENCES

Aarnio, M., Nieminen, J., Pyörälä, E., and Lindblom-Ylänne, S. (2010). Motivating medical students to learn teamwork skills. *Med. Teach.* 32, 199–204. doi: 10.3109/01421591003657469

Asikainen, H., Parpala, A., Lindblom-Ylänne, S., Vanthournout, G., and Coertjens, L. (2014). The development of approaches to learning and perceptions of the teaching-learning environment during Bachelor level studies and their relation to study success. *High. Educ. Stud.* 4, 24–36. doi: 10.5539/hes.v4n4p24

Barrie, S. C. (2006). Understanding what we mean by the generic attributes of graduates. *High. Educ.* 51, 215–241.

Batalden, P., Leach, D., Swing, S., Dreyfus, H., and Dreyfus, S. (2002). General competencies and accreditation in graduate medical education. *Health Affair* 21, 103–111.

Bath, D., Smith, C., Stein, S., and Swann, R. (2004). Beyond mapping and embedding graduate attributes: bringing together quality assurance and action learning to create a validated and living curriculum. *High. Educ. Res. Dev.* 23, 313–328. doi: 10.1080/0729436042000235427

Berkhof, M., Van Rijssen, H. J., Schellart, A. J. M., Anema, J. R., and Van der Beek, A. J. (2011). Effective training strategies for teaching communication skills to physicians: an overview of systematic reviews. *Patient Educ. Couns.* 84, 152–162. doi: 10.1016/j.pec.2010.06.010

Biggs, J. B. (2003). *Teaching for Quality Learning at University: What the Student Does*, 2nd Edn. Ballmoor: Society for Research into Higher Education.

Bokken, L., Linssen, T., Scherpbier, A., van der Vleuten, C., and Rethans, J. J. (2009). Feedback by simulated patients in undergraduate medical education: a systematic review of the literature. *Med. Educ.* 43, 202–210. doi: 10.1111/j.1365-2923.2008.03268.x

Boursicot, K., Kemp, S., Wilkinson, T., Findyartini, A., Canning, C., Cilliers, F., et al. (2021). Performance assessment: consensus statement and recommendations from the 2020 Ottawa Conference. *Med. Teach.* 43, 58–67. doi: 10.1080/0142159X.2020.1830052

Braun, E., and Mishra, S. (2016). "Methods for assessing competences in higher education: a comparative review," in *Theory and Method in Higher Education Research*, eds J. Huisman and M. Tight (Bingley: Emerald Group Publishing Limited), 47–68.

Breen, L., Pike, L. T., and Arco, L. (2003). From postgraduate student to professional: work-based learning in psychology. *Issues Educ. Res.* 13, 13–30.

Bridges, D., Davidson, R. A., Soule Odegard, P., Maki, I. V., and Tomkowiak, J. (2011). Interprofessional collaboration: three best practice models of interprofessional education. *Med. Educ. Online* 16:6035. doi: 10.3402/meo.v16i0.6035

Burch, V. C., Sikakana, C. N. T., Gunston, G. D., and Murdoch-Eaton, D. (2018). Self-reported generic learning skills proficiency: another measure of medical school preparedness. *Afr. J. Health Prof. Educ.* 10, 114–123. doi: 10.7196/AJHPE.2018.v10i2.971

Burgess, A., Bleasel, J., Hickson, J., Guler, C., Kalman, E., and Haq, I. (2020). Team-based learning replaces problem-based learning at a large medical school. *BMC Med. Educ.* 20:492. doi: 10.1186/s12909-020-02362-4

Chan, C. K. Y., and Fong, E. T. Y. (2018). Disciplinary differences and implications for the development of generic skills: a study of engineering and business students' perceptions of generic skills. *Eur. J. Eng. Educ.* 43, 927–949.

Chan, C. K. Y., Fong, E. T. Y., Luk, L. Y. Y., and Ho, R. (2017). A review of literature on challenges in the development and implementation of generic competencies in higher education curriculum. *Int. J. Educ. Dev.* 57, 1–10. doi: 10.1016/j.ijedudev.2017.08.010

Chesluk, B. J., Reddy, S., Hess, B., Bernabeo, E., Lynn, L., and Holmboe, E. (2015). Assessing interprofessional teamwork: pilot test of a new assessment module for practicing physicians. *J. Contin. Educ. Health Prof.* 35, 3–10. doi: 10.1002/chp.21267

Cleland, J. A., Abe, K., and Rethans, J. J. (2009). The use of simulated patients in medical education: AMEE Guide No 42. *Med. Teach.* 31, 477–486. doi: 10.1080/01421590903002821

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erbaum.

Cooper, N., Bartlett, M., Gay, S., Hammond, A., Lillicrap, M., Matthan, J., et al. (2021). UK Clinical Reasoning in Medical Education (CReME) consensus statement group. Consensus statement on the content of clinical reasoning curricula in undergraduate medical education. *Med. Teach.* 43, 152–159. doi: 10.1080/0142159X.2020.1842343

Costello, A. B., and Osborne, J. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Pract. Assess. Res. Eval.* 10:7. doi: 10.7275/jyj1-4868

Cuyvers, K., Donche, V., and Van den Bossche, P. (2015). Learning beyond graduation: exploring newly qualified specialists' entrance into daily practice from a learning perspective. *Adv. Health Sci. Educ.* 21, 439–453. doi: 10.1007/s10459-015-9640-y

D'Amour, D., Ferrada-Videla, M., Rodriguez, L., and Beaulieu, M. D. (2005). The conceptual basis for interprofessional collaboration: core concepts and theoretical frameworks. *J. Interprof. Care* 19, 116–131. doi: 10.1080/13561820500082529

Deveugele, M. (2015). Communication training: skills and beyond. *Patient Educ. Couns.* 98, 1287–1291. doi: 10.1016/j.pec.2015.08.011

Dolmans, D. H. J. M., De Grave, W., Wolfhagen, I. H. A. P., and Van der Vleuten, C. P. M. (2005). Problem-based learning: future challenges for educational practice and research. *Med. Educ.* 39, 732–741. doi: 10.1111/j.1365-2929.2005.02205.x

Dolmans, D. H., and Schmidt, H. G. (2006). What do we know about cognitive and motivational effects of small group tutorials in problem-based learning? *Adv. Health Sci. Educ. Theory Pract.* 11:321. doi: 10.1007/s10459-006-9012-8

El Tantawi, M. M., Abdelaziz, H., AbdelRaheem, A. S., and Mahrous, A. A. (2014). Using peer-assisted learning and role-playing to teach generic skills to dental students: the health care simulation model. *J. Dental Educ.* 78, 85–97.

Entwistle, N., McCune, V., and Hounsel, J. (2002). *Approaches to Studying and Perceptions of University Teaching-Learning Environments: Concepts, Measures and Preliminary Findings. Occasional report 1. ETL project*. Edinburgh: University of Edinburgh.

Entwistle, N., McCune, V., and Hounsell, J. (2003). "Investigating ways of enhancing university teaching-learning environments: measuring students' approaches to studying and perceptions of teaching," in *Unravelling Basic Components and Dimensions of Powerful Learning Environments*, eds E. De Corte, L. Verschaffel, N. Entwistle, and J. van Merrienboer (Oxford: Elsevier Science), 89–108.

Gade, S., and Chari, S. (2012). Case-based learning in endocrine physiology: an approach toward self-directed learning and the development of soft skills in medical students. *Adv. Physiol. Educ.* 37, 356–360. doi: 10.1152/advan.00076.2012

García-Aracil, A., and Van der Velden, R. (2008). Competencies for young European higher education graduates: labor market mismatches and their payoffs. *High. Educ.* 55, 219–239.

Girotto, M., de Andrés, A., and Arisó, A. (2021). Undergraduate business student's self-assessment of meta-competencies in the context of the final year projects. *Int. J. Res. Educ. Sci.* 7, 988–1005.

Golding, R. M., Breen, L. J., Krause, A. E., and Allen, P. J. (2019). The summer undergraduate research experience as a work-integrated learning opportunity and potential pathway to publication in psychology. *Front. Psychol.* 10:541. doi: 10.3389/fpsyg.2019.00541

Haddara, W., and Lingard, L. (2013). Are we all on the same page? A discourse analysis of interprofessional collaboration. *Acad. Med.* 88, 1509–1515. doi: 10.1097/ACM.0b013e3182a31893

Hamilton, K., Morrissey, S. A., Farrell, L. J., Ellu, M. C., O'Donovan, A., Weinbrecht, T., et al. (2018). Increasing psychological literacy and work readiness of Australian psychology undergraduates through a capstone and work-integrated learning experience: current issues and what needs to be done. *Austr. J. Psychol.* 53, 151–160. doi: 10.1111/ap.12309

Herrmann, K. J., Bager-Elsborg, A., and Parpala, A. (2017). Measuring perceptions of the learning environment and approaches to learning: validation of the learn questionnaire. *Scand. J. Educ. Res.* 61, 526–539. doi: 10.1080/00313831.2016.1172497

Hew, K. F., and Lo, C. K. (2018). Flipped classroom improves student learning in health professions education: a meta-analysis. *BMC Med. Educ.* 18:38. doi: 10.1186/s12909-018-1144-z

Hidi, S., and Renninger, K. A. (2006). The four-phase model of interest development. *Educ. Psychol.* 41, 111–127. doi: 10.1207/s15326985ep4102_4

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educ. Psychol. Rev.* 16, 235–266.

Hongbiao, Y., and Zheng, K. (2017). Students' course experience and engagement: an attempt to bridge two lines of research on the quality of undergraduate education. *Assess. Eval. High. Educ.* 42, 1145–1158. doi: 10.1080/02602938.2016.1235679

Hyytinen, H., Ursin, J., Silvennoinen, K., Kleemola, K., and Toom, A. (2021). The dynamic relationship between response processes and self-regulation in critical thinking assessments. *Stud. Educ. Eval.* 71:101090. doi: 10.1016/j.stueduc.2021.101090

Joseph, N., Rai, S., Madi, D., Bhat, K., Kotian, S. M., and Kantharaju, S. (2016). Problem-based learning as an effective learning tool in community medicine: initiative in a private medical college of a developing country. *Indian J. Commun. Med.* 41, 133–140. doi: 10.4103/0970-0218.177535

Karagiannopoulou, E., and Milienos, F. (2018). Experiences of the teaching-learning environment and approaches to learning: testing the structure of the "Experiences of Teaching and Learning" inventory in relation to earlier analyses. *Int. J. Teach. Learn. High. Educ.* 30, 506–521.

Karantzas, G. C., Avery, M. R., Macfarlane, S., Mussap, A., Tooley, G., Hazelwood, Z., et al. (2013). Enhancing critical analysis and problem-solving skills in undergraduate psychology: an evaluation of a collaborative learning and problem-based learning approach. *Aust. J. Psychol.* 65, 38–45. doi: 10.1111/ajpy.12009

Kember, D., and Leung, D. Y. (2005). The influence of the teaching and learning environment on the development of generic capabilities needed for a knowledge-based society. *Learn. Environ. Res.* 8, 245–266. doi: 10.1007/s10984-005-1566-5

Khan, I., and Fareed, A. (2001). Problem based learning variant: transition phase for a large institution. *J. Pak. Med. Assoc.* 51, 271–274.

Knipprath, H. (2017). "How higher education may contribute to the development of graduates' generic competences?," in *Higher Education Transitions - Theory and Research*, eds E. Kyndt, V. Donche, V. K. Trigwell, and S. Lindblom-Ylänne (Routledge: Taylor & Francis Group), 254–269.

Koponen, J., Pyörälä, E., and Isotalus, P. (2012). Comparing three experiential learning methods and their effect on medical students' attitudes to learning communication skills. *Med. Teach.* 34, e198–e207. doi: 10.3109/0142159X.2012.642828

Kridiotis, C. A., and Swart, S. (2017). A learning development module to support academically unsuccessful 1st-year medical students. *Afr. J. Health Prof. Educ.* 9, 62–66. doi: 10.7196/AJHPE.2017.v9i2.694

Krupat, E., Richards, J. B., Sullivan, A. M., Fleenor, T. J. Jr., and Schwartzstein, R. M. (2016). Assessing the effectiveness of case-based collaborative learning via randomized controlled trial. *Acad. Med.* 91, 723–729. doi: 10.1097/ACM.0000000000001004

Kyndt, E., Janssens, I., Coertjens, L., Gijbels, D., Donche, V., and Van Petegem, P. (2014). Vocational education students' generic working life competencies: developing self-assessment instrument. *Vocat. Learn.* 7, 365–392.

Lane, C., and Rollnick, S. (2007). The use of simulated patients and role-play in communication skills training: a review of the literature to August 2005. *Patient Educ. Couns.* 67, 13–20. doi: 10.1016/j.pec.2007.02.011

Liu, J., St John, K., and Courtier, A. (2017). Development and validation of an assessment instrument for course experience in a general education integrated science course. *J. Geosci. Educ.* 65, 435–454.

McGinness, A. K., Wamsley, M., and Rivera, J. (2019). Assessing interprofessional collaboration: pilot of an interprofessional feedback survey for first-year medical students. *J. Interprof. Educ. Pract.* 15, 131–137. doi: 10.1016/j.xjep.2019.03.012

McLaughlin, J. E., Roth, M. T., Glatt, D. M., Gharkholonarehe, N., Davidson, C. A., Griffin, L. M., et al. (2014). The flipped classroom: a course redesign to foster learning and engagement in a health professions school. *Acad. Med.* 89, 236–243. doi: 10.1097/ACM.0000000000000086

McLean, S. F. (2016). Case-based learning and its application in medical and health-care fields: a review of worldwide literature. *J. Med. Educ. Curric. Dev.* 27:3. doi: 10.4137/JMECD.S20377

Mikkonen, J., Heikkilä, A., Ruohoniemi, M., and Lindblom-Ylänne, S. (2009). I study because I'm interested: university students' explanations for their disciplinary choices. *Scand. J. Educ. Res.* 53, 229–244. doi: 10.1080/00313830902917261

Monteiro, S. M., and Norman, G. (2013). Diagnostic reasoning: where we've been, where we're going. *Teach. Learn. Med.* 25(Suppl. 1), 26–32. doi: 10.1080/10401334.2013.842911

Moura, D., Costa, M. J., Pereira, A. T., Macedo, A., and Figueiredo-Braga, M. (2021). Communication skills preparedness for practice: Is there a key ingredient in undergraduate curricula design? *Patient Educ. Couns.* 105, 756–761. doi: 10.1016/j.pec.2021.06.034

Mubuuke, A. G., Louw, A. J. N., and Van Schalkwyk, S. (2016). Utilizing students' experiences and opinions of feedback during problem-based learning tutorials to develop a facilitator feedback guide: an exploratory qualitative study. *BMC Med. Educ.* 16:6. doi: 10.1186/s12909-015-0507-y

Murdoch-Eaton, D., and Whittle, S. (2012). Generic skills in medical education: developing the tools for successful lifelong learning. *Med. Educ.* 46, 120–128. doi: 10.1111/j.1365-2923.2011.04065.x

Murdoch-Eaton, D., Louw, A. J. N., and Bezuidenhout, J. (2016). Effect of curriculum changes to enhance generic skills proficiency of 1st-year medical students. *Afr. J. Health Prof. Educ.* 8, 15–19. doi: 10.7196/AJHPE.2016.v8i1.414

Muthén, L. K., and Muthén, B. O. (eds). (1998–2012). *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.

Myllykoski-Laine, S., Lahdenperä, J., Nikander, L., and Postareff, L. (2022). Students' experiences of the development of generic competences in the Finnish higher education context – the role of the teaching-learning environment and approaches to learning. *Eur. J. High. Educ.* doi: 10.1080/21568235.2022.2058975

Norcini, J. J., Blank, L. L., Duffy, F. D., and Fortna, G. S. (2003). The mini-CEX: a method for assessing clinical skills. *Ann. Intern. Med.* 138, 476–481. doi: 10.7326/0003-4819-138-6-200303180-00012

Norman, G. R., and Schmidt, H. G. (1992). The psychological basis of problem-based learning: a review of the evidence. *Acad. Med.* 67, 557–565. doi: 10.1097/00001888-199209000-00002

Parmelee, D., and Michaelsen, L. K. (2010). Team-based learning: it's here and it works! *Acad. Med.* 85:1658. doi: 10.1097/ACM.0b013e3181f55a35

Parmelee, D., Michaelsen, L. K., Cook, S., and Hudes, P. D. (2012). Team-based learning: a practical guide: AMEE Guide No. 65. *Med. Teach.* 34, e275–e287. doi: 10.3109/0142159X.2012.651179

Parpala, A., and Lindblom-Ylänne, S. (2012). Using a research instrument for developing quality at the university. *Qual. High. Educ.* 18, 313–328.

Parpala, A., Asikainen, H., Ruohoniemi, M., and Lindblom-Ylänne, S. (2017). The relationship between the development of time and effort management and experiences of the teaching-learning environment in a university context. *Int. J. Learn. Change* 9, 170–184. doi: 10.1504/IJLC.2017.084594

Parpala, A., Lindblom-Ylänne, S., Komulainen, E., and Entwistle, N. (2013). Assessing students' experiences of teaching–learning environments and approaches to learning: validation of a questionnaire in different countries and varying contexts. *Learn. Environ. Res.* 16, 201–215. doi: 10.1007/s10984-013-9128-8

Parpala, A., Lindblom-Ylänne, S., Komulainen, E., Litmanen, T., and Hirsto, L. (2010). Students' approaches to learning and their experiences of the teaching-learning environment in different disciplines. *Br. J. Educ. Psychol.* 80, 269–282. doi: 10.1348/000709909X476946

Parra-González, M.-E., López-Belmonte, J., Segura-Robles, A., and Moreno-Guerrero, A.-J. (2021). Spanish adaptation and validation of the teaching and learning experiences questionnaire. *Int. J. Environ. Res. Public Health* 18:3518. doi: 10.3390/ijerph18073518

Piróg, D. (2016). The impact of degree programme educational capital on thetransition of graduates' labour market. *Stud. High. Educ.* 41, 95–109.

Prober, C. G., and Khan, S. (2013). Medical education reimagined. *Acad. Med.* 88, 1407–1410. doi: 10.1097/ACM.0b013e3182a368bd

Razzaq, Z., and Ahsin, S. (2011). PBL wrap up sessions: an approach to enhance generic skills in medical students. *J. Ayub. Med. Coll. Abbottabad* 23, 162–165.

Reeves, S., Boet, S., Zierler, B., and Kitto, S. (2015). Interprofessional education and practice Guide No. 3: evaluating interprofessional education. *J. Interprof. Care* 29, 305–312. doi: 10.3109/13561820.2014.1003637

Reeves, S., Fletcher, S., Barr, H., Birch, I., Boet, S., Davies, N., et al. (2016). A BEME systematic review of the effects of interprofessional education: BEME Guide No. 39. *Med. Teach.* 38, 656–668. doi: 10.3109/0142159X.2016.1173663

Reeves, S., Pelone, F., Harrison, R., Goldman, J., and Zwarenstein, M. (2017). Interprofessional collaboration to improve professional practice and healthcare outcomes. *Cochr. Database Syst. Rev.* 6:CD000072. doi: 10.1002/14651858.CD000072.pub3

Ruge, G., and McCormack, C. (2017). Building and construction students' skills development for employability – reframing assessment for learning in discipline-specific contexts. *Architec. Eng. Design Manage.* 13, 365–383. doi: 10.1080/17452007.2017.1328351

Schot, E., Tummers, L., and Noordegraaf, M. (2020). Working on working together. A systematic review on how healthcare professionals contribute to interprofessional collaboration. *J. Interprof. Care* 34, 332–342. doi: 10.1080/13561820.2019.1636007

Schwartz, R. W., Burgett, J. E., Blue, A. V., Donnelly, M. B., and Sloan, D. A. (1997). Problem-based learning and performance-based testing: effective alternatives for undergraduate surgical education and assessment of student performance. *Med. Teach.* 19, 19–23.

So, H. Y., Chen, P. P., Wong, G. K. C., and Chan, T. T. N. (2019). Simulation in medical education. *J. R. Coll. Phys. Edinburgh* 49, 52–57.

Star, C., and Hammer, S. (2008). Teaching generic skills: eroding the higher purpose of universities, or an opportunity for renewal? *Oxford Rev. Educ.* 34, 237–251. doi: 10.1080/03054980701672232

Tabachnick, B. G., and Fidell, L. S. (2014). *Using Multivariate Statistics*. Essex: Pearson Education Limited.

Trullàs, J. C., Blay, C., Sarri, E., and Pujol, R. (2022). Effectiveness of problem-based learning methodology in undergraduate medical education: a scoping review. *BMC Med. Educ.* 22:104. doi: 10.1186/s12909-022-03154-8

Tuononen, T. (2019). *Employability of University Graduates: The Role of Academic Competences, Learning and Work Experience in the Successful Transition from University to Working Life*. Helsinki: University of Helsinki.

Tuononen, T., Parpala, A., and Lindblom-Ylänne, S. (2019a). Graduates' evaluations of usefulness of university education, and early career success – A longitudinal study of the transition to working life. *Assess. Eval. High. Educ.* 44, 581–595. doi: 10.1080/02602938.2018.1524000

Tuononen, T., Parpala, A., and Lindblom-Ylänne, S. (2019b). Complex interrelations between academic competences and students' approaches to learning - Mixed-methods study. *J. Further High. Educ.* 44, 1080–1097. doi: 10.1080/0309877X.2019.1648776

Tynjälä, P., Virtanen, A., Klemola, U., Kostiainen, E., and Rasku-Puttonen, H. (2016). Developing social competence and other generic skills in teacher education: applying the model of integrative pedagogy. *Eur. J. Teach. Educ.* 39, 368–387.

Utriainen, J., Tynjälä, P., Kallio, E., and Marttunen, M. (2018). Validation of a modified version of the experiences of teaching and learning questionnaire. *Stud. Educ. Eval.* 56, 133–143. doi: 10.1016/j.stueduc.2017.12.007

Virtanen, A., and Tynjälä, P. (2018). Factors explaining the learning of generic skills: a study of university students' experiences. *Teach. High. Educ.* 24, 880–894. doi: 10.1080/13562517.2018.1515195

Wilson, K., Lizzio, A., and Ramsden, P. (1997). The development, validation andapplication of the course experience questionnaire. *Stud. High. Educ.* 22, 33–53.

Winston, K. A., Van der Vleuten, C. P. M., and Scherpbier, A. J. (2012). The role of the teacher in remediating at-risk medical students. *Med. Teach.* 34, 732–742. doi: 10.3109/0142159X.2012.689447

Young, M., Thomas, A., Gordon, D., Gruppen, L., Lubarsky, S., Rencic, J., et al. (2019). The terminology of clinical reasoning in health professions education: implications and considerations. *Med. Teach.* 41, 1277–1284. doi: 10.1080/0142159X.2019.1635686

Zlatkin-Troitschanskaia, O., Shavelson, R. J., and Kuhn, C. (2015). The international state of research on measurement of competency in higher education. *Stud. High. Educ.* 40, 393–411. doi: 10.1080/03075079.2015.1004241

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Education

ORIGINAL RESEARCH
published: 24 June 2022
doi: 10.3389/feduc.2022.884635

Exploring First Semester Changes in Domain-Specific Critical Thinking

Tine Nielsen[1,2], Inmaculada Martínez-García[3] and Enrique Alastor[3]*

[1] Department of Applied Research in Education and Social Sciences, UCL University College [UCL Erhvervsakademi og Professionshøjskole], Odense, Denmark, [2] Department of Psychology, University of Copenhagen, Copenhagen, Denmark, [3] Department of Didactics and School Organization, Faculty of Educational Sciences, University of Málaga, Málaga, Spain

Critical thinking is a common aim for higher education students, often described as general competencies to be acquired through entire programs as well as domain-specific skills to be acquired within subjects. The aim of the study was to investigate whether statistics-specific critical thinking changed from the start of the first semester to the start of the second semester of a two-semester statistics course, where the curriculum contains learning objectives and assessment criteria related to critical thinking. The brief version of the Critical Thinking scale (CTh) from the Motivated Strategies of Learning Questionnaire addresses the core aspects of critical thinking common to three different definitions of critical thinking. Students rate item statements in relation to their statistics course using a frequency scale: 1 = never, 2 = rarely, 3 = sometimes, 4 = often, and 5 = always. Participants were two consecutive year-cohorts of full-time Bachelor of Psychology students taking a two-semester long statistics course placed in the first two semesters. Data were collected in class with a paper-pencil survey 1 month into their first semester and again 1 month into the second. The study sample consisted of 336 students ($n_{cohort\ 1}$ = 166, $n_{cohort\ 2}$ = 170) at baseline, the follow-up was completed by 270 students with 165 students who could be matched to their baseline response. To investigate the measurement properties of the CTh scale, item analysis by the Rasch model was conducted on baseline data and subsequently on follow-up data. Change scores at the group level were calculated as the standardized effect size (ES) (i.e., the difference between baseline and follow-up scores relative to the standard deviation of the baseline scores). Data fitted Rasch models at baseline and follow-up. The targeting of the CTh scale to the student sample was excellent at both timepoints. Absolute individual changes on the CTh ranged from −5.3 to 5.1 points, thus showing large individual changes in critical thinking. The overall standardized effect was small and negative (−0.12), with some variation in student strata defined by, gender, age, perceived adequacy of math knowledge to learn statistics, and expectation to need statistics in future employment.

Keywords: critical thinking, domain-specific, changes, higher education, Rasch model, statistics in psychology

OPEN ACCESS

Edited by:
Edith Braun,
Justus-Liebig Universität, Germany

Reviewed by:
Kristina Walz,
University of Giessen, Germany
Kari Nissinen,
University of Jyväskylä, Finland

****Correspondence:***
Tine Nielsen
tini@ucl.dk

Specialty section:
This article was submitted to Higher Education, a section of the journal Frontiers in Education

Received: *26 February 2022*
Accepted: *23 May 2022*
Published: *24 June 2022*

Citation:
Nielsen T, Martínez-García I and Alastor E (2022) Exploring First Semester Changes in Domain-Specific Critical Thinking. Front. Educ. 7:884635. doi: 10.3389/feduc.2022.884635

INTRODUCTION

Critical thinking is a central concept in higher education, and as it has become relevant at both the individual and societal level it will not only improve students' academic success but also the quality of education (Ren et al., 2020). The scientific literature has investigated the responsibility of educational institutions in training students in competencies that enables them to be future citizens ready to be an acting part of the society by making them critical thinkers (Kuhn, 1999; Paul and Elder, 2005). Thus, there has been a growing interest in the incorporation of critical thinking in the education curricula making critical thinking one of the main aims (Lau, 2015; McGuirk, 2021). With regard to the outcome of higher education, critical thinking is predominantly construed as generic, as it is described in terms of the competencies, students are expected to possess at the completion of a degree program (see **Supplementary Appendix 2** for the competency description for a degree program in this study). However, in terms of incorporating critical thinking into higher education programs, this appears rarely to be in the form of independent courses teaching critical thinking. More often critical thinking seems to be implemented through teaching methods and specifically designed activities within subject courses thus construing critical thinking as domain-specific, or simply by using the term critical thinking in the curriculum description without clear definitions, program- or course-determined approaches to teaching toward critical thinking (c.f. **Supplementary Appendix 2** for the current study). These two levels of implementing critical thinking in higher education tie to the discussion of critical thinking as generic/general or domain- and subject-specific.

There are different ways of understanding critical thinking that involve different implications for practice, so there is no consensus on a single definition (Moseley et al., 2005). Commonly in the literature, there is a distinction between thinking or cognitive skills and dispositional aspects of critical thinking, but as two sides of critical thinking and not separate positions. Two prevalent authors in the field, whose definitions or instruments many draw on in their research, are Facione and Halpern. Facione (1990) conducted a large Delphi study to narrow down the components of critical thinking, and the panel reached a consensus conceptualization of critical thinking as consisting of two dimensions: cognitive skills and affective dispositions. He further defines critical thinking as "*the purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based*" (Facione, 1990, p. 3). Halpern (2014) understand critical thinking as "*the deliberate use of skills and strategies that increase the probability of a desirable outcome*" (p. 450) and that critical thinking is involved in "*solving problems, formulating inferences, calculating likelihoods, and making decisions*" (p. 8), and thus also refer to both skills and dispositions. Facione and Halpern also make the distinction of critical thinking skills being assessed as the abilities to demonstrate critical skills in tasks or assignments, while critical thinking dispositions are assessed by self-report instruments. However, in the empirical studies in the field, there is no consensus on this. Thus, studies using self-report instruments have claimed to assess critical thinking skills (e.g., Ricketts and Rudd, 2005), studies employing critical thinking dispositions self-report instrument has claimed to assess critical thinking skills with this (e.g., Kanbay et al., 2017), and lastly, studies claiming to assess abilities are doing this through to some degree subjective teacher evaluations using short rubrics[1] (e.g., Ralston and Bays, 2015). See the following sections for more details on these studies. At a general level, there appears to be a conceptual shift toward using the term skills and then differentiating between assessed and self-report. Thus, in the remainder of this article, we simply use the term critical thinking skills, while recognizing that we use a self-report instrument to assess this, thus assessing students' perceptions of their critical thinking skills.

Critical Thinking as Generic/General Skills or Domain-Specific Skills

One particularly pertinent discussion in the field is whether critical thinking skills are generic/general skills or whether they are domain-/subject-specific (Tiruneh et al., 2017).

The view of critical thinking as a generic set of skills applicable across domains is based on the common features of critical thinking tasks across a wide variety of domains (e.g., Halpern, 1998; Kuhn, 1999). While Halpern (1998) is a proponent of critical thinking as a set of generic skills, her "Four-Part Model for Enhancing Critical Thinking" to teach critical thinking acknowledges that critical thinking takes place within a knowledge domain and should be taught within this domain. However, this does not mean that Halpern considers critical thinking as domain-specific, but rather that the domain is the learning context for skills, which can be applied more universally across domains after being mastered. The view of critical thinking as domain-specific emphasizes that different domains have different criteria relating to critical thinking and thus the skills required inevitably vary across domains (e.g., McPeck, 1992; Moore, 2011). The issue is more likely not an either/or issue, but an issue of both in combination, as content and critical thinking tasks and skills might differ across domains as they are invariably linked to the domain-knowledge, but there are also commonalities across domains, due to the cognitive processes involved in critical thinking (e.g., Bailin et al., 1999). As such, critical thinking may be regarded as a set of domain-specific skills of which some also belong to the set of generic or general critical thinking skills. Whether there is in fact a transfer effect from the domain-specific learning of critical thinking skills to other domains or adding on to generic critical thinking skills, as suggested in some of the literature, is another pertinent issue in the critical thinking field. However, this is not a central topic in the current study, as we are concerned with domain-specific

[1] Subjective in the sense that the rubrics have so few categories and the descriptions of categories are so general that even an identical scorings can ensure that the behavior or products rated by the teachers as indicating critical thinking is the same (c.f. Ralston and Bays, 2015).

critical thinking skills and their development in the first part of university studies.

Critical Thinking Skills in First-Year University Students and Their Development

First-year university students are particularly interesting when it comes to studying critical thinking skills and how they develop, as many higher education teachers and researchers concur that "first-year students often enter higher education without the ability to use higher-order thinking skills to master their studies" (De Jager, 2012, p. 1374). Much of the research into critical thinking skills of first-year university students and the development of critical thinking skills during university, has been focused on the development of teaching models and methods to enhance critical thinking, assessing their effects, and comparing how different teaching methods affect the critical thinking of the students. One example is Saenab et al. (2021) who developed the ReCODE model (Reading, Connecting, Observing, Discussing, and Evaluating) to improve first-year Biology students' acquisition of critical thinking. The outcome was positive with regard to enhancing students' critical thinking over the course of 3 months, however, it was only used on 38 students. Thomas (2011) developed the "Embedding generic skills in a business curriculum"-program consisting of activities and assessment resources for university teachers to develop critical thinking skills with their first-year students, and emphasize that these skills should be developed in the first year. The suggestions were not tested. On a similar note, Hammer and Green (2011) redesigned a written assessment in the form of a case-based business report for first-year management students in order to facilitate better development of critical thinking as this was part of the requirements for passing. The authors used the percentage of passing students to evaluate the success of the redesign – this went from 78.8 to 84% – but details of the teachers' assessments were not provided, and thus how critical thinking was assessed was not divulged beyond its being a teacher assessment. Ralston and Bays (2015), on the other hand, found that Engineering students' (n = 182) critical thinking increased during the course of their undergraduate studies, which had purposely been designed to incorporate assignments focused on critical thinking. A four-point, holistic critical thinking rubric was designed for the purpose of the study to evaluate domain-specific critical thinking. As a final example, Tiruneh et al. (2017) compared both domain-specific and general critical thinking skills for first-year students in an introductory Physics course (n = 143), using the Halpern Critical Thinking Assessment (HCTA; Halpern, 2015); a standardized scenario-based instrument with 25 everyday scenarios assessing general critical thinking skills by means of computerized scoring in combination with trained grader scoring. The study compared three different instructional designs and found that students in what they termed immersion and infusion designs (intervention) outperformed students in the control design significantly with regard to domain-specific critical thinking as well as course achievement. However, neither of the intervention designs fostered the acquisition of general critical thinking skills.

It is evident that there is an abundance of studies on various methods to enhance students' critical thinking skills in the first year and over the course of university studies. However, critical thinking in first-year students has also been investigated with regard to its "natural" development over time (i.e., no particular design implemented to enhance critical thinking) and how critical thinking is related to other psychological and educational constructs, e.g., emotional intelligence (Kaya et al., 2017; Sahanowas and Halder, 2020) and perceived academic control (Stupnisky et al., 2008). Sahanowas and Halder (2020) used the University of Florida - Engagement, Cognitive Maturity and Innovativeness assessment (UF-EMI, Ricketts and Rudd, 2005), which is a self-report instrument measuring generic critical thinking, in a cross-sectional study with the first-year students in various disciplines (n = 500) found that emotional intelligence was positively related to critical thinking. Kaya et al. (2017) in their study of Nursing students find that they possess a low level of critical thinking at the start of the first academic year, and while critical thinking was positively associated with emotional intelligence at the start, neither developed over the course of the year. Kaya et al. (2017) made use of a Turkish translation of the California Critical Thinking Disposition Scale (Facione et al., 1998), which is a self-report measure of generic critical thinking. Stupnisky et al. (2008) conducted a longitudinal study with Psychology students (n = 1,196) with the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al., 1991), which contains a domain-specific self-report critical thinking scale, and found a reciprocal relationship between critical thinking and perceived academic control, so that students perceived academic control 1 month into the first year predicted critical thinking 6 months later, while critical thinking 1 month into the first year also predicted perceived academic control 6 months later. Another example, of a study on the "natural" development of critical thinking over time, is Kanbay et al. (2017) who assessed critical thinking in Nursing students (n = 46), with the (California Critical Thinking Disposition Scale, see above) at the start of the first year and at the end of the second, third and fourth years of study. Their results revealed a medium level of critical thinking at the beginning and no improvement in critical thinking across the four-year period of time, not statistically and not at the absolute level. In a qualitative study, Özelçi and Çalışkan, 2019, interviewed 11 teacher candidates two times about their critical thinking. The results showed no change in self-perception in critical thinking from the first to the fourth year of study.

Development of Students' Statistics-Related Critical Thinking

Turning to the domain-specific concept of statistics-related critical thinking, several studies have been conducted. Bensley et al. (2010) studied the acquisition of critical thinking skills in instructional different groups of students enrolled in a research methods course by the psychological Critical Thinking Test

(Bensley and Baxter, 2006), which is a domain-specific multiple-choice test. More specifically they compared the acquisition of critical thinking skills for analyses of psychological arguments students who had critical thinking skills infused directly into their course with students where this was not the case. The infusing of critical thinking skills consisted of using a methodologically oriented textbook as well as a critical thinking textbook, as well as examples and practice of critical thinking through exercises and corrective feedback. The non-infusing courses used another textbook that embedded statics instruction within a research design and methodology discussion. The study found that the group of students who had received instruction aimed explicitly at critical thinking showed significantly greater gains in argument analysis skills than the students who had received no explicit critical thinking instruction. Contrary to this, Goode et al. (2018) compared how critical thinking was expressed in early and late writing assignments using specific critical thinking learning objectives recommended by the American Psychological Association (i.e., effective use of critical thinking, use of reasoning in argument, and problem-solving effectiveness) for psychology students assigned at random to a face-to-face and a blended learning versions of a statistics and research design course. Goode et al. (2018) developed a domain-specific scoring rubric with three areas being scored from 'does not meet expectations' to 'far exceeds expectations' for the teachers' assessment of critical thinking. The difference between the two instructional designs was simply that the blended learning version of the course was taught as a 50/50 flipped hybrid of the face-to-face course. Thus, in the blended learning hybrid, students attended face-to-face classes once a week rather than two, and for the second weekly class, they viewed online lectures and worked with other materials outside of the class setting. There was no significant difference in the development of critical thinking between students in the face-to-face and students in the blended learning design. However, an instructor effect was found, showing that student assigned to classes by two instructors increased their critical thinking significantly more than students assigned to two other instructors, and for one instructor both randomly assigned groups of students had a decline in critical thinking during the course. Setambah et al. (2019) evaluated how the critical thinking skills of teacher preparation students in their second semester developed in a basic statistics course employing Adventure-Based Learning (ABL) compared to a control group not receiving ABL. They found that after 10 weeks there was no significant difference, while there was weak evidence for a difference favoring the experimental groups after a further 8 weeks. Lastly, Cheng et al. (2018) showed how the critical thinking of undergraduate students taking introductory statistics classes within various degree programs increased across semester-long courses incorporating assignments, in-class discussion, and Socratic dialog. Cheng et al. (2018) designed a domain-specific rubric with four dimensions related to critical thinking to be scored by domain-specialists and as well as a student self-report survey to assess students' perceptions of improvement in critical thinking.

With regard to the domain-specific statistics-related critical thinking, there appears to be a lack of studies on the "natural" development over time, i.e., without implementation of any specific teaching methodology. As Tiruneh et al. (2017) suggest that "*meaningful instruction in every subject domain inherently comprises the development of CT skills, and therefore, proficiency in CT skills can be achieved as students construct knowledge of a subject-matter domain without any explicit emphasis on the teaching of general CT skills during instruction*" (p. 1067), such studies might contribute to the knowledge of the "natural" development of statistics-related critical thinking.

The Current Study

Drawing on the previous research, the present study intends to contribute to the field by studying specifically the development of statistics-related critical thinking in first-year psychology undergraduate students in a Danish university, where no particular emphasis on teaching critical thinking skills is reflected in the curriculum, but rather implicit references are given and critical thinking is mentioned in the assessment criteria (c.f. Moore, 2011). The primary aim of the study is to investigate whether statistics-related critical thinking changes from the start of the first semester to the start of the second semester of a two-semester-long statistics course, where the curriculum contains learning objectives implying critical thinking and assessment criteria explicitly requiring critical thinking.

At the overall level, we expected all students to have an increase in critical thinking, based on the general goal and performance orientation of these students[2] in combination with the implicit mention in the learning objectives and particularly the explicit mention of critical thinking in the assessment criteria for the first semester of the course. However, such an overall change might mask differentiated subgroup changes, and subgroup changes in opposite directions might also result in no change at the overall level. With regard to subgroups, we expected that the overall change in critical thinking would differ for subgroups of students dependent on their baseline perception of the adequacy of their own mathematical knowledge for learning statistics as well as their expectation to need statistics in their future employment, as would the students' baseline level of critical thinking. Specifically, we expected:

Students who perceived their mathematical knowledge to be inadequate for learning statistics were less inclined toward critical thinking at baseline compared to students who perceived they had an adequate level of mathematical knowledge, due to their lack of insight into the field. We had no set expectation with regard to the direction of difference in the change in critical thinking dependent on the perception of the adequacy of mathematical knowledge, as this could go both ways. For some students, a perceived lack in the prerequisite knowledge required would be a motivating factor making them engage more and thus possibly increase more in critical thinking compared to the other student group. But the opposite is also likely for some students, i.e., perceiving a lack of prerequisite knowledge might be further dis-engaging leading to a smaller increase in critical thinking

[2]In Denmark, psychology is one of the top-10 most difficult higher education degree program to be admitted to, as there are a limited number of vacancies to compete for, and thus it requires almost perfect grades to be admitted.

or even a decrease for this subgroup. In addition, students perceiving adequacy in prerequisite knowledge could be expected to engage more due to their insight and thus increase more in critical thinking than students perceiving their pre-requisite knowledge as inadequate.

Students who did not believe they would need statistics in their future employment were less inclined toward critical thinking at baseline and compared to students believing they would be needing statistics, as they would not be as likely to engage in the cognitively demanding critical because it would be perceived as unnecessary. We would not expect that students' beliefs about their future employment to change much over the cause of the first semester of study, and thus we expected that the largest increase in critical thinking would be seen for the students believing to need statistics in the future, as they would engage more in the subject.

The secondary aim is to investigate further the psychometric properties of the brief version of the Motivated Strategies of Learning Questionnaire critical thinking scale (MSLQ; Pintrich et al., 1991) resulting from a recent validation study, which critically considered the content and construct validity of this scale as well as its cross-cultural validity (Nielsen et al., 2021). As this brief critical thinking scale (CTh) was shown to fit the Rasch model both with a Danish and a Spanish sample of psychology students and have reliability for the Danish sample at the level of those obtained with the original scale, we found the CTh scale to be a good candidate for the current study.

MATERIALS AND METHODS

Instrument

The Critical Thinking scale (CTh) employed in the present study is a brief version of the critical thinking scale from the Motivated Strategies of Learning Questionnaire (MSLQ; Pintrich et al., 1991) resulting from a recent validation study, which critically considered the content and construct validity of this scale as well as its cross-cultural validity (Nielsen et al., 2021). The MSLQ is a multi-scale questionnaire intended to measure aspects of students' motivational orientation and learning strategies in high school and higher education (Pintrich et al., 1991). One of the scales included in the MSLQ is a five-item course-specific critical thinking scale with a seven-point response scale anchored for meaning only at the extremes. Of all the short scales in the MSLQ, the critical thinking scale was originally reported as having one of the highest reliabilities (Cronbach's alpha 0.8) with the development sample of 380 Midwestern college students (Pintrich et al., 1991). More recently, Holland et al. (2018) in their meta-analysis found the reliability of the critical thinking scale to be similar across 344 samples (N = 27,619) stemming from 32 countries and 14 languages (mean Cronbach's alpha 0.78).

In their study of the cross-cultural validity of the critical thinking scale from the MSLQ, Nielsen et al. (2021) analyzed thoroughly the content validity of the scale and found that only three items actually measured critical thinking. Content validity was considered both with a theoretically based approach, i.e., analysis of the item content in relation to three different and prevalent definitions of critical thinking (Facione, 1990; Pintrich et al., 1991; Halpern, 2003), and a statistical and psychometric approach, i.e., analysis of local independence and dimensionality (Kreiner and Christensen, 2004). Both approaches reached the conclusion that two items (the same) should be eliminated in order to improve content validity by eliminating construct contamination.

In addition to eliminating two items, Nielsen et al. (2021) also employed an adapted five-point response scale with meaning anchors for all categories with the brief CTh scale in order to pre-assign the meaning that respondents should infer from the categories and thus prevent a random assignment of meaning to a row of numbers, which would affect the validity in interpretation and reliability (Krosnick and Fabrigar, 1997; Maitland, 2009; Menold et al., 2014). This approach was further supported empirically in previous validity studies of other scales from the MSLQ, e.g., Nielsen (2018) with the motivation scales; Nielsen (2020), Nielsen et al. (2017, 2022) with the self-efficacy scale, where a similar adaption of the response scale had no noteworthy effect on the reliability of the scales compared to the original version.

The three-item CTh scale with the adapted response scale (see below) resulting from the study by Nielsen et al. (2021) had reliability at the level of the original five-item scale with seven response categories for a Danish sample of psychology students (0.82), while slightly lower for a Spanish sample of psychology students (0.73).

The items of the brief CTh scale employed in the present study address the purposeful and inquiring aspect of CTh common to three different definitions of critical thinking (Facione, 1990; Pintrich et al., 1991; Halpern, 2003): how often the student questions things and decide about them (item: I often find myself questioning things I hear or read in this statistics course to decide if I find them convincing); how the student decides about a theory, interpretation or conclusion (item: when a theory, interpretation or conclusion is presented in the statistics course or in the readings, I try to decide if there is good supporting evidence); how the student looks for alternatives (item: whenever I read or hear an assertion or conclusion in this statistics course, I think about possible alternatives) (see also **Supplementary Table A1** in **Supplementary Appendix 1**). Thus, the CTh scale does not cover all aspects of critical thinking, but it covers the core aspects, and more importantly, it is not "contaminated" by items not measuring critical thinking (Nielsen et al., 2021). As with the MSLQ, students rate how they feel that the item statements in the brief CTh scale describe them in relation to a specified course (in this case statistics) in terms of frequency of the thinking described in the items: 1 = never, 2 = rarely, 3 = sometimes, 4 = often, and 5 = always. The Danish item texts can be seen in **Supplementary Appendix 1** with the English equivalents (**Supplementary Table A1**). In this article, CTh items are referenced with their original order from the MSLQ to facilitate comparison to other studies with item-level data.

At baseline, students also provided information on gender and age, whether students perceived their mathematical knowledge to be adequate for learning statistics, and whether they believed they would need statistics in their future employment.

Participants and Data Collection

Participants were two consecutive year-cohorts of first-semester students enrolled in a full-time Bachelor of Psychology program in a major Danish university. The students were all taking a two-semester-long statistics course placed in the first two semesters of the bachelor's program. The course consists of weekly lectures and weakly exercise classes. The learning objectives for the first semester of the course contain implicit references to critical thinking (see **Supplementary Appendix 2**). The course has a separate exam in each of the two semesters, and the first-semester exam is an on-campus written exam assessed as pass/fail using a set of specified criteria. As part of these criteria are both implicit and explicit references to critical thinking (see **Supplementary Appendix 2**).

The students completed the CTh scale as part of a larger survey 1 month into their first semester of the course and again 1 month into their second semester of the statistics course. Data were collected in class with a paper-pencil survey. The data collections were arranged with the responsible lecturer before the start of the course. Students were informed ahead of the lecture that the data collection would take place and that it was voluntary to complete the survey. At the point of the data collection, students were informed of the purpose of the overall study, that participation was voluntary, that their data would be treated according to the prevailing data protection regulations, and that they could ask to have their data deleted up to a specified point in time where they would be anonymized. In addition, students were provided with a written information sheet providing the same information as well as contact information for the responsible researcher.

The study sample consisted of 336 students at baseline ($n_{cohort\ 1}$ = 166, $n_{cohort\ 2}$ = 170), while the follow-up was completed by 270 students with 165 students who could be matched to their baseline response. The matching rate was determined by circumstances related to student enrollment (drop-out and new enrollment), the matching design (asking students for their student ID in handwriting if they wanted to participate again), and chance (students present in the lecture where data were collected). Thus, as various factors contributed to the missingness of data at follow-up, it could not with any certainty be determined whether data were missing at random or not, though the number of contributing factors makes it more likely that they were missing at random. Likewise, the missingness could not be considered in terms of selection bias, due to the external contributing factors. The mean age of the students at baseline was 22.7 years (SD 4.99) and 81% of the 336 students in the baseline sample identified as female, which is a close match to the official gender distribution of the student admitted to the two particular year-cohorts was 81.3% female students (Ministry of Higher Education and Science, 2021). The gender distribution did not change at follow-up, i.e., 82% of the 165 students in the follow-up sample identified as female.

Statistical Analyses

First, we conducted item analysis using the Rasch measurement model (RM; Rasch, 1960) to establish the psychometric properties of the CTh scale both at baseline and at follow-up. The Rasch model was chosen, as Nielsen et al. (2021) have shown the CTh scale to fit the Rasch model in both a Danish and a Spanish sample. Second, we assessed the changes in CTh scores from the start of the first to the start of the second semester as standardized effect sizes.

Item Analyses by Rasch Models

To investigate the measurement properties of the CTh scale (the secondary issue of the study), item analysis by the Rasch model was conducted first on the baseline sample and subsequently in the follow-up sample to confirm the results. The RM provides optimal measurement properties of scales fitting it (Kreiner, 2007, 2013). These properties include:

1. *Unidimensionality* – the scale measures a single latent construct (Critical Thinking).
2. *Local independence of items* (no LD) – responses to a CTh item depends only on the level of Critical Thinking and not on responses to any of the other items on the scale.
3. *Optimal reliability*, as items are locally independent.
4. *Absence of differential item functioning* (no DIF) – responses to a CTh item depends only on the level of critical thinking and not on persons' membership of subgroups such as gender, age, etc.
5. *Homogeneity* – the rank order of the item parameters/item difficulty is the same for all persons.
6. *Score sufficiency* – the sum score is a sufficient statistic for the person's parameter estimates of Critical Thinking.

Homogeneity and sufficiency are properties only provided by the Rasch model, not any other IRT model. The property of sufficiency is particularly desirable when using the summed raw score of a scale, as it is the usual case with the CTh scale. However, fit to the Rasch model facilitates the use of the person parameter estimates resulting from the measurement model (sometimes termed Rasch-scores), and thus either these or the raw scores can be used in subsequent analysis, as preferred by the individual researcher for their specific purpose.

The overall tests of global homogeneity by comparison of item parameters in low and high scoring groups and overall tests of invariance were conducted as overall tests of fit using Andersen (1973) conditional likelihood ratio test (CLR). The fit of individual items to the Rasch model was tested by comparing the observed item-rest-score correlations with the expected item-rest-score correlations under the RM (Kreiner, 2011). Local independence of items and the assumption of no DIF were tested using Kelderman (1984) conditional likelihood ratio test. DIF was tested in relation to five background variables year cohort (1, 2), gender (female and male), median-split age groups (21 years and younger, 22 years and older), as well as baseline perception of the adequacy of mathematical knowledge to learn statistics (not adequate, adequate), and baseline expectancy to need statistics in future employment (yes, maybe, and no).

Reliability was calculated as Cronbach's alpha (Cronbach, 1951). Targeting (whether items provide information in the area of the scale where the sample population is located) was assessed graphically by item maps as well as numerically by two target indices (Kreiner and Christensen, 2013): the test information target index (the mean test information divided by the maximum test information for theta, and the root mean squared error (RMSE) target index (the minimum standard error of measurement divided by the mean standard error of measurement for theta). Both indices should preferably have a value close to one, as this would indicate the degree to which maximum information and minimum measurement error were obtained, respectively. The target of the observed score and the standard error of measurement (SEM) was also calculated. Items maps are plots of the distribution of the item threshold locations against weighted maximum likelihood estimations of the person parameter locations as well as the person parameters for the population (assuming a normal distribution) and the information function.

Critical values were adjusted for false discovery rate (FDR) arising from conducting multiple statistical tests (i.e., controlling type I errors), whenever appropriate (Benjamini and Hochberg, 1995). As recommended by Cox et al. (1977), we distinguished between weak ($p < 0.05$), moderate ($p < 0.01$), and strong ($p < 0.001$) evidence against the model, rather than applying a deterministic 5% critical limit for p-Values.

Analysis of Differences at Baseline and Analysis of Change

To investigate the primary issue of the study, namely changes in critical thinking, the person parameter estimates resulting from the Rasch models, which have equal distance between any two values, were rescaled to the score range of the instrument and used for baseline differences and in the analysis of change. Differences in mean scores for subgroups of students at baseline were tested within the framework of multiple analyses of variance framework to be able to include grouping variables with more than two categories and test for interaction effects. The change was tested using a paired samples t-test approach and change scores at the group level were calculated as the standardized effect size (ES) (i.e., the difference between baseline and follow-up scores relative to the standard deviation of the baseline scores) (Lakens, 2013; Beauchamp et al., 2015). Subgroups of students were defined by our primary independent variables of interest, i.e., perception of the adequacy of their own mathematical knowledge for learning statistics as well as the students' expectations to need statistics in their future employment. As secondary subgroupings, we included gender and age groups, in order to show whether there were any effects of these on baseline levels of critical thinking or on changes that might be imposed on the primary issues.

Software

The item analyses by Rasch models were conducted using DIGRAM (Kreiner, 2003; Kreiner and Nielsen, 2013), while R was used to produce the item maps. Analyses of variance and t-tests were conducted using SPSS. Effect sizes were calculated using Excel.

RESULTS

Psychometric Properties at Baseline and Follow-Up

Results of the item analyses (the secondary research issue) showed that the baseline data fitted the Rasch model, and this was also the case with the follow-up data. Thus, there was no evidence against global homogeneity or invariance (**Table 1**), nor was there any evidence against the fit of the individual items to the Rasch model (**Table 2**). In addition, we found no evidence against local independence of items (**Supplementary Table A2** in **Supplementary Appendix 1**) and no evidence of differential item functioning relative to year cohort, students' baseline perception of the adequacy of mathematical knowledge to learn statistics, students' baseline expectancy to need statistics in future employment, gender, or age (**Supplementary Table A3** in **Supplementary Appendix 1**). Information on Item thresholds, locations, difficulties, targets, and information is also provided in **Supplementary Appendix 1** (**Supplementary Table A4**).

The targeting of the CTh scale to the student sample was excellent at both baseline and follow-up; slightly better at follow-up with a target information index of 86% at follow-up versus 83% at baseline (**Supplementary Table A5** in **Supplementary Appendix 1**). The level of information is highest where most students are located on the CTh scale at both time points (**Supplementary Figure A1** in **Supplementary Appendix 1**). The reliability of the CTh scale was satisfactory for the purpose of statistical analyses at both baseline and follow-up; 0.72 and 0.75 respectively (**Supplementary Table A5** in **Supplementary Appendix 1**).

The conversion from the summed raw scores of the CTh scale to the estimated person parameters resulting from the Rasch model, as well as these person parameters, estimate rescaled to the original range of the CTh scale are provided in **Supplementary Appendix 1** (**Supplementary Table A6**). This allows users of the scales to choose between using the sum scores, which uses the unit of the scale, or to convert these to any of the person parameters estimates, which are continuous and equidistant scores, as preferred for whatever purpose of use.

Differences in Statistics-Related Critical Thinking at Baseline and Changes in Critical Thinking

The primary research question of the study concerned changes in statistics-related critical thinking from the start of the first semester (baseline) to the start of the second semester (follow-up). As we expected the overall change in critical thinking to differ for subgroups of students dependent on their baseline perception of the adequacy of their own mathematical knowledge for learning statistics as well as their expectation to need statistics in their future employment, we first tested baseline differences. To test whether the expected baseline subgroup

difference in critical thinking could be confirmed, we conducted a multivariate analysis of variance using a backward models search strategy, which included the primary independent variables (i.e., perception of mathematical knowledge as adequate or not and expectation to need statistics in future employment) as well as gender and age and all possible two-way interactions between the independent variables. The results showed that only the two primary independent variables defined significant differences for subgroups of students, and there was no interaction effect. Thus, we present simple tests for differences in critical thinking mean scores for subgroups defined by all four of the background variables in **Table 3**. As expected, students who perceived their mathematical knowledge to be inadequate for learning statistics scored lower on statistics-related critical thinking scores at baseline compared to the students who perceived they had an adequate level of mathematical knowledge ($p < 0.001$). Also as expected, students who did not believe they would need statistics in their future employment scored the lowest on statistics-related critical thinking compared to students who thought they might need or would definitely need statistics in future employment ($p < 0.001$).

We then proceeded to analyze the changes in critical thinking. Absolute individual changes on the CTh scale ranged from −5.3 to 5.1 points on the rescaled logit scale (**Supplementary Table A6** in **Supplementary Appendix 1**), thus showing large individual changes in critical thinking from the first to the second semester (**Figure 1**). The overall standardized effect was small and negative (−0.12), and while there were some variations for student strata defined gender, age, perceived adequacy of math knowledge to learn statistics, and expectation to need statistics in future employment, effect sizes remained small for all subgroups (**Table 4**). Thus, while there were large absolute changes in the equidistant scores resulting from the Rasch models at the individual level, effect size estimates show that there were only very small and predominantly negative effects. Our expectation that students overall would increase in critical thinking was rejected. The same was the case with our expectation that students, who at baseline did not expect to need statistics in their future employment would increase less in critical thinking than students expecting to need statistics. Only two subgroups of students showed an increase, though small, in critical thinking. These were the male students and students who at baseline perceived their mathematical knowledge as inadequate for learning statistics.

DISCUSSION

The main aim of the study was to explore changes in statistics-related critical thinking from the start of the first semester to the start of the second semester of a two-semester-long statistics course, where the curriculum contains learning objectives implying critical thinking and assessment criteria explicitly requiring critical thinking. The results showed that the student group as a whole has a low mean score of statistics-related critical thinking at baseline (i.e., a mean score of 8.05 within the possible range of 3 to 15) and that there were no significant differences related to gender or age at baseline. In a previous

TABLE 1 | Global tests of homogeneity and invariance for the Critical Thinking Scale at baseline and follow-up.

Tests of fit	Baseline			Follow-up		
	CLR	df	*p*	CLR	df	*p*
Global homogeneity[a]	9.0	11	0.622	6.4	11	0.844
Invariance						
Year cohort	16.4	11	0.128	9.8	11	0.553
Math adequacy	10.9	11	0.449	7.4	11	0.768
Stat in Future work	40.8	22	0.009[+]	30.2	22	0.113
Gender	14.2	11	0.220	7.3	11	0.775
Age groups	17.4	11	0.097	6.7	11	0.824

CTh, Critical Thinking Scale; CLR, Conditional Likelihood Ratio test.
[a] *The test of homogeneity is a test of the hypothesis that item parameters are the same for persons with low or high scores.*
[+] *The Benjamini-Hochberg adjusted critical level for false discovery rate at the 5% level was p = 0.0083 and at the 1% level p = 0.0017.*

TABLE 2 | Item fit statistics for the Critical Thinking Scale at baseline and follow-up.

	Baseline			Follow-up		
Items	Observed γ	Expected γ	*P*	Observed γ	Expected γ	*p*
CTh1	0.47	0.51	0.438	0.61	0.58	0.665
CTh2	0.62	0.52	0.035[+]	0.64	0.58	0.336
CTh5	0.47	0.52	0.389	0.51	0.58	0.246

γ = *Item-rest score correlations in the form of Goodman and Kruskal's rank correlation for ordinal items.*
[+] *The Benjamini-Hochberg adjusted critical level for false discovery rate at the 5% level was p = 0.0111 and at the 1% level p = 0.0011.*

TABLE 3 | Mean statistics-related critical thinking scores at baseline.

Group (n)	Mean	SD	*p*
All students (336)	8.05	1.90	
Gender			
Male (51)	8.36	1.73	
Female (272)	7.97	1.92	0.172
Age groups			
21 years and younger (199)	8.13	1.82	
22 years and older (131)	7.99	2.00	0.500
Math knowledge to learn statistics			
Not adequate (53)	7.01	1.89	
Adequate (281)	8.25	1.85	<0.001
Expect to need statistics in future work life			
Yes (86)	8.49	1.69	
Maybe (196)	8.07	1.81	
No (52)	7.13	2.06	<0.001[a]

p-Values for "math knowledge to learn statistics" and "expect to need statistics in future work life" are one-sided, due to expectations on the direction of differences. [a]*Post hoc pairwise tests showed that it was the group not expecting to need statistics in their future employment that differed significantly from the remaining two groups.*

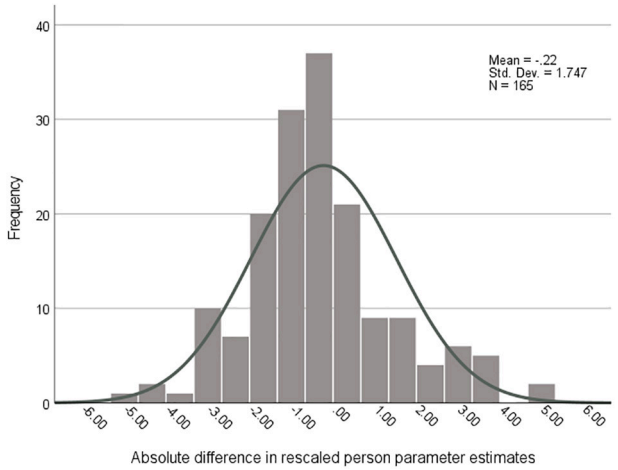


FIGURE 1 | Distribution of differences in Critical Thinking scores (rescaled person parameter estimates) from baseline to follow-up. Differences are shown as follow-up minus baseline so that positive values show an increase and negative values show a decline in critical thinking over time. Distances between any two scores are equal.

cross-cultural study employing the same instrument, statistics-related critical thinking scores were reported at the same level for both Danish and Spanish psychology students, while the mean personality psychology-related critical thinking scores were markedly higher for Danish psychology students, but not the Spanish students (Nielsen et al., 2021). This might very tentatively suggest that domain-specific critical thinking at the start of a semester course varies not only with specific domains with the same academic discipline but also with culture. Two other studies report statistics-related critical thinking at higher levels at the start of a semester course in statistics using different instruments. Bensley et al. (2010) report medium-level scores on one of their subscales for critical thinking, i.e., the argument analysis scale, at the start of a semester prior to introducing different instructional methods to enhance critical thinking in a research methods course for psychology students. Cheng et al. (2018) report high baseline scores on four single items tapping into four dimensions of critical thinking at the start of introductory statistics classes for students from various academic disciplines. The current results open interesting new avenues of research into domain-specific critical thinking in higher education and its development, both within and between academic disciplines, and across cultures.

Furthermore, we found strong evidence that the baseline statistics-related critical thinking scores differed dependent on students' perception of the adequacy of their mathematical knowledge for learning statistics as well as whether they expected to need statistics in their future work life. Thus, students who perceived their mathematical knowledge to be inadequate for learning statistics had a lower level of critical thinking than students perceiving their mathematical knowledge as adequate. The Danish psychology program requires level B mathematics[3] for being admitted to the program but does not require a particular grade for admittance, and thus students can enter with a "just pass"-grade of 02 (see **Supplementary Appendix 2** for the Danish grading scale). As the psychology program is very hard to get into and there is a fixed number of places available, however, only students with a very high-grade point average get in. We assumed that the lack of insight into the field of statistics presents just 1 month into the statistics course and their first semester in the Bachelor of Psychology program might be reflected in their perception of their mathematical basis as adequate or inadequate for learning statistics, and thus also for their inclination toward statistics-related critical thinking at this early point. However, in hindsight, more information on this issue should have been gathered. With regard to baseline differences dependent on the students' expectations to need statistics in their future work life, results were also in line with our assumption, i.e., that confidence in needing statistics in the future would be associated with an enhanced inclination toward statistics-related critical thinking compared to students who were confident they would not need statistics in the future. The results not only confirmed our assumptions but also showed that it was the group of students that were certain to not need statistics in their future work life, who had significantly lowered inclination toward critical thinking compared to both students thinking they might need statistics and students who were sure they would need statistics in the future. The results even showed that there was an ordered relationship in the mean scores for the three groups so that students who expected to need statistics in their future employment had the highest CTh scores, and students who did *not* expect to need statistics in their future employment had the lowest CTh scores and students who thought they might need statistics scored in between. This finding leads us to suggest that future research might explore how the interaction between expectancy-to-need statistics and initial inclination toward statistics-related critical thinking might

[3]Levels are A, B and C, with A being the highest.

TABLE 4 | Overall and stratified mean differences in critical thinking and effect sizes over time.

Group (n)	Mean difference (*p*)	Effect size	(95% CI)
All students (165)	0.23 (0.050)	−0.12	(−0.33 to 0.10)
Gender			
Male (25)	0.23 (0.313)	0.09	(−0.46 to 0.65)
Female (136)	−0.29 (0.018)	−0.16	(−0.40 to 0.07)
Age groups			
21 years and younger (104)	−0.26 (0.069)	−0.14	(−0.41 to 0.14)
22 years and older (60)	−0.16 (0.233)	−0.08	(−0.44 to 0.27)
Math knowledge to learn statistics			
Not adequate (29)	0.09 (0.405)	0.05	(−0.46 to 0.57)
Adequate (136)	−0.29 (0.021)	−0.15	(−0.39 to 0.09)
Expect to need statistics in future work life			
Yes (45)	−0.46 (0.089)	−0.24	(−0.66 to 0.17)
Maybe (101)	−0.09 (0.283)	−0.05	(−0.32 to 0.23)
No (19)	−0.40 (0.145)	−0.22	(−0.85 to 0.42)

Mean differences are shown as follow-up minus baseline so that positive values show an increase and negative values show a decline in critical thinking over time. P-values are one-sided). CI, Confidence Interval. Effect sizes are calculated using the rescaled person parameter estimates, as the distance between any two scores is equal.

be related to the outcome of statistics courses, but also to the actual need for statistics in the first employment of the graduates.

Turning to the main results of the study, namely the lack of an overall increase in statistics-related critical thinking in the first semester, this was the opposite of what we expected. Previous research on the development of statistics-related critical thinking has mainly focused on comparing teaching methods designed to enhance critical thinking with "usual" teaching methods not designed for this purpose, or by simply evaluating the enhancing effect of purposely designed teaching methods. While methods for measuring statistics-related critical thinking differ across studies as does the teaching methods evaluated results are also ambiguous, as some find no effect of the purposely designed teaching compared to the usual teaching without clarifying whether this means there was an effect or no effect for both groups (Goode et al., 2018; Setambah et al., 2019), and others a positive effect for only the students receiving the purposely designed teaching and no change for the students receiving the usual teaching (e.g., Bensley et al., 2010). On the same note, one study evaluating just the effect of a purposely designed teaching method in itself found this to enhance the statistics-related critical thinking of the students (Cheng et al., 2018). The lack of increase in the statistics-related critical thinking in the current study is thus only supported by Bensley et al. (2010), who did not find any change for their control group of psychology students. The current study is not enough to refute that meaningful instruction within a subject domain inherently will entail the development of critical thinking skills even if these are not purposely targeted with teaching activities, as suggested by Tiruneh et al. (2017). However, the current study does show that even the students have a low level of critical thinking at baseline and thus ample room for improvement, one semester's worth of university-level teaching in statistics with lectures as well as small exercise classes, where assessment criteria explicitly mention critical thinking (**Supplementary Appendix 2**), does not enhance the critical thinking of the students as a whole. Thus, Tiruneh et al.'s (2017) notion cannot be supported by our research, as we do not find an overall positive effect on statistics-related critical thinking over the semester. Our study, however, points to the need for developing further research to explore the factors involved in the development of statistics-related critical thinking skills.

The subgroup results in the current study also showed small effects for all subgroups, and thus did not divulge any clear patterns with regard to student factors related to the development of statistics-related critical thinking. The findings, which might suggest areas of interest for future research are the differences in the direction of the development in statistics-related critical thinking found across gender and across perceptions of the adequacy of mathematical knowledge for learning statistics, even if these differences in direction might be random results due to small group sizes. Thus, future research should include additional student characteristics to explore this further, e.g., characteristics such as dispositional characteristics such as personality, e.g., conscientiousness which has consistently been found to be positively associated with academic success in higher education (Richardson et al., 2012; Vedel, 2014), an association, which in relation to learning statistics might very well be mediated by statistics-related critical thinking. Motivation and academic self-efficacy, as both have been linked to student performance (Richardson et al., 2012) and student anxiety (Tahmassian and Jalali Moghadam, 2011; Nguyen and Deci, 2016) and statistics-related anxiety is well-documented among students from other disciplines taking statistics courses and the detrimental effect of anxiety on learning is well-known. We thus propose that motivational factors as well as the belief in one's own ability to learn statistics might moderate the development of statistics-related critical thinking and that this is certainly worth investigating in the future.

Dispositional measures and other student characteristics might also be successfully employed in future studies of increases

and decreases at the individual level, and preferably with more points of measurement (three to six), as they might then contribute to explaining individual student trajectories with regard to statistics-related critical thinking and whether these are one-directional across multiple points of measurement. Such student characteristics might also be useful with larger samples to explore whether certain student profiles are associated with an increase and certain profiles with a decrease in statistics-related critical thinking. In addition, future studies might link to the current research and expand these by including students from other academic disciplines than psychology.

The study has four major strengths. The first strength is that the results concerning change stand on a very strong psychometric foundation as the CTh scale fitted the Rasch model both at baseline and at follow-up and as the scale was very well targeted to the study population of first-year Danish Psychology Bachelor students taking their statistics course. As such, we know that the CTh scale possesses the psychometric properties, we aimed for and that the results of the change analyses and both the differences at baseline and the effect sizes are not biased due to a general lack of invariance or differential item function. The second strength lies in the use of standardized effect sizes to assess changes in statistics-related critical thinking, as this makes it possible for future studies using the same instrument under different conditions to compare the results. The third strength of the study is its contribution to the body of knowledge on the so-called "natural" development of domain-specific critical thinking, by showing that there was no overall increase in critical thinking. The contribution is important, as it showed that even though critical thinking was explicitly mentioned in the assessment criteria and implicitly in the learning objectives for the course as well as the overall competencies to be achieved through the program, no overall increase was found nor were there subgroup-specific increases of any significance. However, equally important is the finding that there were rather large absolute changes in critical thinking at the individual level, both in the form of increases and decreases, as are the findings of baseline differences dependent on students' perception of the adequacy of their mathematical knowledge for learning statistics and their expectancy to need statistics in their future work life.

Likewise, the study has three limitations. The first is the sample size and the subgroup distributions in the longitudinal sample, as this did not allow us to explore any possible interaction effects by stratifying on more than one grouping variable at a time. Thus, it was not possible to explore with any certainty how the differences in statistics-related critical thinking at baseline might affect the development. The second limitation might be considered to be the CTh scale itself, as it only comprises three items covering the purposeful and inquiring aspect of CTh common to three major definitions of critical thinking. However, as thoroughly demonstrated with the content and construct validity analyses conducted by Nielsen et al. (2021), there is no loss in content validity by eliminating two items from the original scale from the MSLQ, as these did in fact not measure critical thinking – not content-wise nor when considering the dimensionality issue. As the brief version, we employed in this study, furthermore fitted the strictest measurement model (i.e., the Rasch model) and was well targeted to the student population in this study and the cross-cultural sample in the study by Nielsen et al. (2021), we do not find the brief version to be inferior to the five-item version from the MSLQ, rather the contrary. However, we do recognize that other and longer instruments might be preferred by other researchers and that such instruments, if appropriately validated, can offer more precise measurement. The third limitation is that we did not collect any qualitative and detailed information from the professor or the students, which might have contributed to a better understanding of the lack of overall increase in statistics-related critical thinking as well as the results at the individual level.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: 10.5281/zenodo.6401225.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

TN construed the study and conducted the analyses. All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

FUNDING

The authors have received to funding for the study.

ACKNOWLEDGMENTS

The authors thank the students and the professor who kindly lend their time to provide data, Martin Andersen for collecting part of the data, Pedro Henrique Ribeiro Santiago for providing the R-code for the item maps, and Mary McGovern for translating the part of the curricula descriptions which was only available in Danish.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2022.884635/full#supplementary-material

REFERENCES

Andersen, E. B. (1973). A goodness of fit test for the Rasch model. *Psychometrika* 38, 123–140. doi: 10.1007/BF02291180

Bailin, S., Case, R., Coombs, J. R., and Daniels, L. B. (1999). Common misconceptions of critical thinking. *J. Curric. Stud.* 31, 269–283. doi: 10.1080/002202799183124

Beauchamp, M. K., Jette, A. M., Ward, R. E., Kurlinski, L. A., Kiely, D., Latham, N. K., et al. (2015). Predictive validity and responsiveness of patient-reported and performance-based measures of function in the Boston RISE study. *J. Gerontol. Med. Sci.* 70, 616–622. doi: 10.1093/gerona/glu227

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *J. R. Statist. Soc. Series B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Bensley, D. A., and Baxter, C. (2006). *The Critical Thinking in Psychology Test. Unpublished manuscript*. Frostburg, MD: Frostburg State Univerisity.

Bensley, D. A., Crowe, D. S., Bernhardt, P., Buckner, C., and Allman, A. L. (2010). Teaching and Assessing Critical Thinking Skills for Argument Analysis in Psychology. *Teach. Psychol.* 37, 91–96. doi: 10.1080/00986281003626656

Cheng, S., Ferris, M., and Perolio, J. (2018). An innovative classroom approach for developing critical thinkers in the introductory statistics course. *Am. Statist.* 72, 354–358. doi: 10.1080/00031305.2017.1305293

Cox, D. R., Spjøtvoll, E., Johansen, S., van Zwet, W. R., Bithell, J. F., and Barndorff-Nielsen, O. (1977). The Role of Significance Tests [with Discussion and Reply]. *Scand. J. Stat.* 4, 49–70.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334. doi: 10.1016/0020-7489(93)90092-9

De Jager, T. (2012). Can first year students' critical thinking skills develop in a space of three months?. Procedia. *Soc. Behav. Sci.* 47, 1374–1381. doi: 10.1016/j.sbspro.2012.06.829

Facione, P. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Research findings and recommendations*. Newark, NJ: American Philosophical Association.

Facione, P. A., Facione, N. C., and Giancarlo, C. A. F. (1998). *The California Critical Thinking Disposition Inventory*. California: Academic Press, 67–79.

Goode, C. T., Lamoreaux, M., Atchison, K. J., Jeffress, E. C., Lynch, H. L., and Sheehan, E. (2018). Quantitative Skills, Critical Thinking, and Writing Mechanics in Blended Versus Face-to-Face Versions of a Research Methods and Statistics Course. *Teach. Psychol.* 45, 124–131. doi: 10.1177/0098628318762873

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *Am. Psychol.* 53, 449–455. doi: 10.1037/0003-066X.53.4.449

Halpern, D. F. (2003). *Thought and knowledge: An introduction to critical thinking*. Mahwah, NJ: Erlbaum.

Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking. 5th Edn*. New York, NY: Psychology Press Taylor & Francis Group.

Halpern, D. F. (2015). *Halpern critical thinking assessment*. Austria: Schuhfried GmbH.

Hammer, S. J., and Green, W. (2011). Critical thinking in a first year management unit: the relationship between disciplinary learning, academic literacy and learning progression. *Higher Educ. Res. Dev.* 30, 303–315. doi: 10.1080/07294360.2010.501075

Holland, D. F., Kraha, A., Zientek, L. R., Nimon, K., Fulmore, J. A., Johnson, U. Y., et al. (2018). Reliability Generalization of the Motivated Strategies for Learning Questionnaire: a Meta-Analytic View of Reliability Estimates. *SAGE Open* 8, 1–29. doi: 10.1177/2158244018802334

Kanbay, Y., Isik, E., Aslan, O., Tektas, P., and Kilic, N. (2017). Critical Thinking Skill and Academic Achievement Development in Nursing Students: four-year Longitudinal Study. *Am. J. Educ. Res. Rev.* 2;12. doi: 10.28933/ajerr-2017-12-0501

Kaya, H., Şenyuva, E., and Bodur, G. (2017). Developing critical thinking disposition and emotional intelligence of nursing students: a longitudinal research. *Nurse Educ. Today* 48, 72–77. doi: 10.1016/j.nedt.2016.09.011

Kelderman, H. (1984). Loglinear Rasch model tests. *Psychometrika* 49, 223–245.

Kreiner, S. (2003). *Introduction to DIGRAM*. Copenhagen: Department of Biostatistics, University of Copenhagen.

Kreiner, S. (2007). Validity and objectivity. Reflections on the role and nature of Rasch Models. *Nordic Psychol.* 59, 268–298. doi: 10.1027/1901-2276.59.3.268

Kreiner, S. (2011). A Note on Item-Restscore Association in Rasch Models. *Appl. Psycholog. Meas.* 35, 557–561. doi: 10.1177/014662161141022

Kreiner, S. (2013). "The Rasch model for dichotomous items," in *Rasch Models in Health*, eds K. B. Christensen, S. Kreiner, and M. Mesbah (London: ISTE Ltd, Wiley), 5–26. doi: 10.1002/9781118574454.ch1

Kreiner, S., and Christensen, K. B. (2004). Analysis of local dependence and multidimensionality in graphical loglinear Rasch models. *Commun. Stat. Theory Methods* 33, 1239–1276. doi: 10.1081/sta-120030148

Kreiner, S., and Nielsen, T. (2013). *Item analysis in DIGRAM 3.04. Part I: Guided tours. Research report 2013/06*. Denmark: University of Copenhagen, Department of Public Health.

Kreiner, S. and Christensen, K. B. (2013). "Person Parameter Estimation and Measurement in Rasch Models", in Rasch Models Health, eds K.B. Christensen, S. Kreiner and M. Mesbah. (London, ISTE and John Wiley & Sons, Inc.) 63-78. doi: 10.1002/9781118574454.ch4

Krosnick, J. A., and Fabrigar, L. R. (1997). "Designing rating scales for effective measurement in surveys," in *Survey measurement and process quality*, eds L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, et al. (New York, NY: John Wiley), 141–164. doi: 10.1002/9781118490013.ch6

Kuhn, D. (1999). A developmental model of critical thinking. *Educ. Res.* 28, 16–25. doi: 10.2307/1177186

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front. Psychol.* 4:863. doi: 10.3389/fpsyg.2013.00863

Lau, J. Y. F. (2015). "Metacognitive education: Going beyond critical thinking," in *The palgrave handbook of critical thinking in higher education*, eds M. Davies and R. Barnett (New York, NY: Palgrave Macmillan), 373–389. doi: 10.1057/9781137378057

Maitland, A. (2009). Should I label all scale points or just the end points for attitudinal questions? *Survey Pract.* 4, 1–4. doi: 10.29115/SP-2009-0014

McGuirk, J. (2021). Embedded rationality and the contextualization of critical thinking. *J. Philosop. Educ.* 55, 606–620. doi: 10.1111/1467-9752.12563

McPeck, J. (1992). "Thoughts on subject specificity," in *The generalizability of critical thinking: Multiple perspectives on an educational ideal*, ed. S. Norris (New York, NY: Teachers College Press), 198–205.

Menold, N., Kaczmirek, L., Lenzner, T., and Neusar, A. (2014). How Do Respondents Attend to Verbal Labels in Rating Scales? *Field Methods* 26, 21–39. doi: 10.1177/1525822X13508270

Ministry of Higher Education and Science (2021). *Ansøgere og optagne fordelt på køn, alder og adgangsgrundlag*. Available online at: https://ufm.dk/uddannelse/statistik-og-analyser/sogning-og-optag-pa-videregaende-uddannelser/grundtal-om-sogning-og-optag/ansogere-og-optagne-fordelt-pa-kon-alder-og-adgangsgrundlag (accessed date 23.12.2021).

Moore, T. (2011). Critical thinking and disciplinary thinking: a continuing debate. *High. Educ. Res. Dev.* 30, 261–274. doi: 10.1080/07294360.2010.501328

Moseley, D., Baumfield, V., Elliott, J., Higgins, S., Miller, J., Newton, D. P., et al. (2005). *Frameworks for thinking: A handbook for teaching and learning*. Cambridge, UK: Cambridge University Press.

Nguyen, T. T., and Deci, E. L. (2016). Can it be good to set the bar high? The role of motivational regulation in moderating the link from high standards to academic well-being. *Learn. Indiv. Diff.* 45, 245–251. doi: 10.1016/j.lindif.2015.12.020

Nielsen, T. (2018). The intrinsic and extrinsic motivation subscales of the Motivated Strategies for Learning Questionnaire: a Rasch-based construct validity study. *Cog. Educ.* 5, 1–19. doi: 10.1080/2331186X.2018.1504485

Nielsen, T. (2020). The Specific Academic Learning Self-efficacy and the Specific Academic Exam Self-Efficacy scales: construct and criterion validity revisited using Rasch models. *Cog. Educ.* 7, 1–15. doi: 10.1080/2331186X.2020.1840009

Nielsen, T., Makransky, G., Vang, M. L., and Dammeyer, J. (2017). How specific is specific self-efficacy? A construct validity study using Rasch measurement models. *Stud. Educ. Eval.* 57, 87–97. doi: 10.1016/j.stueduc.2017.04.003

Nielsen, T., Martínez-García, I., and Alastor, E. (2021). Critical Thinking of Psychology Students: a Within- and Cross-Cultural Study using Rasch models. *Scand. J. Psychol.* 62, 426–435. doi: 10.1111/sjop.12714

Nielsen, T., Martínez-García, I., and Alastor, E. (2022). "Psychometric properties of the Spanish translation of the Specific Academic Learning Self-Efficacy and the Specific Academic Exam Self-Efficacy scales in a higher education context," in *Academic Self-efficacy in Education: Nature, Measurement, and Research*, eds

M. S. Khine and T. Nielsen (New York, NY: Springer), 71–96. doi: 10.1007/978-981-16-8240-7_5

Özelçi, S. Y., and Çalışkan, G. (2019). What is critical thinking? A longitudinal study with teacher candidates. *Internat. J. Eval. Res. Educ.* 8, 495–509. doi: 10.11591/ijere.v8i3.20254

Paul, R., and Elder, L. (2005). *A guide for educators to Critical Thinking Competency Standards*. Santa Barbara, CA: Foundation for Critical Thinking.

Pintrich, P. R., Smith, D. A. F., Garcia, T., and McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ). (Technical Report No. 91-8-004)*. Ann Arbor, MI: The Regents of the University of Michigan.

Ralston, P. A., and Bays, C. L. (2015). Critical thinking development in undergraduate engineering students from Freshman Through Senior Year: a 3-Cohort Longitudinal Study. *Am. J. Eng. Educ.* 6, 85–98. doi: 10.19030/ajee.v6i2.9504

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

Ren, X., Tong, Y., Peng, P., and Wang, T. (2020). Critical thinking predicts academic performance beyond general cognitive ability: evidence from adults and children. *Intelligence* 82:101487. doi: 10.1016/j.intell.2020.101487

Richardson, M., Abraham, C., and Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychol. Bull.* 138, 353–387. doi: 10.1037/a0026838

Ricketts, J. C., and Rudd, R. D. (2005). Critical thinking of selected youth leaders: the efficacy of critical thinking dispositions, leadership and academic performance. *J. Agricult. Educ.* 46, 32–43.

Saenab, S., Zubaidah, S., Mahanal, S., and Lestari, S. R. (2021). ReCODE to Re-Code: an instructional model to accelerate students' critical thinking skills. *Educ. Sci.* 11:2. doi: 10.3390/educsci11010002

Sahanowas, S. K., and Halder, S. (2020). Critical thinking disposition of undergraduate students in relation to emotional intelligence: gender as moderator. *Heliyon* 6:e05477. doi: 10.1016/j.heliyon.2020.e05477

Setambah, M. A. B., Tajudin, N. M., Yaakob, M. F. M., and Saad, M. I. M. (2019). Adventure Learning in Basics Statistics: impact on Students Critical Thinking. *Internat. J. Instruct.* 12, 151–166. doi: 10.29333/iji.2019.12310a

Stupnisky, R. H., Renaud, R. D., Daniels, L. M., Haynes, T. L., and Perry, R. P. (2008). The interrelation of first-year college students' critical thinking disposition, perceived academic control and academic achievement. *Res. High. Educ.* 49, 513–530. doi: 10.1007/s11162-008-9093-8

Tahmassian, K., and Jalali Moghadam, N. (2011). Relationship between self-efficacy and symptoms of anxiety, depression, worry and social avoidance in a normal sample of students. *Iran. J. Psychiatry Behav. Sci.* 5, 91–98.

Thomas, T. (2011). Developing first year students' critical thinking skills. *Asian Soc. Sci.* 7, 26–35. doi: 10.5539/ass.v7n4p26

Tiruneh, D. T., De Cock, M., and Elen, J. (2017). Designing Learning Environments for Critical Thinking: examining Effective Instructional Approaches. *Internat. J. Sci. Mathem. Educ.* 16, 1065–1089. doi: 10.1007/s10763-017-9829-z

Vedel, A. (2014). The Big Five and tertiary academic performance: a systematic review and metaanalysis. *Personal. Indiv. Diff.* 71, 66–76. doi: 10.1016/j.paid.2014.07.011

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 22 July 2022
DOI 10.3389/feduc.2022.886139

OPEN ACCESS

EDITED BY
Tarja Tuononen,
University of Helsinki, Finland

REVIEWED BY
Anna Parpala,
University of Helsinki, Finland
Johanna Rämö,
University of Eastern Finland, Finland

*CORRESPONDENCE
Anne Virtanen
anne.virtanen@jyu.fi

SPECIALTY SECTION
This article was submitted to Higher Education, a section of the journal Frontiers in Education

RECEIVED 28 February 2022
ACCEPTED 01 July 2022
PUBLISHED 22 July 2022

CITATION
Virtanen A, Lauritsalo K, Mäkinen T, Hurskainen H and Tynjälä P (2022) The role of positive atmosphere on learning generic skills in higher education—Experiences of physical education students.
Front. Educ. 7:886139.
doi: 10.3389/feduc.2022.886139

The role of positive atmosphere on learning generic skills in higher education—Experiences of physical education students

Anne Virtanen[1]*, Kirsti Lauritsalo[2], Tommi Mäkinen[2], Heikki Hurskainen[3] and Päivi Tynjälä[1]

[1]Finnish Institute for Educational Research, University of Jyväskylä, Jyväskylä, Finland, [2]Department of Teacher Education, Faculty of Education and Psychology, University of Jyväskylä, Jyväskylä, Finland, [3]Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

The purpose of this study is to examine what kind of role does positive atmosphere play in learning generic skills. The study was carried out in the final year of Finnish physical education (PE) teacher education where teachers of this study module have developed their pedagogical practices for a long time to foster and maintain the positive learning atmosphere of the course. In this study the learning atmosphere was examined from the perspectives of students addressing the following research questions: (1) How do the students perceive and experience the constructed learning atmosphere? (2) What generic skills do PE students report to learn during the final year study module? (3) What kind of relationship exists between the perceived learning atmosphere and learning generic skills? The data of this study were collected *via* an internet questionnaire (N = 189, n = 125) and interviews (n = 19) and analyzed applying qualitative content analysis and using statistical methods. According to the results, the PE students felt that the learning atmosphere of their final year study module was warm and conversational; it was easy to get their own voice heard during the final year. The PE students reported that they had learned versatile and plenty of different generic skills, particularly various social skills. There were statistically significant associations between positive atmosphere and some generic skills, such as the development of creativity, but they were not very strong. To understand this finding, the instruments of the study are discussed, especially from the perspective of the development of the questionnaire statements. Although the relationship between perceived positive atmosphere and learning generic skills was only moderate, the findings are promising. For example, the good practices documented in this study, such as how to construct a warm educational atmosphere, may be applied when developing other study modules.

KEYWORDS

generic skills, atmosphere, learning, students, physical education, higher education

Introduction

Generic skills in the world of work and education

Generic skills (also referred to as generic attributes, generic capabilities, key skills, or soft skills), such as critical thinking skills, problem-solving skills, social skills, and skills concerning creativity, are greatly needed and sought after in today's world of work and education (Kember, 2009; Lavi et al., 2021; Virtanen and Tynjälä, 2022). Kember (2009, p. 52), one of the most influential researchers of generic skills, has stated aptly that "few, if any, would disagree with the need for graduates to possess a range of generic capabilities to equip them for lifelong learning in today's knowledge-based society." In the world of work, the need for generic skills is seen, for example, in competency mappings about the current occupational requirements (e.g., Rekola et al., 2018; Sudirman et al., 2020), in studies on requirements of graduates (Kalfa and Taksa, 2015; Clarke, 2018; Cotronei-Baird, 2020), and in lists and models that describe the needs and demands of future work life (e.g., Fadel et al., 2015; Forbes, 2020). In education, generic skills are integrated into different national and international frameworks (e.g., Gordon et al., 2009; EQF, 2018; P21, 2020; ATC21S, 2022). These frameworks have been formulated for the evaluation, comparability, and development of education programs. Based on these frameworks, generic skills, or twenty-first century skills (Tight, 2021), are incorporated into many countries' curricula (e.g., Ahonen and Kinnunen, 2015).

Although the need for generic skills is recognized in both working life and education, little is known about how these skills are learned and how they should best be taught. Specific generic skills, such as presentation skills and scientific writing, can be prepared as separate courses. In recent years, however, it has been demonstrated that learning generic skills has an interrelationship with the pedagogy used (Kember, 2009; Anthony and Garner, 2016; Virtanen and Tynjälä, 2019, 2022). In other words, generic skills can be learned as a part of regular teaching, but certain forms of teaching and learning are needed. For example, according to Kember (2009), students' intellectual capabilities develop well when teaching activities demand critical thinking, self-managed learning, and problem-solving. Similarly, Virtanen and Tynjälä (2019) found that the pedagogical practices in which students assess knowledge critically support the development of students' critical thinking skills.

Group work and other methods that demand working together foster students' collaborative skills (Kember, 2009; Kostiainen et al., 2018; Mäkinen et al., 2022; Virtanen and Tynjälä, 2022). This is a logical and self-evident finding, but it is worth stressing for developing teaching and pedagogical practices in different educational institutes. Tuononen et al. (2019) found that the graduates most often mentioned challenges related to their social and presentation skills which they felt were not taught well enough in their institutions. Research has shown that pedagogy that utilizes interactive and collaborative work fosters not only the learning of social skills but also other generic skills. For example, Virtanen and Tynjälä (2019) found that teaching that requires working together develops students' decision-making and problem-solving skills. In their study, acting at the interface between theory and practice also played an important role in explaining students' learning of generic skills. Anthony and Garner (2016) found similar results. They noticed that assignments involving real-world applications and examples, and projects that engaged students or had practical applications, were the most helpful activities for business students to learn generic skills.

The studies presented above are in line with Kember's statement: "It might be noted that a very common model of university teaching—a professor lecturing to students who sit listening—does not provide practice in the generic capabilities" (2009, p. 53). In contrast, the practices that seem to support and foster the learning of generic skills demand students to participate rather than just sit and listen. Many studies also suggest that, in interactive and collaborative learning and teaching, teachers must put effort into constructing a positive and safe learning atmosphere (Vila et al., 2012; Kiuru et al., 2015; Mäkinen et al., 2022). Next, we will review research on the topic.

The role of a positive and safe atmosphere for learning

Feeling safe is a prerequisite for all learning. Recent findings support the view that the quality of the learning atmosphere has an even greater impact on learning than the ability of the learner (Pakarinen et al., 2014; Ratmawaty, 2018; Calavia et al., 2021; Visiers-Jiménez et al., 2021). For example, a study carried out in six countries reported that the quality of the supervisory relationship and pedagogical atmosphere in which learning took place were closely linked to the students' learning success and their satisfaction with their education (Visiers-Jiménez et al., 2021). A positive and safe atmosphere can affect a student's attitude toward learning, which is considered the most important factor for improving their achievement and confidence (Ratmawaty, 2018; Suyatno et al., 2019). Therefore, increased attention is being paid to the atmosphere in which learning will take place.

In many studies on the learning atmosphere, particular attention is paid to the teacher-student relationship (e.g., Pianta et al., 2008; Kember, 2009; Kiuru et al., 2015; Kostiainen et al., 2018; Pöysä et al., 2019). For example, Ratmawaty (2018) states that a good classroom atmosphere is characterized by a pleasant interaction between the teacher and student during the learning process. In studies of work-related learning, it has been shown that it is the job supervisor rather than the teacher who has the

role of the important person (e.g., Visiers-Jiménez et al., 2021). In other words, the construction of a learning atmosphere seems to emanate from the adult that guides the student's learning; for example, from a teacher or job supervisor.

Positive and safe atmosphere offers the basis for productive collaboration (e.g., Alles et al., 2019), and as we concluded above, the learning of generic skills seems to require interactive and collaborative pedagogy. However, we found only a limited number of studies on the relationship between learning atmosphere and the learning of generic skills. For example, Zeng (2021) found that there is a positive association between the two; this association, however, was not very strong. Calavia et al. (2021) perceived that a positive learning atmosphere fosters students' creativity, which Virtanen and Tynjälä (2019) also confirmed.

The aim of this study is to examine more closely the relationship between a positive learning atmosphere and learning generic skills. The study should be, then, carried out in such an environment where the learning atmosphere is known to be good. Therefore, this study was conducted in a context where extensive development has resulted in a learning atmosphere where students would feel safe, trust each other and have the courage to take part in class by engaging in discussion with the teacher and other students. This kind of developmental work is in line with current studies on a positive learning atmosphere, which emphasizes trust, general appreciation (i.e., talking to each other politely, listening to each other, and letting each other finish speaking) and pleasant teacher-student and student-student interactions during the learning process (Kiuru et al., 2012; Ratmawaty, 2018; Alles et al., 2019; Suyatno et al., 2019). The above-mentioned characterization also describes how a positive learning atmosphere is understood in this study. Next, the context of this study and the developmental work concerning the learning atmosphere will be described in depth.

Context of the study

The research context was the final year of Finnish physical education (PE) teacher education, when most PE students carry out the major part of their pedagogical studies. In Finland, PE teacher education is a master's degree program (300 ECTS credits), which is highly valued; only 5–6 percent of applicants are accepted annually. PE students carry out one-third of their pedagogical studies in their own faculty where they are integrated as a part of their own major subject, sport pedagogy. Two-thirds of the pedagogical studies are taken in the Faculty of Education, in the Department of Teacher Education. This study focuses particularly on the separate pedagogical study module that PE students carry out at the Department of Teacher Education in their final academic year. It is a large study module (28 ECTS credits) and lasts one academic year. There are 45–60 students in these modules annually, taught by two professors (later in this text called teachers).

The final-year study module consists of three theoretical courses and three practical training periods, taken simultaneously. The final module has been developed over several years (Klemola et al., 2013; Tynjälä et al., 2016; Lauritsalo et al., 2019). As this final-year study module is a key component in the professional development of teachers, all courses and training periods have been developed in a way that best supports the developing teacher.

The pedagogical practices of the final year study module are based on self-determination theory (Ryan and Deci, 2000) and Gordon's (2003) interaction model (Lauritsalo et al., 2019). Feeling safe and belonging to the community are core feelings that teachers hope students experience during the pedagogical practices of the final year. Therefore, they consciously constructed the final year study module to incorporate practices that support and foster a sense of community and solidarity. For example, at the beginning of the year, the teachers assigned different duties and responsibilities to their students. While one PE student group was responsible for the wellbeing of peer students, another was in charge of guiding students' learning tasks and a third organized the party at the end of the year. In addition, one female student was elected and named "mama" and one male student "papa" for the year. The mama and the papa of the final year acted as messengers between the students and the teachers. The aim of these above-mentioned examples of the responsibilities was to support and maintain a sense of community throughout the year.

Teachers stressed that student commitment to the pedagogical practices in the final year was important because there were only two teachers for every 45–60 students annually, and their teaching in the final year occurred primarily through discussion and collaboration. As such, students were required to trust and work with each other throughout the year. The teachers considered the first week of the final year critical. During that time, they implemented plenty of instruction and face-to-face teaching and included several practices to help the students trust and get to know each other better. The teachers conveyed that the work and actions in this first week were crucial to the success of the entire study module. After the first week, nurturing the learning atmosphere is a goal in every teaching situation throughout the year. Students were asked to take care of their peers and organize different activities supporting the learning atmosphere also in their spare time.

At the beginning of the final year, teachers also introduced students to the Three K model, which is a pedagogical practice based on three different values. Values are respect (*kunnioitus*), encountering (*kohtaaminen*), and presence (*kiireettömyys*) (the letter "K" refers to the first letter of each word in Finnish). The model's values were carried out in all formal and informal situations and meetings throughout the year. Respect referred to every person's importance and value: "I'm important, you're

important, and we're important;" encountering the described thought, "With an open mind and interest, I watch other people and new things. I ask for help, and I give help;" and presence denoted the presence of all students and teachers in all formal and informal situations: "I'm present right now." Teachers returned to this model several times throughout the year to remind the class of its existence.

Aim and research questions of the study

Teachers of this final-year study module have done a lot of work to develop the learning atmosphere, making this a suitable context for this study which aim is to examine the relationship between the learning atmosphere and learning generic skills. Before examining this relationship, the experiences of PE students with the learning atmosphere and the learning of generic skills during this final year study module are investigated. More specifically, we addressed the following research questions: (1) How do the students perceive and experience the constructed learning atmosphere? (2) What generic skills do PE students report to learn during the final year study module? (3) What kind of relationship exists between the perceived learning atmosphere and learning generic skills?

Materials and methods

This study was carried out in a collaboration between teachers and researchers. Teachers planned and took care the teaching, whereas researchers conducted the study (data gathering, analyses). Despite having these different tasks, our experience was that we conducted this study together. For example, teachers helped the researchers to collect data as a part of their own teaching, analysis was discussed jointly and teachers felt that they had received valuable information for the development of their teaching.

This study's data were collected *via* an internet questionnaire and interviews. Participation in this study was voluntary. The questionnaire data was gathered from all PE students (N = 189) of four final year study modules during four different years; 125 of them (66%) answered the questionnaire. The questionnaire has been developed and tested in different educational contexts over the course of many years (e.g., Tynjälä and Virtanen, 2005; Virtanen et al., 2014; Virtanen and Tynjälä, 2019). The questionnaire consisted of two different themes: (1) learning outcomes, which included questions about learning generic skills; and (2) learning processes, which included questions related to the integration of theory and practice, assessment and feedback, the nature of teaching and learning, and the learning atmosphere. In the study, only certain parts of the themes of the questionnaires were used: the 20 highest-scoring generic skills from the theme of learning outcomes, and statements relating to the learning atmosphere from the theme of learning processes.

The questionnaire was analyzed using different statistical methods (e.g., a comparison of the mean values, Pearson's correlation coefficient, and regression analysis). First, using a five-point scale, students were asked to assess the learning atmosphere of their final year study module with the help of six statements. Statements are based on current studies on a positive learning atmosphere, which emphasize trust, general appreciation and pleasant teacher-student and student-student interactions during the learning process (Kiuru et al., 2012; Ratmawaty, 2018; Alles et al., 2019; Suyatno et al., 2019). These statements, with their received mean values and standard deviations, are presented in **Table 1**. For further analysis, the aggregate scale "Positive learning atmosphere" (α = 0.68) was formed from the statements shown in **Table 1**. This aggregate scale was used in the correlation analysis (Pearson correlation coefficient), where the relationship between positive learning atmosphere and learning generic skills was examined. This relationship was also assessed *via* regression analysis in a confirmatory setting. In this design, the seven generic skills (see **Table 2**) were the dependent variables and the aggregate scale "Positive learning atmosphere" was the independent variable. The regression analysis followed a stepwise pattern so that the final model included only the variables that were associated with strong explanatory factors. Only the highest coefficient of determination is reported, as all others were minor (R^2 = 0.91–0.96). Second, students were asked to assess their learning of generic skills during the final year study module on a five-point scale (1 = nothing, 5 = a great deal). Altogether, the list of generic skills included 55 different skills. It included 22 basic academic skills (BAS), such as critical thinking skills, problem-solving skills, and skills for knowledge acquisition; 12 social skills (SS), such as interaction skills, collaboration skills, and the ability to listen to others; and 19 other skills (OS), such as independent working skills, the ability to operate in

TABLE 1 Mean values of statements describing the learning atmosphere of the final year study module assessed by PE students.

Statements concerning learning atmosphere of the final year study module	Mean values min. 1, max. 5	Standard deviations (*SD*)
Communication with the teacher felt natural.	4.81	0.40
We had good team spirit in this course.	4.75	0.44
Collaboration with other students was smooth.	4.74	0.44
The threshold to ask for clarifications was low.	4.71	0.59
It was easy to get one's own voice heard during the final year.	4.58	0.64
It was easy to share one's own opinions and thoughts.	4.48	0.78

TABLE 2 Associations between a perceived positive learning atmosphere and learning of the generic skills (Pearson Correlation Coefficient).

Generic skills	Correlations with positive learning atmosphere
Resourcefulness, innovativeness, or creativity (OS)	0.404**
Planning one's own career (OS)	0.310**
Basic skills of one's occupation/field (BAS)	0.309**
Continuing learning skills (OS)	0.308**
Written communication skills (BAS)	0.307**
Oral communication skills (BAS)	0.304**
Project-work skills (BAS)	0.302**

BAS, Basic academic skills; SS, Social skills; OS, Other skills; **p < 0.01.

TABLE 3 Mean values of variables describing perceived learning of generic skills among PE students during their final year study module (n = 119).

Generic skills	Mean values (min. 1, max. 5)	Standard deviations (*SD*)
Interaction skills (SS)	4.48	0.68
Holistic thinking (BAS)	4.39	0.65
Ability to listen to others (SS)	4.36	0.70
Seeing things from the perspective of others (SS)	4.31	0.67
Increased awareness of one's know-how (OS)	4.26	0.71
Taking responsibility for one's work (OS)	4.24	0.77
Developing an overall picture of one's field (BAS)	4.18	0.75
Collaboration skills (SS)	4.18	0.78
Ability to apply learned skills and knowledge in different situations (OS)	4.17	0.68
Assessing one's own work (OS)	4.17	0.68
Ability to operate in new situations (OS)	4.14	0.71
Planning and organization skills (BAS)	4.14	0.76
Basic skills of one's occupation/field (BAS)	4.14	0.78
Deep understanding of things learned (BAS)	4.13	0.70
Increasing one's self-confidence (OS)	4.07	0.75
Ability to evaluate the actions of others (SS)	4.01	0.76
Oral communication skills (SS)	4.00	0.75
Independent working skills (OS)	4.00	0.83
Critical thinking skills (BAS)	3.97	0.79
Continuing learning skills (OS)	3.97	0.80

BAS, Basic academic skills; SS, Social skills; OS, Other skills.

new situations, and increasing one's initiative. In addition, our questionnaire included two negative outcomes: disadvantages of the field and bad practices. Due to our extensive list of generic skills, we reported only the 20 highest scores of generic skills assessed by students (Table 3).

Interviews (n = 19) with PE students involved three different final year study modules. Teachers asked volunteers for the interview. The researcher contacted the prospective interviewees, informed them of the purpose of the study and carried out the interviews, each of which lasted about 45–60 min. These interviews were collected 6 months after the end of the final year study module. Student interviews were intentionally carried out several months after the final year study module in order to make sure that students would have an overall impression of the final year. Student interviews included questions about structure, forms of activities, teachers, learning atmosphere, assessment and feedback, practical training of the final year, and the role of the final year as a part of the PE qualification. Specific pedagogical practices covered in the final year, such as the Three K model, and different duties and responsibilities of PE students, were discussed in the interviews as well. All interviews were transcribed verbatim.

The transcribed interviews were analyzed by applying qualitative content analysis (Elo and Kyngäs, 2008). Students' answers to questions concerning the learning atmosphere and the Three K model were extracted from the interview transcripts. The analysis focused on the same issues that the teachers assessed as significant for the construction of a positive learning atmosphere. These included descriptions of the students' experiences and views of their duties and responsibilities, the first week of the final year study module and the Three K model. Summaries of these views may be found in the results of this study.

Results

RQ1: Learning atmosphere of the final year study module—Students' perspectives

As shown in Table 1, all mean values of the statements concerning learning atmosphere were rather high (4.48–4.81, max. 5). Thus, the students perceived the final year's learning atmosphere as positive. For example, PE students felt that the communication with the teacher (4.81) and with other PE students (4.74) were natural and smooth.

The findings from the student interviews are in line with the quantitative analysis regarding the learning atmosphere of the final year assessed by PE students above: the students experienced that the teachers succeeded in their goal of constructing a positive atmosphere for learning and professional development. The students reported that the learning atmosphere was warm, dialogic, and interactive. For example, one participant described the learning atmosphere of the final year: "I had such a warm and safe feeling here."

According to the students, the teachers not only put effort into the learning atmosphere at the beginning of the final year

but maintained it in various ways throughout the whole year. For example, one student said that, at the start of almost all lessons, the teachers asked the students for news or carried out certain exercises that required them to take part in action or in discussion. After this kind of active beginning, the students found it easier to take part in the interactive and collaborative work required for their lessons.

While the students initially felt that the responsibilities they received were slightly separate from their other studies and the work required of them for their final year, they found that they became naturally connected to the pedagogical practices and actions of their studies. One student said, "Responsibilities were one part of the sense of community during our final year. We were all responsible for how things happened during the year." Moreover, students said that teachers gave them responsibility over their own learning, helping them feel that the teachers completely trusted them.

According to the students, the Three K model was highly visible and present during the final year. The students felt that the teachers themselves manifested the model they developed. An excerpt from an interview with one student illustrates this perspective:

> . . .Teachers are so warm. They convey very strongly that they are interested in us, and they want to help us. And we [the students and teachers] are working together this year. They do not always have the right answers or solutions, but they are ready to work and think together with us. . . When they said that they would like to foster such things [the student is referring to the Three K model], it is clear in their actions. For example, if they talk about interaction skills and say that active listening is very important in interactions, they themselves use active listening. In other words, they don't just say that this is the thing that you must learn but act on the lessons and elsewhere they teach (Student 4).

As noted in the excerpt above, the students recognized that the teachers themselves acted in line with their own demands. The students also claimed that they would like to emulate their teachers in the future.

RQ2: Perceived learning of generic skills among physical education students during their final year study module

As shown in **Table 3**, among the 20 highest scores for skills were eight generic skills from the category "other generic skills," as well as six generic skills from the "basic academic skills" and "social skills" categories. These categories are represented by the following abbreviations: BAS, basic academic skills; SS, social skills; and OS, other skills, and are included with each generic skill provided in **Table 3**.

According to the PE students' assessments, the highest scores were given to various social skills: interaction skills (mean value was 4.48, max. 5), ability to listen to others (4.36), seeing things from the perspectives of others (4.31), and collaboration skills (4.18). From the BAS category, students particularly reported learning holistic thinking (4.39) and developing an overall picture of one's field (4.18). From the OS category, students assessed that they learned increased awareness of one's know-how (4.26) and taking responsibility for one's work (4.24). Thus, the pedagogical practices of the final year study module seem to offer opportunities to develop versatile social skills and increase students' awareness of themselves as professional agents (increased awareness of one's know-how) and their profession in the wider context (developing an overall picture of one's field, holistic thinking).

RQ3: Relationship between perceived learning atmosphere and learning generic skills

As shown in **Table 2**, it was established that a positive learning atmosphere was associated with the learning of resourcefulness, innovativeness, creativity; ability to plan a career; basic field-specific occupational skills; continuing professional development, and few generic academic skills (i.e., written and oral communication skills, project-work skills). These associations, however, were not strong ($p < 0.01$).

Due to moderate associations between the generic skills and the positive learning atmosphere, the coefficients of determination were also minor. The highest coefficient of determination was for "resourcefulness, innovativeness or creativity" ($R = 0.404$, $R^2 = 0.163$, $F = 15.768$, $p < 0.001$)—that is, the positive learning atmosphere explained 16 percent of the learning of resourcefulness, innovativeness or creativity.

Discussion

In this study, the role of a positive learning atmosphere for learning generic skills was examined during the final year of Finnish physical education (PE) teacher education. This context was suitable for examining this topic because teachers of the module have developed their pedagogical practices for a long time and were therefore able to foster a sense of community and solidarity in the class (e.g., Tynjälä et al., 2016; Lauritsalo et al., 2019; Mäkinen et al., 2022). The study focused on (1) PE students' experiences of atmosphere during the final year of the module, (2) their assessments of learning generic skills, and (3) the relationship between perceived atmosphere and learning generic skills.

The results concerning the atmosphere indicate that the teachers succeeded in their objective of constructing a

positive environment for PE students' learning and professional development. Students reported that the atmosphere was warm and conversational; they also felt that it was easy and natural to speak their minds. It could be said that the students felt equal to their teachers. They felt that their teachers trusted them completely, giving them considerable responsibility in terms of their own learning, thinking and collaboration.

Furthermore, perceived learning of generic skills was very strong. In this study, only the 20 most learned generic skills were reported. The mean values for all of them were rather high; almost all had a 4 or above, with a maximum of 5. In other words, PE students assessed they learned these skills considerably well during the year. Earlier studies have shown that interactive and collaborative teaching practices particularly support and foster the learning of generic skills (Kember, 2009; Anthony and Garner, 2016; Virtanen and Tynjälä, 2019). This study supports those findings as teaching in the course was interactive and the PE students reported learning many different social, basic academic and other skills.

There were associations between positive learning atmosphere and some generic skills, such as the development of creativity, but the correlations were only moderate ($p < 0.01$). This finding is also in line with earlier, limited number of studies on the topic (e.g., Zeng, 2021). For example, Virtanen and Tynjälä (2019) and Calavia et al. (2021) also found that a perceived positive atmosphere supports creativity. We discuss our findings more closely below.

During the final year, the module in question aimed to support, above all, the professional growth of PE teachers, whereas the learning of different generic skills is not the primary goal.

The teachers created a positive and safe atmosphere to foster their students' professional development. According to the findings of this study, the teachers succeeded in their aim of constructing a positive and safe learning atmosphere for their students. In further studies it would be interesting to examine whether a strong relationship exists between perceived positive atmosphere and professional development of student teachers. However, the findings do not show how the learning atmosphere of the final year study module should be developed in order to support and foster the learning of generic skills.

The relationship between learning atmosphere and learning generic skills was examined *via* a questionnaire. Although the analysis of the questionnaire data shows that PE students assessed the learning atmosphere very highly, the statements (see **Table 1**) were related to the atmosphere in a rather general way. For example, they did not investigate the teacher-student relationship in detail despite the fact that it has been found to be a significant factor in constructing a positive learning atmosphere in earlier studies (e.g., Kiuru et al., 2015; Kostiainen et al., 2018; Pöysä et al., 2019). Kember (2009) found that an interactive teacher-student relationship was associated with the development of higher-order thinking capabilities. For this reason, the questionnaire statements should be further developed so that they may more accurately convey the teacher-student relationship.

As to limitations of the study, attention can be paid to the fact that the results of the analysis of the questionnaire data are based on the respondents' self-assessed answers. As a method, self-assessment has been criticized because assessments of one's own actions are considered unreliable (e.g., Paulhus and Vazire, 2007). For example, Braun and Brachem (2017) state that self-reported competences cannot be equivalent to an objective assessment of competences. In the Finnish educational context, however, an interesting phenomenon has been found. For example, Virtanen (2013) found in her doctoral thesis that Finnish students assessed their learning and professional development during their workplace learning periods more critically than their teachers or workplace trainers did. Moreover, in her study, the workplace trainers saw the learning and professional development of students at workplaces in the most positive light. A strong correlation between assessments by teachers and students has also been demonstrated in other studies (e.g., Falchikov and Goldfinch, 2000; Wang et al., 2009; Asikainen et al., 2014). It is also noteworthy that the questionnaire used in this study has been developed, tested and applied over a long period of time across several studies (e.g., Tynjälä and Virtanen, 2005; Virtanen et al., 2014; Virtanen and Tynjälä, 2019). Reliability of the research findings can also be supported with the use of triangulation of the methods. The learning atmosphere findings were examined with the help of both interviews and questionnaires and both results were completely in line with each other.

To conclude, although this study found only moderate relationship between perceived positive atmosphere and learning generic skills, its other findings are promising. The research context's atmosphere was considered very positive and safe. Moreover, the learning of generic skills was assessed as strong and versatile. Therefore, the good practices documented in this study, such as how to construct a warm educational atmosphere, may be applied when developing other study modules.

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Ahonen, A., and Kinnunen, P. (2015). How Do Students Value the Importance of Twenty-first Century Skills? *Scand. J. Educ. Res.* 59, 395–412. doi: 10.1080/00313831.2014.904423

Alles, M., Seidel, T., and Gröschner, A. (2019). Establishing a positive learning atmosphere and conversational culture in the context of a video-based teacher learning community. *Prof. Dev. Educ.* 45, 250–263. doi: 10.1080/19415257.2018.1430049

Anthony, S., and Garner, B. (2016). Teaching soft skills to business students: an analysis of multiple pedagogical methods. *Bus. Prof. Commun. Q.* 79, 360–370. doi: 10.1177/2329490616642247

Asikainen, H., Virtanen, V., Postareff, L., and Heino, P. (2014). The validity and students' experiences of peer assessment in a large introductory class of gene technology. *Stud. Educ. Eval.* 40, 197–205. doi: 10.1016/j.stueduc.2014.07.002

ATC21S (2022). *Assessment & Teaching of 21st Century Skills*. Available online at: http://www.atc21s.org/ (accessed February 28, 2022).

Braun, E., and Brachem, J.-C. (2017). "The labour market's requirement profiles for higher education graduates," in *Higher Education Transitions. Theory and Research*, eds E. Kyndt, V. Dochne, K. Trigwell, and S. Lindblom-Ylänne (London: Routledge), 219–237. doi: 10.4324/9781315617367

Calavia, M. B., Blanco, T., and Casas, R. (2021). Fostering creativity as a problem-solving competence through design: think-Create-Learn, a tool for teachers. *Think. Skills Creat.* 39:100761. doi: 10.1016/j.tsc.2020.100761

Clarke, M. (2018). Rethinking graduate employability: the role of capital, individual attributes and context. *Stud. High. Educ.* 43, 1923–1937. doi: 10.1080/03075079.2017.1294152

Cotronei-Baird, V. S. (2020). Academic hindrances in the integration of employability skills development in teaching and assessment practice. *High. Educ.* 79, 203–223. doi: 10.1007/s10734-019-00405-4

Elo, S., and Kyngäs, H. (2008). The qualitative content analysis process. *J. Adv. Nurs.* 62, 107–115. doi: 10.1111/j.1365-2648.2007.04569.x

EQF (2018). *The European Qualifications Framework: Supporting Learning, Work, and Cross-Border Mobility*. Available Online at: http://www.ehea.info/Upload/TPG_A_QF_RO_MK_1_EQF_Brochure.pdf (assessed Feb 28, 2022).

Fadel, C., Bialik, M., and Trilling, B. (2015). *Four-Dimensional Education: The Competencies Learners Need to Succeed*. Boston: Center for Curriculum Redesign.

Falchikov, N., and Goldfinch, J. (2000). Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks. *Rev. Educ. Res.* 70, 287–322. doi: 10.2307/1170785

Forbes (2020). *The Top 10 Skills Recruiters are Looking for in 2021*. Jersey City: Forbes.

Gordon, J., Halasz, G., Krawczyk, M., Leney, T., Michel, A., Pepper, D., et al. (2009). *Key Competences in Europe: Opening Doors for Lifelong Learners Across the School Curriculum and Teacher Education. CASE Network*. Reports No. 87. Warsaw: CASE, Center for Social and Economic Research.

Gordon, T. (2003). *Teachers' Effectiveness Training: The Program Proven to help Teachers Bring out the best in Students of all Ages*. New York, NY: Three Rivers Press.

Kalfa, S., and Taksa, L. (2015). Cultural capital in business higher education: reconsidering the graduate attributes movement and the focus on employability. *Stud. High. Educ.* 40, 580–595. doi: 10.1080/03075079.2013.84221

Kember, D. (2009). Nurturing generic capabilities through a teaching and learning environment which provides practice in their use. *High. Educ.* 57, 37–55. doi: 10.1007/s10734-008-9131-7

Kiuru, N., Aunola, K., Lerkkanen, M.-K., Pakarinen, E., Poskiparta, E., Ahonen, T., et al. (2015). Positive teacher and peer relations combine to predict primary school students' academic skill development. *Dev. Psychol.* 51, 434–446. doi: 10.1037/a0038911

Kiuru, N., Poikkeus, A.-M., Lerkkanen, M.-K., Pakarinen, E., Siekkinen, M., Ahonen, T., et al. (2012). Teacher-perceived supportive classroom climate protects against detrimental impact of reading disability risk on peer rejection. *Learn. Instr.* 22, 331–339. doi: 10.1016/j.learninstruc.2011.12.003

Klemola, U., Heikinaro-Johansson, P., and O'Sullivan, M. (2013). Physical education student teachers' perceptions of applying knowledge and skills about emotional understanding studied in PETE in a one-year teaching practicum. *Phys. Educ. Sport Pedagogy* 18, 28–41. doi: 10.1080/17408989.2011.630999

Kostiainen, E., Ukskoski, T., Ruohotie-Lyhty, M., Kauppinen, M., Kainulainen, J., and Mäkinen, T. (2018). Meaningful learning in teacher education. *Teach. Teach. Educ.* 71, 66–77. doi: 10.1016/j.tate.2017.12.009

Lauritsalo, K., Mäkinen, T., Virtanen, A., Klemola, U., and Tynjälä, P. (2019). "Final year of physical education studies – Supporting physical education students' professional development as a teacher," in *Paper presented at the Conference "International Association for Physical Education in Higher Education" (AIESEP "Association Internationale des Écoles Supérieures d'Éducation Physique")* (New York, NY).

Lavi, R., Tal, M., and Dori, Y. J. (2021). Perceptions of STEM alumni and students on developing 21st century skills through methods of teaching and learning. *Stud. Educ. Eval.* 70:101002. doi: 10.1016/j.stueduc.2021.101002

Mäkinen, T., Kostiainen, E., and Klemola, U. (2022). "Significant in life: Core learning outcomes of a social and emotional course in physical education teacher education," in *International Approaches to Promoting Social and Emotional Learning in Schools: A Framework for Developing Teaching Strategy*, eds M. Talvio and K. Lonka (London: Routledge), 167–189. doi: 10.4324/9781003093053-13

P21 (2020). *Partnership for 21st Century Learning*. Available online at: http://www.p21.org/ (accessed October 20, 2020).

Pakarinen, E., Aunola, K., Kiuru, N., Lerkkanen, M.-K., Poikkeus, A.-M., Siekkinen, M., et al. (2014). The cross-lagged associations between classroom interactions and children's achievement behaviors. *Contemp. Educ. Psychol.* 39, 248–261. doi: 10.1016/j.cedpsych.2014.06.001

Paulhus, D. L., and Vazire, S. (2007). "The self-report method," in *Handbook of Research Methods in Personality Psychology*, eds R. W. Robins, R. C. Fraley, and R. F. Krueger (New York, NY: Guilford Press), 224–239.

Pianta, R. C., La Paro, K. M., and Hamre, B. K. (2008). *The Classroom Assessment Scoring System. Manual*. Baltimore, MD: Brookes.

Pöysä, S., Vasalampi, K., Muotka, J., Lerkkanen, M.-K., Poikkeus, A.-M., and Nurmi, J.-E. (2019). Teacher-student interaction and lower secondary school

student' situational engagement. *Br. J. Educ. Psychol.* 89, 374–392. doi: 10.1111/bjep.12244

Ratmawaty, S. (2018). How to improve classroom atmosphere and undergraduate nutrition students' performance in learning nutrition care process? *Int. J. Learn. Teach. Educ. Res.* 17, 154–174. doi: 10.26803/ijlter.17.11.10

Rekola, M., Nippala, J., Tynjälä, P., and Virtanen, A. (2018). Modelling competences and anticipating the future competence needs in the forest sector. *Silva Fenn.* 52, 1–19. doi: 10.14214/sf.9983

Ryan, R. M., and Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemp. Educ. Psychol.* 25, 54–67. doi: 10.1006/ceps.1999.1020

Sudirman, I., Siswanto, J., and Aisha, A. N. (2020). Software entrepreneurs' competencies based on business growth. *J. Res. Mark. Entrepreneursh.* 22, 111–132. doi: 10.1108/JRME-12-2017-0055

Suyatno, A., Wantini, M., Pambudi, D. I., and Amurdawati, G. (2019). The impact of teacher values, classroom atmosphere, and student-teacher relationship towards student attitude during learning process. *Int. J. Learn. Teach. Educ. Res.* 18, 54–74. doi: 10.26803/ijlter.18.8.4

Tight, M. (2021). Twenty-first century skills: meaning, usage and value. *Eur. J. High. Educ.* 11, 160–174. doi: 10.1080/21568235.2020.1835517

Tuononen, T., Parpala, A., and Lindblom-Ylänne, S. (2019). Graduates' evaluations of usefulness of university education, and early career success – A longitudinal study of the transition to working life. *Assess. Eval. High. Educ.* 44, 581–595. doi: 10.1080/02602938.2018.1524000

Tynjälä, P., and Virtanen, A. (2005). Skill Learning at Work: Investigations into Student Experiences of On-the-Job Learning. *Learn. Skills* 7, 106–116. doi: 10.1097/OPX.0000000000001594

Tynjälä, P., Virtanen, A., Klemola, U., Kostiainen, E., and Rasku-Puttonen, H. (2016). Developing social competence and other generic skills in teacher education: applying the model of integrative pedagogy. *Eur. J. Teach. Educ.* 39, 368–387. doi: 10.1080/02619768.2016.1171314

Vila, L. E., Perez, P. J., and Morillas, F. G. (2012). Higher education and the development of competencies for innovation in the workplace. *Manag. Decis.* 50, 1634–1648. doi: 10.1108/00251741211266723

Virtanen, A. (2013). *Opiskelijoiden oppiminen ammatillisen peruskoulutuksen työssäoppimisen järjestelmässä [Students' workplace learning in Finnish Vocational Education and Training]*. Ph.D. thesis. Jyväskylä: University of Jyväskylä.

Virtanen, A., and Tynjälä, P. (2019). Factors explaining the learning of generic skills: a study of university students' experiences. *Teach. High. Educ.* 24, 880–894. doi: 10.1080/13562517.2018.1515195

Virtanen, A., and Tynjälä, P. (2022). Pedagogical practices predicting perceived learning of social skills among university students. *Int. J. Educ. Res.* 111:101895. doi: 10.1016/j.ijer.2021.101895

Virtanen, A., Tynjälä, P., and Eteläpelto, A. (2014). Factors promoting vocational students' learning at work: study on student experiences. *J. Educ. Work* 27, 43–70. doi: 10.1080/13639080.2012.718748

Visiers-Jiménez, L., Suikkala, A., Salminen, L., Leino-Kilpi, H., Löyttyniemi, E., Henriques, M. A., et al. (2021). Clinical learning environment and graduating nursing students' competence: a multi-country cross-sectional study. *Nurs. Health Sci.* 23, 398–410. doi: 10.1111/nhs.12819

Wang, L., MacCann, C., Zhuang, X., Liu, O. L., and Roberts, R. D. (2009). Assessing teamwork and collaboration in high school students: a multimethod approach. *Can. J. Sch. Psychol.* 24, 108–124. doi: 10.1177/0829573509335470

Zeng, L. M. (2021). How much does it differ? How much does it matter? The research experience of Mainland Chinese and Hong Kong students in a Hong Kong University. *Stud. High. Educ.* 46, 606–623. doi: 10.1080/03075079.2019.1647412

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 28 July 2022
DOI 10.3389/feduc.2022.886726

OPEN ACCESS

EDITED BY
Edith Braun,
Justus-Liebig Universität, Germany

REVIEWED BY
Lars Müller,
University of Giessen, Germany
Ulrike Schwabe,
German Centre for Higher Education Research and Science Studies (DZHW), Germany

*CORRESPONDENCE
Hanni Muukkonen
hanni.muukkonen@oulu.fi

SPECIALTY SECTION
This article was submitted to Higher Education, a section of the journal Frontiers in Education

RECEIVED 28 February 2022
ACCEPTED 05 July 2022
PUBLISHED 28 July 2022

CITATION
Muukkonen H, Lakkala M, Ilomäki L and Toom A (2022) Juxtaposing generic skills development in collaborative knowledge work competences and related pedagogical practices in higher education. *Front. Educ.* 7:886726. doi: 10.3389/feduc.2022.886726

Juxtaposing generic skills development in collaborative knowledge work competences and related pedagogical practices in higher education

Hanni Muukkonen [1,2]*, Minna Lakkala [2], Liisa Ilomäki [2] and Auli Toom [3]

[1]Learning and Learning Processes Research Unit, Faculty of Education, University of Oulu, Oulu, Finland, [2]Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland, [3]Centre for University Teaching and Learning, Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

This study employs the term knowledge work competence to address generic aspects of higher education graduates' expected learning outcomes. Twenty-eight higher education courses were investigated: 1069 students responded to the Collaborative Knowledge Practices (CKP) questionnaire to rate their self-evaluated competence development. From the same courses, 56 teachers provided descriptions of the course pedagogical practices. First, students' self-reported generic collaboration competence gains were analyzed statistically for differences between courses. Second, qualitative categorization of the pedagogical practices based on rich description of pedagogical designs and teachers' reflective responses was carried out. This offered a categorization with elaborated descriptions and a clustering to three types of enacted pedagogical practices. Finally, the study juxtaposed these previous two results to investigate how the pedagogical features were related to students' self-evaluations on collaboration competence gains. The findings highlighted one cluster of pedagogical practices, collaborative knowledge creation with systematic support for epistemic and group work, as most beneficial for student competence gains. In it, professional ways of working were explicitly modeled and practiced, teacher support for knowledge creation during contact teaching was available, and time was reserved for reflection with students. Such pedagogical practices are important to ensure graduates' fluent transition to complex knowledge work.

KEYWORDS

generic competence development, pedagogical practice, knowledge creation, collaboration, multi-method analysis

Conceptual background and research questions

Higher education is expected to prepare future academic experts for the knowledge-driven global world (Barrie, 2012; Karlgren et al., 2020a). Successful learning and working in today's knowledge-based society demands competence that exceeds individual expertise and engages individuals in joint collaboration and knowledge creation in teams (Binkley et al., 2012). Such competence is embedded in people's actions, social interaction and the socio-material affordances of their environments as they co-develop knowledge objects (e.g., Damşa and Muukkonen, 2020), engage in epistemic practices (Markauskaite and Goodyear, 2017) and regulate their collaborative learning and working as a team (e.g., Borge et al., 2018; Splichal et al., 2018). To understand how such competence develops in higher education, we need to examine both the experienced competence gains and pedagogical settings in which these competences are nurtured. Although many different factors influence learning, there is a need to further decipher the role pedagogical practices may have on competence development.

Many countries are struggling to keep up with the demands of a highly skilled workforce (OECD, 2018), and the current COVID-19 pandemic has created sudden changes and challenges as team members are forced to work remotely and devise novel practices for collaboration. Professional teamwork has taken a major shift from disciplinary to interdisciplinary teams to respond to the growing complexity and dynamic nature of tasks and to seek better ways to tackle ambiguous challenges (Benoliel and Somech, 2015). However, educational objectives and practices may not be truly aligned with the changes in professional work (Markauskaite, 2020). Particularly in interdisciplinary collaboration, both discipline specific and generic (e.g., critical and analytical thinking, problem-solving, self-management of learning, communication skills, and information and digital literacy, Binkley et al., 2012) competences are needed in an intertwined manner to produce novel ideas, syntheses, designs or practices (Goodyear and Zenios, 2007). How students are directed to engage in learning activities is instrumental for competence development (Goodyear and Zenios, 2007; Puntambekar et al., 2007).

A meso-level investigation of competence development originates theoretically from sociocultural paradigm: Learning is regarded as embedded in social processes, practices, and tool use rather than being an individual venture (Säljö, 2010). This study employs the term knowledge work competence to address the generic aspects of higher education learning. Knowledge work competence (Damşa and Muukkonen, 2020; Karlgren et al., 2020b; Muukkonen et al., 2020) for higher education graduates refers to capacities for advanced knowledge work, i.e., understanding and creating knowledge, orchestrating collaboration, and self and co-regulating performance. As such, defining "work" for higher education graduates and exploring the relationship between work and learning are complex tasks considering interdisciplinarity and dynamics of external environments across professional fields (Jung, 2022). However, graduates from higher education need to be equipped with competences to solve complex authentic problems regardless of field, take part in creating knowledge in real working life settings and promote novel solutions by using the community's collective, technology-mediated efforts.

This paper carries out a multi-method investigation. First, on a dataset of self-reported student assessments of own generic competence development, more specifically in collaborative knowledge work competence. The self-reports were collected following a specific higher education study unit, referred to as a course. This offers the student perspective on which types of competence development was central in the examined course. Second, the study offers a framework for analyzing pedagogical features in the same courses. The pedagogical features were mapped through a survey to the same courses' teachers, followed by a categorization of the pedagogical practices based on rich description of pedagogical designs and reflection responses by the teachers. This offers a categorization and elaborated description of the pedagogical practices and their clustering to three types of enacted practices. Third, as the core result, the study juxtaposes these previous two results to investigate how certain types of pedagogical practices may contribute to generic competence development, particularly collaborative knowledge work competences. The courses represented authentic higher education instructional practices: organized as lecture, project, inquiry, and interdisciplinary courses, which all involved some type of collaboration between peers. Hence it was meaningful to examine the variation of competence development in relation to the pedagogical practices.

Previous studies have investigated knowledge work competence by structuring it 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 (e.g., Karlgren et al., 2020b; Muukkonen et al., 2020; Vesikivi et al., 2020). This builds on the theoretical background of socio-cultural theories of learning and particularly on the knowledge creation metaphor (Paavola and Hakkarainen, 2005). Acquisition and participation as two metaphors of learning were put forward by Sfard (1998). The acquisition metaphor of learning addresses assimilation of knowledge and the individual's mental models and strategies of learning. The participation metaphor refers to adaptation to the existing cultural and communal practices and the dialogical practices of learning. As a third metaphor, Paavola and Hakkarainen (2005) added the knowledge creation metaphor. It introduced the presence of artifacts, products and practices i.e., objects,

and collaboration to advance them as pivotal (Paavola and Hakkarainen, 2005). The objects can for instance be a report or essay co-authored together, or a procedure description, website, or plan co-created in collaboration. Briefly expressed, object-orientedness is a concept formulated in the cultural-historical theory, referring to a shared motive or tangible object for a learning or working community (Miettinen and Virkkunen, 2005). The object mediates knowledge advancement as the participants collaborate to negotiate meanings, extend, and version it. For students, educational activities emphasizing knowledge creation metaphor often involve more open-ended and complex assignments which integrate collective efforts around iterative development of knowledge objects.

Pedagogical design includes many aspects. "Design involves making invitations to other people to act in certain kinds of ways. These invitations can be clear and explicit, but they are sometimes encoded into the affordances of materials. Designers' knowledge has to include ways of predicting, or at least imagining, how other people will respond to these invitations" (Goodyear, 2015, p. 39). Making these designs involves pre-active aspects of planning but also post-active phases of reflection, evaluation, and assessment. Resulting course documents (e.g., course plans, instructions, assignment descriptions) and digital tool choices can be considered material instantiations of the teachers' ideas and decisions regarding the organization of tasks, activities, and responsibilities (Goodyear, 2015; Esterhazy et al., 2021).

Next, the introduction of previous research will review object-orientedness, role of integration of efforts in collaborative learning, feedback, cross-fertilization, and digitalization in higher education. Competence development and pedagogical practices are addresses as an intertwined phenomenon, which is further elaborated in the empirical investigation.

Object-orientedness

Evidence in higher education is building on how students engage in meaningful interactions with peers, knowledge resources and objects, and the social and digital-material environment in which such activities take place (Damşa and Muukkonen, 2020). Learning addressed as a process of knowledge creation brings it closer to professional practices, which takes place through interactive practices that contribute to ideas being materialized into (shared) knowledge objects (Paavola et al., 2011). In higher education, such objects may be for instance reports, designs or products ideated and co-created in student collaboration. In a study comparing two anatomical sciences courses with different pedagogical designs, the students reported more competence gains when they had a shared object to prepare, the teaching presentation, compared to just taking part in interaction with peers on an assignment (Laakkonen and Muukkonen, 2019). The shared knowledge object intensified the need to collaborate and learn about planning, coordinating, and sharing responsibility during collaboration as well as the integration of individual and collaborative contributions.

From the pedagogy point of view, the objects are elaborated through intermediate and mediating artifacts and tools, and iterative development of tangible artifacts, such as draft and sketches (Miettinen and Paavola, 2018; Damşa and Muukkonen, 2020). This requires the teacher to make specific choices about the intensity and extend of collaboration, how collaboration is assessed, to plan a process involving iterative cycles of feedback, editing and monitoring the epistemic challenge.

Integration of efforts in collaborative learning

Research emphasizes that engaging in productive co-construction of knowledge does not happen automatically (Baker et al., 2013). Individuals and groups vary in the extent of their competence to collaborate with others and to respond to the situation-specific learning and interaction challenges in authentic educational settings (Näykki et al., 2014). During collaboration, students are expected to negotiate task aims and standards, to act strategically based on monitoring their group activities, to revise processes and outcomes, to select and use suitable digital tools, and to productively deal with any challenges groups face (e.g., Splichal et al., 2018). Through extended practice, successful learners and team members use a repertoire of skills and strategies to regulate their learning processes on cognitive, social, and emotional levels (Baker et al., 2013; Hadwin et al., 2017).

Regarding pedagogy, Vogel et al. (2017) meta-analysis found that computer-supported collaborative learning scripts were particularly effective for domain-specific learning when they prompted transactive activities in which learners build on the contributions of peers and when additional content-specific scaffolding such as worked examples were available. The present study was motivated by the need to better understand content-generic aspects of learning, and, further, aims to examine at an elaborate detail the design of collaboration with peers and scaffolding for generic competence development and its impact on student collaboration competence learning.

Feedback

Student-centered methods in higher education emphasize students' central role in regulating their own learning. This involves generating and soliciting feedback on their own learning (Boud and Malloy, 2013). Further, orchestration of collaboration extends the competence demands to proactive feedback on both individual and collaborative learning.

Making such a role feasible presumes that teachers need to plan productive feedback opportunities in which students can engage actively with and employ the feedback for future learning. These include dialogical processes and activities which can support and inform the student on the task at hand, while catering for the ability to self-regulate performance on future tasks (Carless et al., 2011; Esterhazy et al., 2021). Similarly, regarding engagement on the shared object, collective feedback is important for co-development and competences for advancing collective outcomes. Further, Esterhazy et al. (2021) showed that productive feedback should not be understood as a prescribed model or solution across all disciplines but contextualized in disciplinary or interdisciplinary objectives and pedagogy.

Cross-fertilization

Modern teamwork is often organized as collaboration in online communities, with heterogeneous and temporary convergence (Faraj et al., 2011). Students entering work life should be ready to act as agentic collaborators who can participate proactively in solving interdisciplinary and ambiguous challenges. Therefore, cultivating competence for working in interdisciplinary teams and creating joint knowledge objects are increasingly considered important objectives in higher education (Cooke and Hilton, 2015).

In pedagogical practice, cross-fertilization refers to interaction between different areas of expertise or organizations, for instance, by collaboration, problem solving or new product development for purposes extending beyond educational institutions (Paavola et al., 2011). Similarly, Cremers et al. (2016) used the term 'hybrid learning configurations' to define designs which connect formal learning with workplace experiences by integrating settings for studying and working. Interdisciplinary co-creational activities with ill-defined and authentic tasks are central in such configurations. Project courses are commonly used methods by involving various stakeholders inside and outside of educational institutions. Projects engage students in producing tangible and meaningful results, sometimes in cooperation with professionals, generating, potentially, outcomes for continued use in an organization (e.g., Viswanathan et al., 2012).

Digitalization in higher education

Two important assumptions are in need of consideration in the digitalization of higher education: First, the assumption that technology is an instrumental issue that is neutrally implemented and second, that students became fluent users of technology in a self-directed way (Castañeda and Selwyn, 2018). Both of these assumptions are tightly tied to generic competences. The way that digital tools are integrated to pedagogical design can have a considerable influence on the kinds of practices that can be designed for and enacted in collaboration. Also, students do not necessarily have the required competence to engage in technologically mediated knowledge work, collaboration, or expert-like practices of writing and co-creation, without instructed and guided engagement.

Pedagogy is inherently part of any educational technology use in higher education (Castañeda and Selwyn, 2018) although this is often reduced to learning management systems serving very basic information distribution and communication needs. Theory and pedagogy informed technology design has had considerable efforts invested through research and development, but the mainstream technology use remains designed for the support of logistical processes rather than for pedagogical change (e.g., Collis and Moonen, 2008).

Research questions

The study investigated how higher education students in twenty-eight courses evaluated their learning and competence gains in the generic competences of collaborative knowledge work practices. Further, the study examined the courses' pedagogical practices to provide a combined, juxtaposed, understanding of how the pedagogical practices were related to student learning. The following research question were examined:

1. Were there differences in students' self-assessed competence gains between courses?
2. What kinds of design of collaboration did the courses' pedagogical design reflect?
3. How the pedagogical features were related to students' self-evaluations on collaboration competence gains?

Materials and methods

The general investigative approach was an explanatory multiple case study (Yin, 2014) and a multi-methods approach was used in the data collection and analysis (De Laat et al., 2007; Cresswell, 2009). The aim of the approach was to gain a triangulated understanding of course pedagogical design, enacted practices and student learning. Teachers were invited to take part in the study, by answering a questionnaire on course design and reflection responses and by passing forward a link to an e-questionnaire to the students and encouraging their participation. Students were asked to answer the questionnaire at the end of their course. All participants were asked to provide their informed consent electronically, and those responses without a consent were excluded from the study.

Courses and participants

The data included responses from twenty-eight courses in two large Finnish universities. The data was not intended to be representative of specific fields, rather, the aim has been to involve multiple fields and types of collaboration to investigate the variation. The fields of study included education, educational psychology, philosophy, life sciences, law, and economics. The courses, participants, and response rates are detailed in Table 1. Courses typically lasted for one period of c. 8-9 weeks and were obligatory courses in the degree program. Courses were included which received more than 7 responses per course from students and a teacher response was available. The response rate to the CKP questionnaire for students was 55.9% (varied between 25-58%), as a total of 1,912 students completed these investigated courses. In total 56 teachers were included in the data, with mean age 48.5 years and 55% female. Some courses had multiple teachers, especially project-type courses or larger courses and thus we obtained several teacher responses per course. In total 1,069 student responses were included for analysis 19.8% male, 79.3% female and 0.8% reported other or missing. 755 students were enrolled in a first-year course, other courses were in later bachelor or master's degree studies. Students' average age was 24.9 (SD = 6.6) reflecting the rather high university starting age in Finland.

Data collection

The first data consisted of higher education students' responses to the Collaborative Knowledge Practices questionnaire (CKP; Muukkonen et al., 2020). The CKP

TABLE 1 Participants.

Course field	Course ID	ECTS	Teacher responses	Student responses	Student age		Student gender		Total students completing the course
			n	N	*M*	*SD*	Male	Female	n
Plant sciences	ID02	3	1	49	24.4	6.1	15	34	60
Economics	ID03	5	1	21	24.7	6.8	5	16	79
Environmental change & economics	ID04	5	2	16	26.8	7.8	4	12	33
Economics	ID60	5	1	53	22.6	3	14	39	79
Veterinary medicine	ID61	4	2	34	22.8	6.6	3	31	66
Philosophy	ID63	5	1	25	33.3	9.5	2	23	40
Agricultural sciences	ID64	3	2	26	26.8	7.1	12	14	29
Agricultural sciences	ID65	5	3	8	30	6.6	1	7	8
Aquatic sciences	ID66	5	3	11	25.3	2.7	1	9	14
Agricultural sciences	ID67	5	3	17	26.4	6.6	5	12	60
Veterinary medicine	ID68	3	1	20	24.7	4.4	3	17	69
Veterinary medicine	ID81	4	1	16	21.9	3.1	0	16	70
Philosophy	ID82	5	2	25	28	8.5	2	23	51
Educational psychology	ID83	5	1	7	28	5.4	2	5	25
Agricultural sciences	ID84	3	3	17	23.1	3.1	3	14	26
Educational sciences	ID85	5	2	71	25	7.3	9	61	70
Agricultural sciences	ID87	5	6	13	26.3	5.4	1	12	15
Agricultural sciences	ID88	5	3	21	24.6	8.7	3	18	71
Educational psychology	ID89	5	1	17	32.3	8.3	2	15	31
Educational sciences	ID91	5	5	263	24.9	6.7	32	228	375
Philosophy	ID92	5	2	20	30.6	9.6	0	20	48
Law	ID94	5	3	133	22.8	4.3	45	87	178
Forest Sciences	ID95	5	1	40	24.1	4.6	15	25	38
Agricultural sciences	ID96	3	2	34	26.1	7	6	28	41
Agricultural sciences	ID97	5	3	9	31.6	12.3	4	5	10
Agricultural sciences	ID110	5	3	47	23	5.6	15	29	100
Educational psychology	ID111	5	1	10	25.5	4.8	1	9	27
Educational psychology	ID113	3	1	46	23.5	4.5	7	39	199
Total			60	1069	24.9	6.6	212	848	1912

questionnaire has been developed and validated for use as a generic self-evaluation tool for students on course-based learning outcomes on generic collaborative knowledge work competences (Karlgren et al., 2020b; Muukkonen et al., 2020). The CKP does not measure content learning, it thus complements other content-related evaluation measures employed in a given course. The scales of CKP were used to measure course-related learning. The seven scales are: Collaborate on shared objects, Integrating individual and collaborative working, Development through feedback, Persistent development of knowledge-objects, Understanding various disciplines, Interdisciplinary collaboration, and Exploit digital technology. Students were asked to evaluate how each statement (27) corresponded to their competence learning on the seven scales of the CKP. "During the course I have learned . . .," e.g., "to develop ideas further together with others," "to understand the value of commenting on work in progress," and "to use various digital applications and use them together whenever needed" (please see for details Muukkonen et al., 2020; Karlgren et al., 2020b). The statements were on a five-point Likert-scale (1 = not at all – 5 = very much).

The second data consisted of teachers' open answers (N = 56) to an online questionnaire about the practices (collaboration design, types of tasks, use of digital tools, guidance, assessment) and reflection of experiences in 28 courses. Additional materials collected from the investigated courses, including course descriptions, task guidelines, teaching materials, digital platform content, lesson observations, or students' feedback forms, were used as complementary data in the analysis of pedagogical practices.

Data analysis

Collaborative knowledge practices questionnaire

Student responses were screened for outliers and seventeen participants were removed from data due to missing data or unvarying responses. In the first two courses, the data was collected with an option "not applicable" (0). We replaced the 'not applicable' responses with "not at all" in the scale in order to aid interpretation of the data (please see Muukkonen et al., 2020 for details). The seven scales of the CKP questionnaire were used to examine course related self-reported learning. The scale reliabilities were good (Cronbach's alpha 0.73-0.86).

Pedagogical practices of the courses

The pedagogical practices analysis aimed to examine the design features based on identified pertinent pedagogical design elements for collaborative practices. The variation within the pedagogical practices was outlined by describing in detail the pedagogical practices for every course. Initial categories were based on related theories such as knowledge building (Bereiter, 2002), metaphors of learning (Paavola et al., 2011), and authentic learning (McCune, 2009) as well as empirical studies (e.g., Lakkala et al., 2015, 2020; Ilomäki et al., 2017. The unit of analysis was the whole course, and the analysis covered, first, the teachers' questionnaire answers about the course practices and, second, all other data available from the courses. The categories were created through abductive use of theory-informed and data-grounded data analysis (Timmermans and Tavory, 2012).

The pedagogical design of each course was described in a table under the initial categories. The different ways to design collaboration, use scaffolds and modeling for collaboration were documented. Based on these descriptions, categories and subcategories were further elaborated. For instance, process-like emphasis has been raised in prior literature as an important quality of collaborative learning and knowledge building (Muukkonen and Lakkala, 2009; Scardamalia and Bereiter, 2014) to improve student produced contributions. Three subcategories were identified related to this: Iterativeness describing the longitudinal versioning and improvement of outcomes; availability and designed points of Feedback to support the process; and Reflection of practices explicitly included and modeled as part of the learning process. These sub-categories were named as features of pedagogical practices and positioned in three exclusive levels. Level 1 not involving the described features, level 2 to some extend and level 3 to wide extend. Four researchers created the categories and sub-categories together in several joint analysis sessions and made a preliminary analysis with a sub-set of 16 courses. After the preliminary analysis, one researcher made the analysis of all courses, after which the analysis results were, again, discussed together between the researchers in several sessions, clarifying unclear definitions and making decisions about the final categorization. The discussions were carried out until there was an agreement between two coders for the entire data and four coders for a c. 50% of the units of analysis as it was discussed during the development process. Category and subcategory descriptions were written. Finally, each course was scored with all sub-categories using levels 1-3 to explicate the extent and nature of the pedagogical features in the course practices (see **Table 2**). All the highest levels (score 3) aim to describe a pedagogical practice where the targeted competences are modeled and supported by various design decisions implemented in the course.

The main categories are the following (see **Figure 1**):

- Object-orientedness refers to the degree that the course collaboration is organized around shared knowledge objects, such as a report, website, design or a product. The extensiveness of the shared object influences its role in the collaboration. How the developed shared object is planned to be reused by students or other stakeholders may add re-use value to the knowledge object.
- Epistemic challenge is outlined by the wider or more narrow problem space where the student-centered activities are embedded in. Explicit modeling of professional

TABLE 2 Categorization, level descriptions and examples from the analysis of pedagogical practices.

Main category	Subcategory	Level description	Example from qualitative analysis description
Object-orientedness	Extent of tangible outcomes to develop	(1): No artifact development, only oral discussions or answers to teacher-defined questions	Hands-on activities for learning of an animal body and oral conceptualization in discussions and negotiations (ID61)
		(2): Several minor artifact production tasks	Two group reports for answering teacher-created questions; small group tasks during lectures; e.g., filling a worksheet (ID60)
		(3): One-two major artifact production tasks (in addition to smaller ones)	One large project work including a project plan, presentation and report for a real client's challenge (ID65)
	Reuse of knowledge artifacts	(1): No re-use	No explicit reuse, the reports were for course completion (ID111)
		(2): Some artifacts reusable by the students in the course or afterward	Students made essays individually from a self-chosen topic. The goal was that the essays can be used in their future studies (e.g., in master thesis) (ID82)
		(3): Planned re-use across courses or by external stakeholders	A solution made for the client to be used later; solutions from previous courses as examples (ID04)
Epistemic challenge	Problem space	(1): Narrow, well-defined tasks	Individual homework tasks (e.g. calculations), narrow essay-type tasks, well-defined and same for all (ID88)
		(2): Limited problem space or pre-defined task structure (comparison, analysis, review)	Weekly applied group tasks (e.g., analyzing law cases); same tasks for all groups (ID94)
		(3): Open, ill-defined and challenging problems or authentic task challenge	A solution applicable in an authentic context from a topic chosen by the group and developed throughout the course (ID95)
	Student-centered activities	(1): Meetings mainly based on lecturing	Course meetings were mass lectures, pair essays were done outside meetings (ID116)
		(2): Meetings include both lecturing and students' own working	Lectures, group work and visits to organizations (ID02)
		(3): Practically no lecturing in meetings, mainly students' own working	Course meetings mainly included project work in teams (ID65)
	Modeling of professional practices	(1): Content learning practices	The course was organized as a flipped teaching design for content learning (ID03)
		(2): Simulating professional practices but only in some tasks or inadequately implemented (e.g., very short time, no explicit phasing or modeling)	Writing a Wikipedia article and scientific term bank definitions in addition to other tasks (presentations, learning log) (ID02)
		(3): Simulating professional and real-life working processes with explicit modeling.	The progressive inquiry model used to simulate professional research practices (ID83)
Process-like emphasis	Iterativeness	(1): Narrow tasks without versioning or elaboration	No iterations in producing weekly case analyses (ID94)
		(2): A knowledge creation process without clear iteration points (or only one)	Solutions produced longitudinally, but only random guidance from the teacher in course meetings; no explicit iteration phases
		(3): Longitudinal process with several versioning and iteration phases	Individual essays were elaborated, commented on in thematic groups and improved through several iterations (ID92)
	Role of feedback	(1): No feedback or only joint discussions	No feedback, only if the essay was not accepted, a possibility to improve it (ID116)
		(2): Random oral feedback from the teacher or discussions with peers during or after the process	Peer evaluation of reports between groups before the final submission; discussion of solutions in lessons (ID03)
		(3): Explicit feedback from peers, teachers or external stakeholders at several points	Repeatedly given comments from clients and lecturers affected the progress of project work and finalization of plans and reports (ID87)
	Reflection of practices	(1): No	No organized reflection (ID81)
		(2): Oral reflection discussions or reflection only at the end	Students evaluated their contribution in group work at the end (ID03)
		(3): Explicit reflection activities during the process with models and templates provided	The groups evaluated their practices through templates at the middle of the course; the group and course practices were also discussed in the last meeting (ID64)
Intensity of collaboration	Centrality of collaboration	(1): Tasks mainly individual, or small-scale group activities in the meetings	No collaboration instead of small group activities during lectures (ID110)
		(2): A mixture of individual and group tasks	Essays written individually, peer feedback in organized thematic groups (ID82)
		(3): Main course tasks based on group work	Project work made in groups throughout the course (ID66)
	Integration of individual and collaborative tasks	(1): No collaboration or separate individual and group tasks	Home exams and weekly group tasks separated; group work based on groups' independent regulation (ID67)

(Continued)

TABLE 2 (Continued)

Main category	Subcategory	Level description	Example from qualitative analysis description
		(2): Individual contribution to group tasks relevant and expected, but not well structured, guided or followed	Contribution in weekly group tasks expected but not followed. Absence from group presentation had to be replaced by a separate individual reflection task (ID94)
		(3): Individual and group tasks highly integrated and systematically structured, guided or followed	Groups were formed based on students' interests; project work completed in groups and division of labor had to be decided; each student kept a log of their working hours (ID97)
Cross-fertilization	Multidisciplinarity	(1): No multidisciplinarity	No, all students were from the same major (ID61)
		(2): Integration of a couple of fields or sub-fields and majors	Student were from different sub-majors working in mixed groups (ID95)
		(3): Fully multidisciplinary demonstrated in content and/or group compositions	The course was organized between two faculties and disciplines, groups were formed based on the diversity of backgrounds (ID04)
	External collaborators	(1): No external collaborators	No external collaborators, but three university lecturers as experts (ID61)
		(2): Some contacts, visits or interaction with external stakeholders	Communication with an educational researcher from the faculty: meetings, discussions and reading the researchers' articles (ID85)
		(3): Intensive or multiple type of collaboration with external stakeholders (experts, professionals)	Collaboration with an external client in group projects (ID66)
Digitalization	Use of digital tools	(1): A course platform, the Web and basic office applications in use	Moodle platform for course organization and peer commenting, Word for individual essays (ID82)
		(2): Also other apps, cloud services or professional tools in use; freedom to choose apps to be used in group work	Moodle platform for course organization, Wiki for sharing materials and project work activities, and tools chosen by the groups (ID66)
		(3): Versatile and integrated use of various types of applications for different purposes; joint agreements and models for digital practices in groups	Moodle platform for course organization, co-authoring tools with templates for group activities (e.g., OneDrive documents), getting familiar with various cloud services, writing a blog post in groups (ID84)
Assessment foci	Versatility of assessment methods	(1): No assessment or holistic assessment made by the lecturers	Pass/fail grading by the teacher based on group work (ID68)
		(2): Grading made by the lecturers based on a combination of tasks	Assessment by the teacher based on multiple tasks (tasks in Moodle, essays, peer-commenting) (ID84)
		(3): Versatile assessment methods and assessors (individual, group, mixed; grades or pass/fail; teacher, peers, experts)	Grading of the group reports by the teacher; group self-assessment (with an evaluation matrix) had an effect on the final grading (ID111)
	Assessment of generic competences	(1): No assessment or focusing on content acquisition	Assessment focusing on acquiring the course content (ID03)
		(2): Learning of generic competencies included in learning objectives and tasks, but not explicitly graded	Real project work as the object of learning but skills not separately assessed (ID04)
		(3): The learning of generic competencies (e.g., group working, ways of commenting, argumentation) explicitly assessed	Project work progression and working in groups assessed in addition to the quality of outcomes (ID97)

practices is a way to support students in tackling complex epistemic challenges.

- Process-like emphasis involves a design that includes iterations in the process, allowing for revising and improving outcomes. Offering and receiving feedback during the process is a central for improvement as well as the collective and individual reflection of practices.
- Intensity of collaboration is enhanced by design decisions that emphasize the centrality of collaboration and employ tasks which require the integration of individual and collaborative efforts.
- Cross-fertilization refers to involving multiple disciplinary expertise and professional practices in the course activities or participants. External collaborators may also take varying roles in the course, e.g., by giving assignments or being clients.
- Digitality describes the use level of digital tools in a course, e.g., for collaboration, communication, disciplinary activities, or teaching.
- Assessment foci highlights how the pedagogical design incorporates versatile assessment methods to support competence development and, especially, considers also more generic types of competences, such as knowledge work competences.

A K-means Hierarchical cluster analysis was conducted using the scores of sub-categories given to each course for grouping the courses. One-way ANOVA and

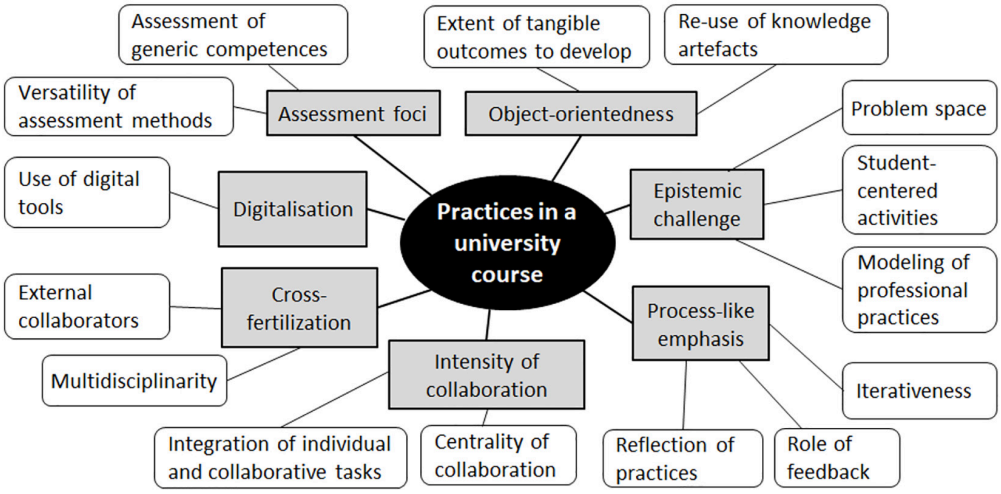


FIGURE 1
Overview of categories of pedagogical practices.

Post hoc (Tamphane's T2) analysis were applied for comparing cluster means.

Finally, an ANOVA analysis of variance was carried out to examine how students' self-evaluations about learning of knowledge work competence were related to the three clusters of pedagogical practices. For this comparison, we used the clusters found for grouping the courses (RQ2) and examined how students' self-reported competence learning scale means and standard deviations (RQ1) were distributed across these clusters.

Results

Students' self-assessed competence gains across courses

To examine whether there were differences between students' self-reported competence learning between courses, scale means and standard deviations were calculated for each course (Table 3). The evaluations showed statistically significant variation between courses (Table 4). Scales related to Understanding various disciplines and Interdisciplinary collaboration were on average scored the lowest. Scales related to learning to Collaborate on objects and Integrate efforts were scored on average the highest.

Pedagogical features related to the design of collaboration and student activities

The analysis of pedagogical practices through cluster analysis (Table 5) uncovered three prototypical types of enacted practices (a) activating learning promoting content learning, (b) self-directed individual or collaborative knowledge creation, and (c) challenging collaborative knowledge creation with systematic support (Figure 2).

Courses in Cluster 1 can be illustrated to represent *activating learning practices promoting content learning*. Courses primarily focused on the acquisition of domain content through activating lectures and/or small-scale individual and collaborative knowledge creation tasks both in contact sessions and as assignments. It is noteworthy, that all ten courses in Cluster 1 were for first year students and there were, on average, 96.2 students in a course (min 60, max 178). The courses were from various subject domains: economics, biosciences, law, agricultural sciences and education. It appeared that the pedagogical practices were designed for large class activities, with primary emphasis on engaging students with activating methods on content learning.

As an example of Cluster 1, one course (ID94) was a 5 ECTS course about legal thinking for the first-year law students including about 180 participants. Students had weekly group tasks to analyze law cases and write up one own law. Groups organized their group work time themselves outside contact sessions. In addition to attending expert lectures, students presented their group outcomes and/or were opponents to some other group in contact sessions. The tasks were applied tasks, and although the tasks were the same for all groups, the problem space was open. Course platform Moodle was used for sharing materials and group discussions. In addition, lecturers demonstrated professional databases for students. Groups had a freedom to use digital applications of their choice in presentations. The course was graded on a scale pass-fail

TABLE 3 Collaborative knowledge practices scale means and standard deviations.

Course field	N	Collaborate on objects		Integrate efforts		Feedback		Persistent development		Various disciplines		Interdisciplinary collaboration		Exploit technology	
		M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Plant sciences	49	3.84	0.56	3.83	0.60	3.44	0.70	3.48	0.68	3.58	0.74	2.78	0.96	3.27	0.78
Economics	21	3.75	0.68	3.67	0.87	3.46	0.88	3.57	0.79	3.08	0.87	3.08	0.90	3.87	0.90
Environmental change & economics	16	4.45	0.73	4.34	0.58	4.27	0.74	4.27	0.58	4.05	0.74	3.92	0.67	3.42	0.85
Economics	53	3.43	0.61	3.56	0.68	3.15	0.67	3.34	0.63	2.41	0.69	2.49	0.65	3.27	0.77
Veterinary medicine	34	3.16	0.65	3.77	0.60	2.93	0.80	3.25	0.75	2.25	0.75	1.91	0.78	2.84	0.78
Philosophy	25	2.24	0.72	2.81	0.77	3.49	0.62	2.63	0.73	1.83	0.62	1.84	0.75	1.84	0.74
Agricultural sciences	26	3.37	0.52	3.34	0.56	2.95	0.57	3.19	0.49	2.88	0.58	3.09	0.72	3.40	0.68
Agricultural sciences	8	4.03	0.63	4.06	0.48	3.75	0.58	4.09	0.40	3.75	0.38	3.25	0.53	3.75	0.48
Aquatic sciences	11	3.64	0.53	3.70	0.81	2.98	0.70	3.64	0.64	2.66	0.64	2.70	0.60	2.25	0.82
Agricultural sciences	17	3.20	1.05	2.93	0.98	2.68	0.76	2.90	0.78	2.35	0.78	2.14	0.93	2.81	0.86
Veterinary medicine	20	3.89	0.63	4.15	0.61	3.00	0.75	3.44	0.58	2.33	0.57	1.85	0.79	2.48	0.72
Veterinary medicine	16	3.61	0.94	4.15	0.83	3.36	1.08	3.44	1.08	2.53	0.76	2.00	0.66	2.17	0.86
Philosophy	25	2.60	0.76	2.98	0.76	3.55	0.72	2.69	0.69	1.99	0.82	1.82	0.71	2.35	0.83
Educational psychology	7	3.89	0.63	3.71	0.53	3.46	0.77	3.61	0.24	2.64	0.67	2.38	1.04	2.71	0.98
Agricultural sciences	17	3.51	0.60	3.51	0.42	3.15	0.75	3.24	0.35	2.74	0.35	2.61	0.64	3.09	0.57
Educational sciences	71	3.84	0.59	3.95	0.59	3.21	0.66	3.46	0.55	2.85	0.62	2.86	0.81	3.13	0.81
Agricultural sciences	13	4.17	0.66	4.06	0.61	3.87	0.50	4.04	0.60	3.33	0.84	3.10	0.81	2.98	0.75
Agricultural sciences	21	3.43	0.64	3.44	0.62	2.64	0.72	3.00	0.59	2.61	0.74	2.13	0.74	2.90	0.85
Educational psychology	17	4.01	0.58	3.96	0.66	3.91	0.73	3.78	0.54	3.01	0.77	2.67	1.03	2.97	1.03
Educational sciences	263	381	0.62	3.87	0.61	3.26	0.61	3.43	0.55	2.99	0.68	2.83	0.76	3.21	0.82
Philosophy	20	2.64	0.74	3.09	0.61	3.64	0.56	2.91	0.55	2.00	0.83	2.02	0.95	2.44	0.73
Law	132	3.41	0.66	3.35	0.72	2.87	0.69	3.25	0.64	2.56	0.67	2.11	0.77	3.05	0.78
Forest Sciences	40	3.17	0.64	3.01	0.69	2.83	0.74	2.96	0.69	2.31	0.76	2.45	0.84	2.72	0.59
Agricultural sciences	34	3.64	0.54	3.55	0.52	3.24	0.53	3.29	0.55	2.81	0.51	2.60	0.60	3.18	0.77
Agricultural sciences	9	4.00	0.51	3.94	0.45	3.67	0.50	3.86	0.42	3.36	0.81	3.04	0.75	3.22	0.58
Agricultural sciences	47	2.03	0.88	2.35	0.93	2.37	0.77	2.45	0.79	2.71	0.84	2.27	0.85	2.37	0.87
Educational psychology	10	3.98	0.51	4.08	0.73	3.80	0.60	3.83	0.46	2.88	0.60	2.23	0.94	2.85	0.87
Educational psychology	46	3.69	0.63	3.56	0.54	2.71	0.78	3.11	0.59	2.79	0.68	2.12	0.77	2.93	0.71
Total	1067	3.61	0.82	3.58	0.78	3.15	0.76	3.30	0.70	2.76	0.80	2.52	0.88	3.00	0.86

and the evaluation was group-based. If students were absent from the group presentations, they had to make an individual reflection task.

Cluster 2 represents pedagogical practices that can be described as *self-directed individual or collaborative knowledge creation practices with content-focused contact teaching*. Courses

TABLE 4 Collaborative knowledge practices scale reliabilities, one-way analyses of variance, and correlations.

Scale	Cronbach alpha	M	SD	F(27,1065)	η^2	Collaborate on objects	Integrate efforts	Feedback	Persistent development	Various disciplines	Interdisciplinary collab.	Exploit technology
Collaborate on objects	0.86	3.61	0.82	23.05***	0.38	1						
Integrate efforts	0.82	3.58	0.78	16.62***	0.30	0.80	1					
Feedback	0.79	3.15	0.76	10.54***	0.22	0.58	0.61	1				
Persistent development	0.77	3.30	0.70	11.16***	0.23	0.78	0.75	0.65	1			
Various disciplines	0.75	2.76	0.80	13.65***	0.26	0.51	0.49	0.47	0.53	1		
Interdisciplinary collaboration	0.73	2.52	0.88	11.31***	0.23	0.46	0.45	0.44	0.49	0.72	1	
Exploit technology	0.85	3.00	0.86	8.73***	0.19	0.54	0.53	0.43	0.53	0.46	0.41	1

***$p < 0.001$. All correlations significant at 0.01 level.

TABLE 5 Pedagogical practices cluster analysis descriptors.

	Cluster		Error		F	Sig.
	Mean square	df	Mean square	df		
Tangible object	5.845	2	0.211	25	27.70	<0.001
Reuse	3.070	2	0.463	25	6.63	0.005
Problem space	5.170	2	0.255	25	20.27	<0.001
Student-centered activities	3.288	2	0.267	25	12.31	<0.001
Modeling of professional practices	6.927	2	0.303	25	22.86	<0.001
Iterativeness	5.000	2	0.120	25	41.67	<0.001
Role of feedback	2.245	2	0.499	25	4.50	0.021
Reflection of practices	4.289	2	0.324	25	13.23	<0.001
Centrality of collaboration	2.113	2	0.351	25	6.02	0.007
Integrating individual and collaborative tasks	2.739	2	0.448	25	6.11	0.007
Multidisciplinarity	0.316	2	0.299	25	1.057	0.362
External collaborators	7.570	2	0.183	25	41.364	<0.001
Digital tools	1.345	2	0.171	25	7.863	0.002
Versatility of assessment methods	0.907	2	0.516	25	1.758	0.193
Assessment of generic competences	9.707	2	0.132	25	73.539	<0.001

included lectures or hands-on sessions and one major open-ended individual or collaborative knowledge creation task (in addition to smaller ones) elaborated mainly outside course meetings. Individual students or groups received occasional and tailored guidance from the lecturers in the meetings to complete the tasks. Six of the courses in Cluster 2 were master level courses, two were for second- or third-year undergraduate students. The average number of participants in the courses was 41.1 (min 25, max 69). The domains of the courses included education, biosciences and forest sciences.

An example in Cluster 2 is a 5 ECTS course (ID92) about the philosophy of science for master students in education for about 25 participants by two lecturers. In the course, each student prepared an individual theoretical essay on a topic chosen by themselves. Students were organized in thematic peer groups where the members commented on each other's essays at certain points during the course. Weekly contact sessions included lectures with discussions and sometimes also working in the thematic groups. The essays were elaborated and commented on mainly as homework. A course platform Moodle was used for sharing materials, peer commenting and task submissions, MS Word for essay writing. The assessment was based on grading from 1 to 5 done by the lecturers; both the quality of essays and peer commenting activity were taken into account in the assessment.

Cluster 3 included courses where practices represented challenging *collaborative knowledge creation with systematic support for epistemic and group work*. Courses were shaped by one major open-ended, authentic and challenging collaborative knowledge creation task (in addition to smaller ones) elaborated both in contact sessions and out-of-class assignments. Professional ways of working were explicitly modeled and

practiced with the students. Two courses in this cluster were for first year students, the others were targeted for third year undergraduate students and/or master students. The average number of students in the courses or study groups (one course was divided in three study groups because the total number of participants was 375) was 51.8 (min 8, max 124). The subject domains of the courses included biosciences, agricultural sciences and education; one course was a multidisciplinary project course.

In Cluster 3, one example of a course (ID96) is a 3 ECTS obligatory project management and work life skills course for third year undergraduate students including about 40 students. During the course, students completed various assignments relevant for their future careers. Individually they made a CV, a LinkedIn profile and a portfolio, gave a personal elevator pitch and interviewed a professional in the field. In groups they made a summary of interviews as a blog post and produced a project plan in groups through a longitudinal process including several phases and sub-tasks. The project topic was given by experts from another university unit, based on a real task from an existing project. Groups competed on the best solution to the project assignment, and the winner was chosen based on the group report and pitching of the solution in the last course meeting. The course was graded on a scale pass-fail, but all sub-tasks had to be completed acceptably and many of them were commented on and assessed both by the course lectures and experts from other university units.

To summarize, what differentiated cluster 1 from the remaining two clusters, based on the level analysis, was that there was no emphasis on an artifact development, iterativeness in the process, structured feedback during the process nor explicit assessment of generic competences. Cluster 2 was differentiated from cluster 3 especially by more emphasis on lecturing and less on student-centered activities, less self-reflection on collaboration process and fewer involvement of external collaborators. All clusters had a rather low level of multidisciplinarity, but quite systematic use of group work practices. This was affected by the fact that all courses included in the study had some type of collaboration task included in the course, because otherwise completing the CKP questionnaires would not have been meaningful for the participants.

Relationship of students' self-evaluations on competence gains and pedagogical practices

We compared the means of the seven scales of CKP against the cluster membership. There were 435, 164, and 467 students in clusters 1, 2, and 3, respectively. A one way ANOVA provided evidence that the clusters had statistically significant differences on each of the CKP scales ($F(2,1065) = 18.7 - 79.3$, $p < 0.001$). A Scheffe *post hoc* test showed that in the scales of Integrate efforts, Persistent development, and Interdisciplinary collaboration, clusters 1 and 2 did not have statistically significant differences, while on the scale Feedback cluster 2 and 3 did not have statistically significant differences. All other comparisons were statistically significantly different ($p < 0.05$). **Figure 3** displays the scale means by pedagogical practices cluster.

Overall, cluster 2 self-directed knowledge creation held the lowest average CKP scale scores, with the exception of Feedback. The students' self-evaluated learning of competence gains in generic competences of collaborative knowledge work was the highest in cluster 3 supported collaborative knowledge creation. As described above, cluster 3 had been qualitatively analyzed as pedagogical practices most specifically designed to support collaborative knowledge creation.

Discussion

The study undertook an examination of generic collaboration competences from three directions. Twenty-eight courses in higher education contributed data to the study, each involving some form of collaborative assignments for the students. Further, each course had some field specific expected learning outcomes regarding the content learning as well as some objectives for gaining knowledge work competences, representing generic competences in collaboration and professional epistemic practices.

First, we investigated how the students evaluated self-assessed competence gains across the courses. We found that there were statistically significant differences between courses in how students rated their learning on the seven scales of the Collaborative Knowledge Practices CKP questionnaire.

Second, we analyzed the pedagogical design decisions made in these courses and developed a categorization of the pedagogical features. This enabled a more detailed examination of how the courses were intended to model and support complex knowledge work competences. This analysis provided three clusters of pedagogical practices. Cluster 1 was considered to represent *activating learning practices promoting content learning*. They emphasized acquisition of domain content through activating lectures and/or small-scale individual and collaborative knowledge creation tasks. These were most prevalent in first year courses in the data. Cluster 2 was named *self-directed individual or collaborative knowledge creation practices* with content-focused contact teaching. Courses included lectures or hands-on sessions and one major open-ended individual or collaborative knowledge creation task elaborated mainly outside course meetings. The knowledge creation challenge was clearly present in the assignments, but strong facilitation for how to carry it out was missing. Cluster 3 was named *collaborative knowledge*

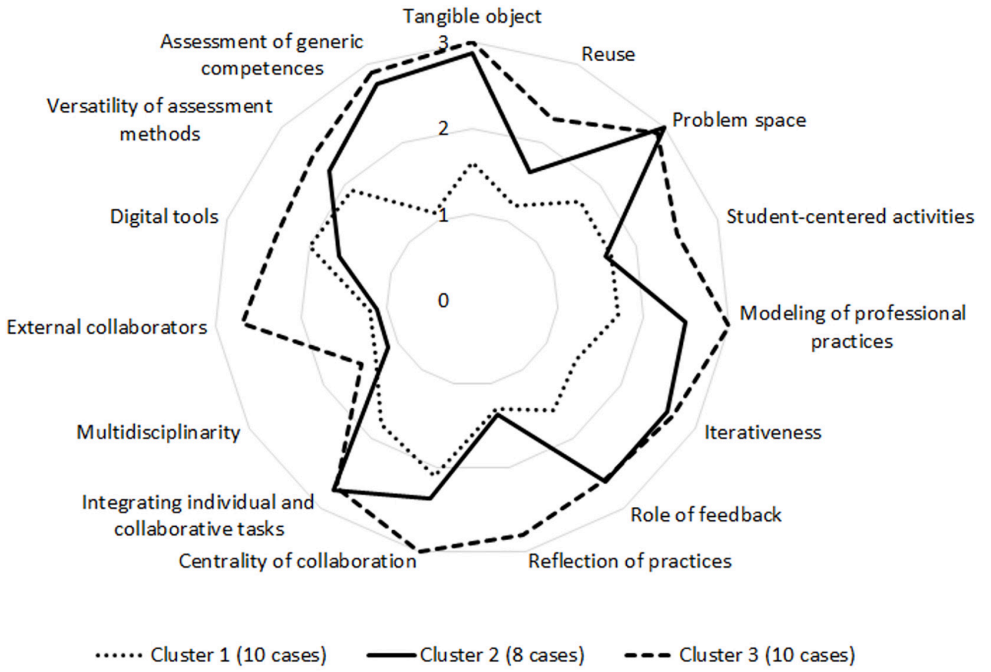


FIGURE 2
Three clusters of pedagogical practices.

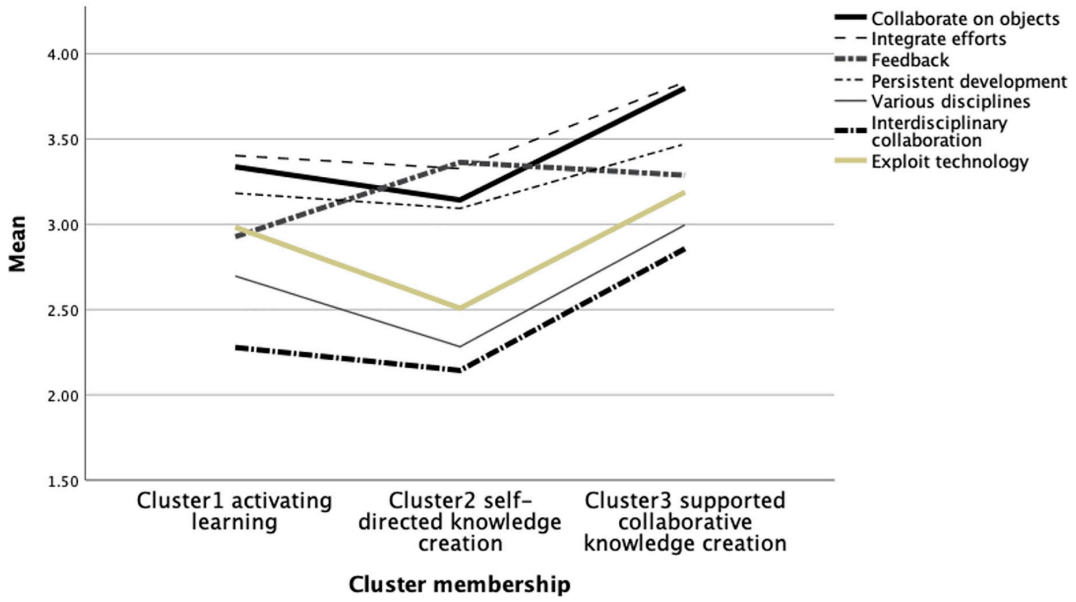


FIGURE 3
Learning of knowledge work competences by pedagogical practices.

creation with systematic support for epistemic and group work. Professional ways of working were explicitly modeled and practiced with the students. This third is argued to be an important addition to current educational practices to ensure graduates' fluent transition to knowledge work.

Thirdly, we investigated how the pedagogical features were related to students' self-evaluations on competence gains. This provided evidence that the three clusters of pedagogical practices were related to differing student evaluations on competence development on the seven CKP scales. Pedagogical

practices of cluster 3 were systematically evaluated to generate more competence gains. Comparison of clusters 1 and 2 offered interesting evidence that the activating learning practices were considered to generate higher competence gains. Further investigation is needed to fully understand these differences. One explanation might be that the self-directed nature of group tasks and shortcomings in the pedagogical support to complete them in cluster 2 courses did not produce strong learning experiences in collaborative knowledge work competences. A parallel result was found in a study focusing on the pedagogical quality of international summer courses (Lakkala et al., 2018): Courses representing traditional teacher-centered lecturing combined with self-directed academic studying outside contact sessions received, on average, lower scores from students in the course evaluation survey, compared with courses that followed practices of activating learning or shared expertise.

The object-orientedness was raised in the qualitative analysis in cluster 2 and 3, but related competence learning was highlighted in cluster 3, which offers crucial information for scaffolding and designing complex collaboration processes. Engaging in integration of efforts is a means to initiate and practice valuable generic competences for collaboration highlighted here, as also pointed out in many prior studies on co-regulation and object-oriented learning (e.g., Borge et al., 2018; Splichal et al., 2018; Damşa and Muukkonen, 2020). But the finding also suggested that mere setting up of collaboration in a course is not enough without modeling and scaffolding. Feedback and cross-fertilization between fields play a key role in all professional practices, therefore becoming competent in them requires effortful practice. On using digital tools in higher education, the outcomes showed that for the most part they were an integrated part of the pedagogical design of courses, with few exceptions.

Theoretically, emergence of the clusters contrast various pedagogical design prototypes for collaborative learning, where Cluster 1 and 3 can be recommended, but cluster 2 points out further support needs for students. The pedagogical features framework offers a tool to examine the design features of collaboration in a structured fashion. Specific aspects targeted in pedagogical designs were also, on average, rated higher by the students on competence gains namely Collaborate on objects and Integrate efforts. The findings suggested higher education can employ student self-evaluations as measures of generic skills development in knowledge work competences. Furthermore, curriculum design can benefit from a systemic approach to mapping and specifying both the features of pedagogical practices and expected learning outcomes on generic skills (see this issue).

This study was conducted as a multi-methods investigation. Contrasting the outcomes of qualitative findings on pedagogical practices and the scaled responses enabled to juxtapose the experienced learning of generic competences and the analyzed pedagogical practices. The sample size of students was quite large and included students from different study years and ages. A limitation of the study is that the age of participants and other background factors like previous work experience were not included as independent variables in the analysis. Further studies should investigate the influence of age on the outcomes. Preliminary examinations with CKP data have suggested that there is considerable individual variation and that a young first year students might provide very different self-evaluation than first year students with prior work experience, so status as first year student is not enough information. This phenomenon might have a relationship with that the questionnaire asks to evaluate the extend of learning in a certain course, and if student experiences that the learning has taken place earlier, then their evaluation may remain moderate. This is a methodological question pertaining to the nature of self-reports more widely. As suggested also by Vogel et al. (2017), the development of collaboration competence might depend on the amount of practice in the corresponding activities, hence previous work experience in knowledge work might influence self-evaluation of new competence gains.

Limitations regarding self-reports of learning have been discussed repeatedly. Although self-report measures are considered suitable for higher education (Roth et al., 2016), there are concerns over the self-report's closeness to actual behavior. Benton et al. (2013) pointed out that the validity of student self-evaluation of learning depends on that relevant learning objective have been identified and, further, whether students can offer accurate evaluations of their learning. Earlier development process of the CKP questionnaire identified and removed those items that were ambiguous or uninformative, thus contributing to relevance in term of targeted and validated questionnaire. The relevance in terms of the course specific learning objectives was evaluated by the students, which aligned with the teacher descriptions in a theoretically and pedagogically integrated way. Further research will need to add parallel means to evaluate student learning, e.g., by pre-post or teacher evaluations.

Some courses had multiple teachers and how this teacher-student ratio influences the pedagogical choices and abilities to support students is an important future research question. Contextual issues have an effect on the pedagogical decisions, e.g., in mass courses it is not often possible to implement complex design for collaborative knowledge practices. Nevertheless, among the investigated courses there were also courses with large number of participants (e.g., a course in education, about 100 participants in one teaching group) and the courses were implemented representing collaborative knowledge creation practices with systematic support (cluster 3).

The transformation to fully online and hybrid teaching is a further design challenge for higher education. Previous research has suggested that it is rare in online learning settings that students engage in high-quality activities or knowledge creation spontaneously (Kobbe et al., 2007). The current findings evidenced that for teaching collaboration and

generic collaboration competences, there are a multitude of pedagogical design decisions to make. The courses in cluster 3 had a distinct emphasis on modeling, offering teacher support during contact teaching, and reserving time for reflection. This demands significant before class preparation from teachers and using time to introduce various tools and models, e.g., for expected learning outcomes for generic competence or orchestration of multidisciplinary collaboration. It is aligned with the call for *design for learning* (Goodyear, 2015), which entails investing more heavily in the planning phase and recognition of the primary role of design for student learning. Further, the results point out a need for a well-structured digital environment for sharing materials or links to other services, practical information, and guidance about online learning practices and assessment. These are needed to enable students to concentrate on the learning goals instead of trying to find out what is expected of them. Besides content learning objectives, teachers' social presence is especially important for modeling demanding generic competences and interaction around knowledge creation.

Data availability statement

The datasets presented in this article are not readily available because not all data can be rendered anonymous. Anonymized raw student responses supporting the conclusions of this article can be made available by the authors. Teacher data cannot be made fully anonymous, metadata therefrom can be made available. Requests to access the datasets should be directed to (HM, Hanni.muukkonen@oulu.fi).

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Funding

This work was partly supported by the Academy of Finland (grant number: 285806).

Acknowledgments

We are very thankful to the students and teachers participating in the study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Baker, M. J., Andriessen, J., and Järvelä, S. (2013). *Affective Learning Together: Social and Emotional Dimensions of Collaborative Learning.* Milton Park: Routledge. doi: 10.4324/9780203069684

Barrie, S. C. (2012). A research-based approach to generic graduate attributes policy. *High. Educ. Res. Dev.* 31, 65–77. doi: 10.1080/07294360.2012.642842

Benoliel, P., and Somech, A. (2015). The role of leader boundary activities in enhancing interdisciplinary team effectiveness. *Small Group Res.* 46, 83–124. doi: 10.1177/1046496414560028

Benton, S. L., Duchon, D., and Pallett, W. H. (2013). Validity of student self-reported ratings of learning. *Assess. Eval. High. Educ.* 38, 377–388. doi: 10.1080/02602938.2011.636799

Bereiter, C. (2002). *Education and Mind in the Knowledge Age*. Hillsdale: Erlbaum.

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., et al. (2012). "Defining twenty-first century skills," in *Assessment and Teaching of 21st Century Skills*, eds P. Griffin, B. McGaw, and E. Care (New York, NY: Springer), 17–66. doi: 10.1007/978-94-007-2324-5_2

Borge, M., Ong, Y. S., and Penstein Rosé, C. (2018). Learning to monitor and regulate collective thinking processes. *Int. J. Comput. Support. Collab. Learn.* 13, 61–92. doi: 10.1007/s11412-018-9270-5

Boud, D., and Malloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assess. Eval. High. Educ.* 38, 698–712. doi: 10.1080/02602938.2012.691462

Carless, D., Salter, D., Yang, M., and Lam, J. (2011). Developing sustainable feedback practices. *Stud. High. Educ.* 36, 395–407. doi: 10.1080/03075071003642449

Castañeda, L., and Selwyn, N. (2018). More than tools? Making sense of the ongoing digitizations of higher education. *Int. J. Educ. Technol. High. Educ.* 15:22. doi: 10.1186/s41239-018-0109-y

Collis, B., and Moonen, J. (2008). Web 2.0 tools and processes in higher education: quality perspectives. *Educ. Media Int.* 45, 93–106. doi: 10.1080/09523980802107179

Cooke, N. J., and Hilton, M. L. (eds) (2015). *Enhancing the Effectiveness of Team Science*. Washington, DC: NAP.

Cremers, P. H. M., Wals, A. E. J., Wesselink, R., and Mulder, M. (2016). Design principles for hybrid learning configurations at the interface between school and workplace. *Learn. Environ. Res.* 19, 309–334. doi: 10.1007/s10984-016-9209-6

Cresswell, J. W. (2009). Mapping the field of mixed methods research. *J. Mix. Methods Res.* 3, 95–108. doi: 10.1177/1558689808330883

Damşa, C., and Muukkonen, H. (2020). Conceptualising pedagogical designs for learning through object-oriented collaboration in higher education. *Res. Pap. Educ.* 35, 82–104. doi: 10.1080/02671522.2019.1677751

De Laat, M. F., Lally, V., Lipponen, L., and Simons, P. R. J. (2007). Online teaching in networked learning communities: a multi-method approach to studying the role of the teacher. *Instr. Sci.* 35, 257–286. doi: 10.1007/s11251-006-9007-0

Esterhazy, E., Nerland, M., and Damsşa, C. (2021). Designing for productive feedback: an analysis of two undergraduate courses in biology and engineering. *Teach. High. Educ.* 26, 806–822. doi: 10.1080/13562517.2019.1686699

Faraj, S., Jarvenpaa, S. L., and Majchrzak, A. (2011). Knowledge collaboration in online communities. *Organ. Sci.* 22, 1224–1239. doi: 10.1287/orsc.1100.0614

Goodyear, P. (2015). Teaching as design. *Rev. High. Educ.* 2, 27–50.

Goodyear, P., and Zenios, M. (2007). Discussion, collaborative knowledge work and epistemic fluency. *Br. J. Educ. Stud.* 55, 351–368. doi: 10.1111/j.1467-8527.2007.00383.x

Hadwin, A. F., Järvelä, S., and Miller, M. (2017). "Self-regulation, co-regulation and shared regulation in collaborative learning environments," in *Handbook of Self-Regulation of Learning and Performance*, eds D. Schunk and J. Greene (Milton Park: Routledge), 65–86. doi: 10.4324/9781315697048-6

Ilomäki, L., Lakkala, M., and Toom, A. (2017). "Knowledge work assignments in upper secondary school: Results of 13 cases," in *Proceedings of the 17th Biennial Conference for Research on Learning and Instruction (EARLI)*, Tampere.

Jung, J. (2022). Working to learn and learning to work: research on higher education and the world of work. *High. Educ. Res. Dev.* 41, 92–106. doi: 10.1080/07294360.2021.2002274

Karlgren, K., Paavola, S., and Ligorio, B. (2020a). Introduction: what are knowledge work practices in education? How can we study and promote them? *Res. Pap. Educ.* 35, 1–7. doi: 10.1080/02671522.2019.1677761

Karlgren, K., Lakkala, M., Toom, A., Ilomäki, L., Lahti-Nuuttila, P., and Muukkonen, H. (2020b). Assessing the learning of knowledge work competence in higher education – cross-cultural translation and adaptation of the Collaborative Knowledge Practices Questionnaire. *Res. Pap. Educ.* 35, 8–22. doi: 10.1080/02671522.2019.1677752

Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., Häkkinen, P., et al. (2007). Specifying computer-supported collaboration scripts. *Int. J. Comput. Support. Collab. Learn.* 2, 211–224. 2-007-9014-4 doi: 10.1007/s11412-007-9014-4

Laakkonen, J., and Muukkonen, H. (2019). Fostering students' collaborative learning competencies and professional conduct in the context of two gross anatomy courses in veterinary medicine. *Anat. Sci. Educ.* 12, 154–163. doi: 10.1002/ase.1811

Lakkala, M., Ilomäki, L., Mikkonen, P., Muukkonen, H., and Toom, A. (2018). Evaluating the pedagogical quality of international summer courses in a university program. *Int. J. Res. Stud. Educ.* 7, 89–104. doi: 10.5861/ijrse.2017.1781

Lakkala, M., Muukkonen, H., Ilomäki, L., and Toom, A. (2020). "Framework for evaluating the pedagogical features of university courses representing collaborative knowledge work practices," in *The Interdisciplinarity of the Learning Sciences, 14th International Conference of the Learning Sciences (ICLS) 2020*, eds M. Gresalfi and I. S. Horn (Nashville: International Society of the Learning Sciences), 1757–1758.

Lakkala, M., Toom, A., Ilomäki, L., and Muukkonen, H. (2015). Re-designing university courses to support collaborative knowledge creation practices. *Australas. J. Educ. Technol.* 31, 21–536. doi: 10.14742/ajet.2526

Markauskaite, L. (2020). Commentary: learning for knowledge work practices in the wild. *Res. Pap. Educ.* 35, 105–115. doi: 10.1080/02671522.2019.1677762

Markauskaite, L., and Goodyear, P. (2017). *Epistemic Fluency and Professional Education*. New York, NY: Springer. doi: 10.1007/978-94-007-4369-4

McCune, V. (2009). Final year biosciences students' willingness to engage: teaching–learning environments, authentic learning experiences and identities. *Stud. High. Educ.* 34, 347–361. doi: 10.1080/03075070802597127

Miettinen, R., and Paavola, S. (2018). "Beyond the distinction between tool and sign: objects and artefacts in human activity," in *The Cambridge Handbook of Social–Cultural Psychology. Second Edition*, eds A. Rosa and J. Valsiner (Cambridge: Cambridge University Press), 148–162. doi: 10.1017/9781316662229.009

Miettinen, R., and Virkkunen, J. (2005). Epistemic objects, artefacts and organizational change. *Organization* 12, 437–456. doi: 10.1177/1350508405051279

Muukkonen, H., and Lakkala, M. (2009). Exploring metaskills of knowledge-creating inquiry in higher education. *Int. J. Comput. Support. Collab. Learn.* 4, 187–211. doi: 10.1007/s11412-009-9063-y

Muukkonen, H., Lakkala, M., Lahti-Nuuttila, P., Ilomäki, L., Karlgren, K., and Toom, A. (2020). Assessing the development of collaborative knowledge work competence: scales for higher education course contexts. *Scand. J. Educ. Res.* 64, 1071–1089. doi: 10.1080/00313831.2019.1647284

Näykki, P., Järvelä, S., Kirschner, P. A., and Järvenoja, H. (2014). Socio-emotional conflict in collaborative learning - A process-oriented case study in a higher education context. *Int. J. Educ. Res.* 68, 1–14. doi: 10.1016/j.ijer.2014.07.001

OECD (2018). *Education at a Glance 2018: OECD Indicators*. Paris: OECD Publishing. doi: 10.1787/eag-2018-en

Paavola, S., and Hakkarainen, K. (2005). The knowledge creation metaphor—An emergent epistemological approach to learning. *Sci. Educ.* 14, 535–557. doi: 10.1007/s11191-004-5157-0

Paavola, S., Lakkala, M., Muukkonen, H., Kosonen, K., and Karlgren, K. (2011). The roles and uses of design principles for developing the trialogical approach on learning. *Res. Learn. Technol.* 19, 233–246. doi: 10.3402/rlt.v19i3.17112

Puntambekar, S., Stylianou, A., and Goldstein, J. (2007). Comparing classroom enactments of an inquiry curriculum: lessons learned from two teachers. *J. Learn. Sci.* 16, 81–130. doi: 10.1080/10508400709336943

Roth, A., Ogrin, S., and Schmitz, B. (2016). Assessing self-regulated learning in higher education: a systematic literature review of self-report instruments. *Educ. Assess. Eval. Account.* 28, 225–250. doi: 10.1007/s11092-015-9229-2

Säljö, R. (2010). Digital tools and challenges to institutional traditions of learning: Technologies, social memory and the performative nature of learning. *J. Comput. Assist. Learn.* 26, 53–64. doi: 10.1111/j.1365-2729.2009.00341.x

Scardamalia, M., and Bereiter, C. (2014). "Knowledge building and knowledge creation: Theory, pedagogy, and technology," in *Cambridge Handbook of the Learning Sciences*, ed K. R. Sawyer (Cambridge: Cambridge University Press), 397–417. doi: 10.1017/CBO9781139519526.025

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educ. Res.* 27, 4–13. doi: 10.3102/0013189X027002004

Splichal, J. M., Oshima, J., and Oshima, R. (2018). Regulation of collaboration in project-based learning mediated by CSCL scripting reflection. *Comput. Educ.* 125, 132–145. doi: 10.1016/j.compedu.2018.06.003

Timmermans, S., and Tavory, I. (2012). Theory construction in qualitative research: from grounded theory to abductive analysis. *Sociol. Theory* 30, 167–186. doi: 10.1177/0735275112457914

Vesikivi, P., Lakkala, M., Holvikivi, J., and Muukkonen, H. (2020). The impact of a project-based learning curriculum on first-year retention, study experiences, and

knowledge work competence. *Res. Pap. Educ.* 35, 64–81. doi: 10.1080/02671522.2019.1677755

Viswanathan, L., Whitelaw, G. S., and Meligrana, J. (2012). Evaluating the role of the project course in professional planning education and its influence on planning policy and practice. *Plan. Pract. Res.* 27, 387–403. doi: 10.1080/02697459.2012.673329

Vogel, F., Wecker, C., Kollar, I., and Fischer, F. (2017). Socio-cognitive scaffolding with computer-supported collaboration scripts: a meta-analysis. *Educ. Psychol. Rev.* 29, 477–511. doi: 10.1007/s10648-016-9361-7

Yin, R. K. (2014). *Case Study Research. Design and Methods (5th ed)*. Thousand Oaks: Sage Publications.

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 08 August 2022
DOI 10.3389/feduc.2022.886375

OPEN ACCESS

EDITED BY
Edith Braun,
Justus-Liebig Universität, Germany

REVIEWED BY
Bettina Hannover,
Freie Universität Berlin, Germany
Ilka Benner,
University of Giessen, Germany

*CORRESPONDENCE
Soo Jeung Lee
soojlee@sejong.ac.kr

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Higher Education, a section of the journal Frontiers in Education

RECEIVED 28 February 2022
ACCEPTED 05 July 2022
PUBLISHED 08 August 2022

CITATION
Lee A and Lee SJ (2022) Korean university students' significant learning experiences and associated generic skills: A qualitative essay review. *Front. Educ.* 7:886375. doi: 10.3389/feduc.2022.886375

Korean university students' significant learning experiences and associated generic skills: A qualitative essay review

Ahram Lee† and Soo Jeung Lee*†

Department of Education, Sejong University, Seoul, South Korea

The purpose of the present study was to explore significant learning experiences of Korean university students and examine associated generic skills. The study implemented a document analysis approach to investigate essays collected from 33 students in a 4-year university in Seoul, South Korea. A total of 102 excerpts were coded, forming 14 sub-themes which were categorized into five themes that describe students' significant learning experiences. The five themes are interacting with others, learning by oneself and about oneself, realizing applicability to real-life, venturing into advanced learning, and experiencing a respectful learning atmosphere. Also, 18 generic skills were identified which were categorized into four clusters, namely comprehensive thinking skills, information utilization skills, interpersonal skills, and personal attributes. The results of the present study provide the groundwork for understanding students' perceptions of significant learning experiences and associated generic skills.

KEYWORDS

significant learning experiences, generic skills, core competencies, document analysis, university students

Introduction

The rapid changes of the twenty-first century are marked by factors such as the advancement of technology and recession (Blustein, 2019), a sharp decline in the school-age population (Ban, 2016; Kim, 2022), and prolonged COVID-19 (Karalis and Raikou, 2020), calling for university education to change. In order to meet the varying needs of the ever-changing society, discussions on *well-teaching universities* have actively taken place in South Korea since the late 2000s. Specifically, demands for university education innovation have emerged, starting with the 2008 University Education Capacity Enhancement (UECE) project and the 2010 Advancement of College Education (ACE) project, focusing on enhancing the quality of education for students (Byoun, 2018).

These projects aim to foster *well-teaching universities* by establishing competencies-based education systems and curricula (Lee, 2017). As such, the key components of the education innovation taking place in South Korea are twofold; to teach well and to foster core competencies. Thus, the universities in South Korea are actively devising ways to take part in the nationwide innovation movement in order to provide more meaningful learning to students and to assess the learning outcomes based on the established core competencies.

In alignment with such education innovation, it would be imperative to first understand students' perceptions of what constitutes meaningful and significant learning. Significant learning involves experiences in which students apply knowledge across various situations, actively explore knowledge, and immerse themselves in the learning process (Strange and Banning, 2015). Fink (2003) pointed out that when classes are designed to foster significant learning experiences, students are motivated and can actively participate in learning. However, previous studies mainly focused on designing instructions to facilitate significant learning from the perspective of instructors (Saulnier, 2003; Levine et al., 2008; Trudeau and Kruse, 2014; Sanchez et al., 2020). There are only a few studies in South Korea that examined students' perceptions of significant learning experiences (Han and Hwang, 2021). Therefore, the current study intended to explore and understand students' own experiences of significant learning.

Moreover, the study tried to identify the generic skills associated with students' significant learning experiences. Generic skills are transferrable skills that can be applied in varying contexts beyond the boundaries of specific disciplines (Barrie, 2006) and are also referred to as core skills or core competencies (Bratianu and Vatamanescu, 2017; Virtanen and Tynjälä, 2019). In South Korea, the term *core competencies* is used to refer to generic skills that need to be fostered through university education. In accordance with various government projects that are leading the innovation of university education in South Korea, such as the aforementioned ACE project, the Ministry of Education and the South Korea Research Institute for Vocational Education and Training developed the Korea Collegiate Essential Skills Assessment (K-CESA) in order to measure six basic competencies of university students, namely communication competency, global competency, resource information technology utilization competency, comprehensive thinking competency, interpersonal competency, and self-management competency. Individual universities have also established their own core competencies, based on which they assess students and rearrange the curriculum and extracurricular activities (e.g., Park and Chung, 2017; Choi, 2020; Shin et al., 2021). Such emphasis on competencies has led Korean universities to evaluate learning outcomes, identify *good class*, and foster competitive global personnel required by the future job market based on the core competencies that they established (Jin et al., 2011).

While the importance of establishing core competencies or generic skills at the university level is recognized in South Korea, it should also be noted that the development of such skills need to be understood in the context of students' significant learning experiences. The development of generic skills is inseparable from the learning process (Drummond et al., 1998; Bath et al., 2004). Thus, it would be important to examine generic skills that are associated with learning experiences that students perceive as being significant.

Thus, the current study investigated university students' own experiences of learning generic skills in the context of their learning process, by first exploring the significant learning experiences identified by students, and then analyzing generic skills entailed in those experiences. In order to reflect the context of higher education in South Korea, in which core competencies-based education program is recently being emphasized, this study applied the document analysis method to examine essays written by students studying in a university located in Seoul, South Korea. The research questions are as follows.

(1) What experiences do students identify as being significant learning experiences at university?

(2) What generic skills are associated with students' significant learning experiences?

Literature review

Significant learning experiences involve applying knowledge across varying contexts, actively exploring new knowledge, and being immersed in the learning process (Strange and Banning, 2015). Engaging in significant learning experiences is different from merely cramming knowledge in that students not only understand what they learn and apply it to real situations but also get to know more about themselves and the world around them (Bae and Hwang, 2021). Thus, significant learning experiences are often discussed in relation to higher education (Fink, 2003).

It is not easy to define significant learning experiences. Learning is a personal, intentional, dynamic, and interactive process that tentatively generates an output characterized by specific knowledge produced at a given moment and context (Agra et al., 2019). Ausubel (2000, p. 1) used the term *meaningful reception learning* and defined it as "the acquisition of new meanings from presented learning material," and explained that meaningful learning takes place when students gain understanding by linking new knowledge and experiences to what they already know. The interaction between potentially new meanings and related ideas in the students' cognitive structure generates actual or psychological meanings. Since each student's cognitive structure is different, all newly acquired meanings are unique in themselves. Ausubel (1963) viewed meaningful reception learning as a result of the teacher's

successful leadership of the teaching-learning process. Based on Ausubel's theory, Joyce et al. (2000) proposed the instructional design principles to frame learner performance into goals and tasks, divide these tasks into small component tasks, design training activities, and align learning events in order to promote the transfer of prerequisite learning.

In order to provide a comprehensive understanding of significant learning, Fink (2013) suggested dynamic linkages of the following six elements that change students in significant ways: (1) *foundational knowledge* refers to basic knowledge that constitute the content of a course that is necessary for one to understand and remember in a course; (2) *application* indicates applying the previously learned knowledge to other situations so that students can acquire and develop advanced knowledge and skills; (3) *integration* refers to expanding one's thinking by linking acquired knowledge and ideas, such as identifying similarities and differences between subjects or different theories; (4) *human dimension* indicates actively interacting with others through which one gains deeper understanding of him or herself and the word; (5) *caring* is the process of changing in feelings, interests, or values in relation to a subject, and is also related to motivations for learning; (6) *learning how to learn* refers to learning skills that help one become a better and more self-directed learner and keep on learning even after the course is over.

Fink (2013) mentioned that any course can apply these six elements to create significant learning experiences. It was reported that in the class based on significant learning experiences with the six aforementioned elements, students learned how to think creatively, learned together with others, and learned how the content of the class affected them and the world (Levine et al., 2008; Evans et al., 2016). The significant learning experiences have strength in that they mean a holistic learning experience in which students learn knowledge and apply the learned knowledge to real problems and get to know themselves and the world through them (Strange and Banning, 2015).

However, previous studies on significant learning experiences have mainly focused on the instructors and how they can design and create significant learning experiences for students across disciplines (Saulnier, 2003; Levine et al., 2008; Trudeau and Kruse, 2014; Sanchez et al., 2020). These studies examined significant learning experiences proposed by Fink (2013) from the perspective of instructors. Very few studies, to the authors' knowledge, have examined significant learning experiences from the students' perspectives (e.g., Bae and Hwang, 2021; Han et al., 2021). Therefore, the current study intended to investigate what experiences students identify as being significant to themselves.

Moreover, since significant learning is not simply learning and memorizing content material but integrating knowledge and applying it in varying life situations, such learning experiences are closely related to generic skills. Generic skills are transferrable skills that can be applied in different contexts beyond the boundaries of specific disciplines (Barrie, 2006). Such skills are also referred to as core skills, key skills, generic attributes, core competencies, or employability skills (Bratianu and Vatamanescu, 2017; Virtanen and Tynjälä, 2019).

While courses specifically designed for generic skills may not be very effective (Hattie et al., 1996), there is a widespread belief that generic skills can be developed through various learning experiences at university (Kember et al., 2007). Kember et al. (2007) tried to connect generic skills to the teaching and learning environment and proposed that the characteristics of the teaching and learning environment that particularly developed students' generic skills were active learning, teaching for understanding, assessment, coherence of curriculum, teacher-student interaction, feedback to assist learning, assistance from teaching staff, relationship with other students, and cooperative learning. They addressed that the teaching and learning environment has a greater-than-expected impact on the development of generic skills (Kember and Leung, 2005; Leung and Kember, 2005; Kember et al., 2007). Virtanen and Tynjälä (2019) examined the types of pedagogical practices related to students' learning of generic skills and found that the learning of generic skills was not affected by any specific pedagogical practice, but rather required the use of various teaching methods and pedagogical practices.

It would not be possible to provide an exhaustive list of what constitutes generic skills. Attempts have been made to extract core competencies and sub-elements. The Organization for Economic Co-operation and Development (OECD) initiated the Definition and Selection of Competencies (DeSeCo) project in 1997 and defined core competencies as "the ability to meet complex demands, by drawing on and mobilizing psychosocial resources (including skills and attitudes) in a particular context" (OECD, 2005, p. 4). The core competencies in the DeSeCo project were classified into three broad, interrelated categories such as "use tools interactively," "interact in heterogeneous groups," and "act autonomously." Based on the DeSeCo project, the OECD Education 2030 project has identified three further categories of competencies such as "creating new value," "reconciling tensions and dilemmas," and "taking responsibility" (OECD, 2018). Previous studies have also examined various skills as being generic skills. Badcock et al. (2010) examined critical thinking, interpersonal understandings, problem-solving, and written communication to assess generic skills. Braun and Leidner (2009) suggested six domains of competencies: knowledge processing, systematic, presentational, communication, cooperation, and personal competence. In addition to these, Bath et al. (2004) included skills such as intellectual curiosity and rigor, ethical awareness and practice, integrity, and tolerance. Virtanen and Tynjälä (2019) examined pedagogical approaches that can foster the

following eight generic skills: resourcefulness, innovativeness, creativity, ability to operate in new situations, decision-making skills, ability to solve occupational problems, continuing learning skills, and self-assessment skills. The current study investigated students' own experiences of engaging in generic skills in the context of their learning process and identified generic skills associated with students' significant learning experiences.

Method

Data collection

This study used essays written by students attending a 4-year university in Seoul, South Korea. The essays were collected in the fall semester of 2021 from two courses in the Department of Education, namely Educational Administration and Adult Learning and Counseling, conducted by the corresponding author and the first author, respectively. The students enrolled in these courses were asked to write a reflection paper on their significant learning experiences as a part of their course assignments. Students were instructed to think about their significant learning experience(s) that they had during their university years and to freely describe their experience(s) in detail. The following guiding questions were provided: What were the activities you participated in?; What did you experience?; What made the experience significant?; and What factors were related to your experience?. Each essay was one to two pages long. After the end of the semester, when all the course evaluation was completed, the description of the research and a consent form permitting the use of the essays for research was announced to the students *via* the learning management system (LMS) used in the university.

Initially, a total of 43 essays were collected during the semester, and among them, 33 students voluntarily sent their written consent to allow the use of their essays. The 10 students who did not send the consent may not have wanted to participate in the research or they may not have checked the announcement sent through the LMS since the semester was over. Among the 10 students, two were international students from the Department of Hotel Management, and the rest were from the Department of Education; seven were female and three were male.

The final sample used in this study was 33 essays. Most students who consented to research were from the Department of Education with few exceptions: two from the Department of Engineering, two from the Department of Korean Language and Literature, one from the Department of English Language and Literature, and one from the Department of Humanities. The grade level of students ranged from sophomore to senior. The number of female students was 24 and male students was 9.

Analysis

The current study used a document analysis approach to assess essays written by university students. Document analysis involves a systematic procedure to examine and interpret texts and images of varying forms, including both public and private records, to gain meaning and understanding of the chosen phenomenon (Bowen, 2009). The documents used for analysis in this study were students' essays on significant learning experiences that they had during their university years. These documents were selected to examine students' perceptions on significant learning and to reflect the context of Korean education.

The study adhered to the following analytic procedure proposed by Braun and Clarke (2006) to derive themes and identify patterns and meaning from the collected data: (1) familiarizing yourself with your data; (2) generating initial codes; (3) searching for themes; (4) reviewing themes; (5) defining and naming themes; and (6) producing the report. First, the authors individually read the students' essays multiple times to absorb the rich details and to understand the essence of what the students were describing in their essays. While reading and re-reading the essays, we noted our initial ideas for the possible codes. Second, we compared our ideas for codes, examined the exemplary quotations, and generated initial codes. We listed the codes and went back to the essays, systematically coding the data and collating excerpts relevant to each code. In the process of finding supporting excerpts, the codes were added or removed, combined or divided, and renamed. For instance, an initial code named "participating in discussions" was divided into "saying my opinion in discussions" and "listening to others in discussions." In the third step, the revised codes were grouped into potential overarching themes, and we put initial names for each theme. Fourth, we reviewed the themes as well as the codes and their supporting excerpts within each theme. Similar themes were merged while themes without enough supporting data were removed, and sub-themes were generated. We checked for a coherent pattern of the codes within each theme. We also went through the essays again to make sure that the themes were accurately reflecting the meaning. In the fifth step, the finalized themes and sub-themes were named. Finally, we tried to build a narrative for each theme which is provided in the following Results section.

Although the main purpose of the analytic procedure was to derive themes for significant learning experiences, we also wanted to know what generic skills were associated with each of the themes. Therefore, in the process of coding, we also extracted generic skills that students directly mentioned in their essays and formed a separate list of codes for generic skills. After finalizing the themes for significant learning experiences, we listed the codes for generic skills within each theme and grouped similar codes into clusters, based on the classification by Braun and Leidner (2009) and K-CESA. Then, we named the clusters for

the generic skills. We tried to secure the validity of the study in terms of consistency and neutrality (Guba and Lincoln, 1985) and tried to maintain objectivity by seeking opinions on data analysis and results through a continuous consultation process so that the results derived through data collection and analysis were consistent, excluding any bias in the research process.

Results

The essays described various course-related and extracurricular activities that were identified as a significant learning experience. A total of 102 excerpts were coded and formed into 14 sub-themes, which were reorganized into five themes about significant learning experiences, namely interacting with others, learning by oneself and about oneself, realizing applicability in real-life, venturing into advanced learning, and experiencing respectful learning atmosphere.

From these experiences, 18 generic skills were induced which were categorized into four clusters. First is comprehensive thinking skills including critical thinking, creative thinking, problem-solving skills, recognizing and reducing prejudice, synthesis, and intellectual curiosity. Second is information utilization skills including information management skills and ability to apply knowledge into practice. Third is interpersonal skills including openness to others, collaborative skills, and communication skills. Fourth is personal attributes including the capacity to learn actively, autonomy, passionate attitude, self-confidence, self-understanding, self-esteem, and establishing a sense of value.

The five themes related to significant learning experiences and four clusters of associated generic skills are illustrated in **Figure 1**. Specific descriptions of the five themes and 14 sub-themes of significant learning experiences and the associated generic skills are provided below.

Interacting with others

Most of the students identified learning experiences in which they were engaged with one another as being significant. Interaction-based learning experience encompasses asserting one's own ideas, listening to others, receiving feedback, and engaging in team activities.

Asserting one's own ideas

Asserting one's own ideas to others in discussions or debates was identified as a significant experience. By trying to verbalize their perspectives, students were able to "construct knowledge" in their own way and "freely generated ideas." One student described her experience of having unguided discussions as "feeling as if I were an ancient philosopher sitting around a Greek bath sharing scholarly knowledge" with colleagues. Moreover, having an audience to deliver one's perspective led students to become more familiar with the content knowledge and memorize important information better.

> Each class was divided into groups for discussion, and a representative was selected to lead the discussion. Befitting the title the professor gave us, I found myself explaining what I knew to classmates like a real teacher. . . The reason why I have to call it a significant learning experience is that the contents I shared with my classmates are still perfectly remembered in my mind. For me, true learning occurs when I am actively using the knowledge so that it is stored in my long-term memory, and I can explain it to others whenever I want (Student 6).

Students reporting these experiences needed to grasp the essence of content materials and generate their own perspectives integrating what they have learned. Thus, they expressed these experiences as significant learning experiences when utilizing comprehensive thinking skills, such as creative thinking and synthesis. Also, they had to process and organize knowledge and opinion in such a way so that they could deliver it to others, involving information management skills.

Not only was asserting oneself meaningful in terms of knowledge, but it was also significant to some students at a personal level. Students reported that they were able to become more open to others and share their personal experiences by having discussions. One student also reported that he could establish his own values more firmly while he was trying to persuade his friends. The skills involved in these experiences are named openness to others and establishing a sense of value, respectively.

Listening to others' perspectives

Listening to others' opinions in discussions was an essential element of a significant learning experience. Most of all, students could gain different perspectives and expand their worldview by sharing others' viewpoints.

> By sharing opinions and conducting a discussion, we can rethink the topic in a direction different from what we initially thought and advocated. Through this, I think critical reflection on my opinion occurred, subsequently making changes in knowledge that I previously knew (Student 9).

Listening to others' ideas generated new inquiries, allowed students to synthesize different theories, and led the thought process in a different direction. Such broadening of perspective led students to think more critically and solve problems in innovative ways. It even guided students to reflect on themselves and face their own prejudice and rigidity, leading to a critical evaluation of oneself. One student also reported that he was "able to change his negative thoughts into positive ones" by listening to other opinions in discussions.

> During the discussion, I was able to look at the subject from various perspectives through critical thinking, and I was able to look back on myself, breaking the prejudice that I implicitly held. I felt a change in my thoughts, breaking my own stereotypes and looking from a new perspective (Student 21).

Students reported that these experiences were significant to them because they were involved in critical thinking, recognizing and reducing prejudice, and engaging in synthesis. Generic skills integrated in the experience of listening to others' perspectives were mostly comprehensive thinking skills.

Receiving feedback

Receiving feedback from lecturers and peers led to meaningful learning outcomes. Feedback was closely related to motivation. Receiving immediate feedback made students look forward to the class and participate more actively in activities, regardless of whether the feedback was positive or negative.

> A specific learning method that was meaningful to me was task-based learning and immediate feedback. In each semester, one or two courses offered weekly quizzes (multiple choice or written essay). In these courses, immediate feedback was provided each week, whether it be multiple choice or written essay. In the reconstructing the knowledge experienced in these courses, the immediate feedback provided by the instructor each week was the main motivator for learning (Student 2).

These students were motivated and proactive to learn and reported experiencing synthesis of knowledge by obtaining different perspectives through feedback. Students who received positive feedback reported gaining self-confidence and generating new and better ideas. Students receiving negative feedback also reported such an experience as being significant because it allowed them to identify areas of growth.

The most notable generic skill identified by students, whose significant learning experience was receiving feedback, was the capacity to learn actively. Students' perception of feedback being motivational and encouraging was related to actively participating in the learning process. Other skills included self-confidence and comprehensive thinking skills such as synthesis and intellectual curiosity.

Engaging in team activities

Engaging in team activities was significant interaction-based learning experience. Team activities may encompass the aforementioned discussions, debate, and feedback, but a notable aspect to consider would be togetherness. Students reported experiencing "a sense of belongingness" which allowed them to explore more since they felt reassured within their team. While trying to tackle a problem together, students engaged in productive discussions and critical thinking to find the appropriate solution.

> A long-term project was carried out as a small group activity. As a transfer student, I had a lot of difficulty in the beginning. . . since I did not know the campus culture or the overall atmosphere of the school. However, I was able to have social interactions with upperclassmen through the small group activities, and as a result, I found out that the campus museum was not being operated actively. By working with my excellent colleagues, I was able to find a solution to the problem that I could not have been able to solve on my own. As such, my knowledge construction takes place in a socio-cultural environment, so I came to realize again that learning cannot occur independently and that learning is a result of social interaction (Student 29).

In the process, most students described their experience of collaboration and communication with the teammates as being meaningful. A student expressed satisfaction of being able to produce high-quality assignment through collaboration. Accordingly, most emphasized generic skills in team activities were interpersonal skills such as collaboration skills and communication skills due to teammates striving for a common goal. Enhanced critical thinking and problem-solving skills were reported as a result of the collaboration.

Learning by oneself and about oneself

Learning that is led by one's directiveness and centers around oneself was a prominent experience students identified as being significant. There are largely two categories within this theme, namely self-directed learning and self-reflective learning.

Self-directed learning

Self-directed learning was described by twelve students and one of the common characteristics of the experience was that students were more motivated, passionate, and proactive about what they were learning. It was reported that students were able to focus on knowledge that they wanted to learn which made them look forward to learning.

> I initiated and continued learning based solely on my own will, organizing the contents and methods according to my own preference, and freely filled in the missing points scattered here and there. When there were no tests, I was learning what I wanted to learn, and not the knowledge that would be tested on an exam. I could focus more on the knowledge that I need and want to know right now. Also, passionate learning with complete intrinsic motivation was possible without getting pressure from anything (Student 3).

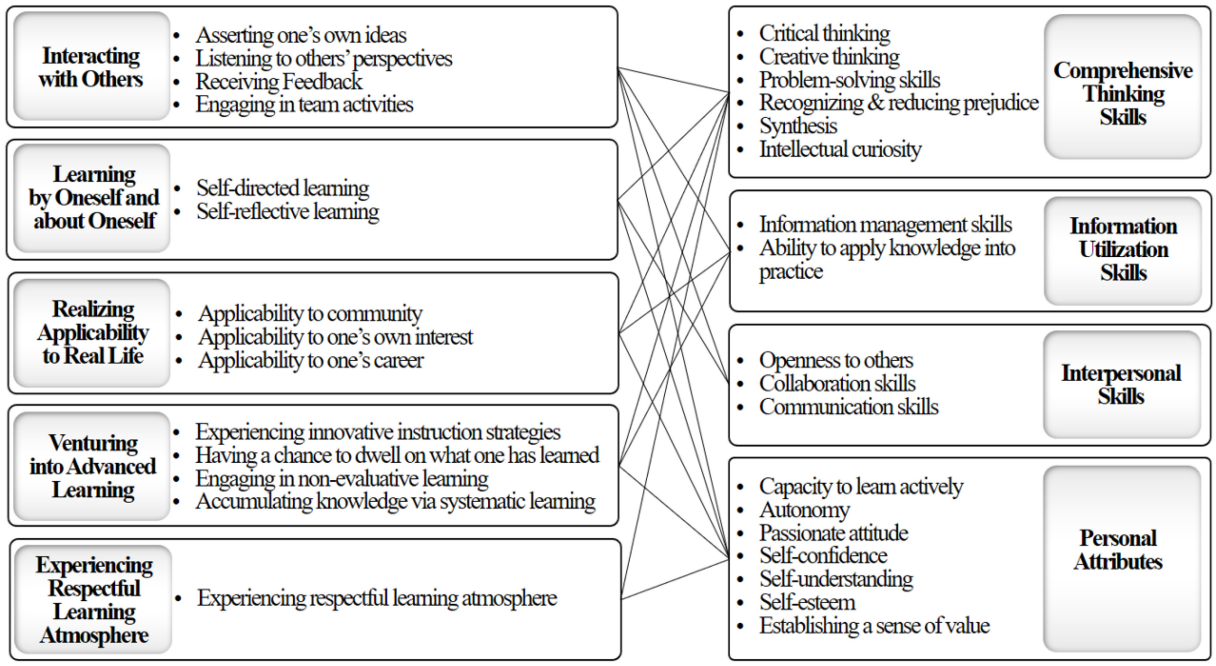


FIGURE 1
Themes related to significant learning experiences and clusters of associated generic skills.

Learning was also experienced as being "fun," "enjoyable," and "attractive" when it was self-directed as a result of students feeling "no pressure." Through self-directed learning, one student reported gaining "accurate assessment of the depth of her knowledge" and also the areas of growth which was a significant experience. Other aspects that made the self-directed learning experience significant was feeling the achievement, leading to in-depth learning, and being able to construct one's own knowledge and internalize knowledge. Also, self-directed learning was reported to lead to active communication with other learners. From these experiences, students were experiencing autonomy, a passionate attitude toward learning, and the capacity to learn actively. Also, intellectual curiosity and communication skills were reported in some students' experiences.

Self-reflective learning

Self-reflective learning was a powerful experience that allowed students to explore oneself and gain self-esteem, confidence, and understanding of oneself. While some self-reflective learning naturally occurred in the process of engaging in discussions with others or doing a written assignment, some were intentionally prompted in the context of specific courses, such as Art Therapy or Creating a Happy Family.

> Among the courses, the courses that made me focus on "myself" were especially meaningful. This is because I had more time to think quietly at home due to the COVID-19, and thinking about my future life naturally leads to interest in and exploration of myself. . . In order to create a family in the future that is full of happiness, the first priority was to have a deep insight and understanding of myself. There was an impressive activity, which was about writing my own 300 strengths. In the process, I was able to find my strengths even in trivial daily things. I was able to spend time exploring myself while fully concentrating on the task. I was able to define myself to some extent by thinking about my actions, tone of voice, personality, and values, and I was able to become more aware of my existence as a being (Student 11).

Students found self-reflective learning as being significant because they could find their own strengths, discover their true selves and accept them, and redefine their values. One student reported being able to produce "learning outcome that contain my identity" as a result of self-reflective learning. One student reported experiencing the change of behavior through reflecting on her own thoughts and discovering new and interests. From these experiences, several personal attributes were identified. Particularly, students reported enhancing self-understanding, self-esteem, and establishing a sense of value.

Realizing applicability to real life

Learning became meaningful when it was related to the real-life situations, whether directly to one's own life or to the community with which one is affiliated. Applicable learning occurred in the classroom and also through extracurricular activities. Categories of the theme include applicability to the community, to one's own interest, and to one's career.

Applicability to the community

Learning that was applicable to real-life situations in the societal level was meaningful to students. Because students could relate the learning material to the situations and people they recognized. One student recalled an experience of developing an educational program for school-aged students from a dual-income family and reported that the experience was meaningful because she was also "in the same situation, and there are actually a lot of people with dual-income parents." Such relation with the real situation encouraged her to participate more actively and be more passionate about the program. One student participating in a club activity reported that she was able to apply the learning theories that she had learned in class to systematically plan for the program.

> I believe that knowing how to apply knowledge to society is having a true learning experience. In college, meaningful learning experiences can happen through club activities and extracurricular activities (contests and volunteering). In my case, as a leader of a club, I had to search for activities the club would do over the course of a year. Education planning was the first step in any educational activity. I was able to plan activities more systematically by applying the educational model and learning theories that I had learned in the Department of Education. . . There were times when what I have planned did not work in reality. I realized that I had planned with certain stereotype, thinking "It will work," and also realized that I had to take into account various factors such as the participants, the characteristics of the participants, and the environmental factors with an open mind. This experience made me feel the necessity of the major classes, and my experiences helped me understand the content materials better (Student 10).

Other students also reported that being able to apply the learning contents to real-life situations allowed them to recognize the necessity of learning, which resulted in increased motivation, enjoyment, and active participation. Moreover, students took the time to investigate their community, identify problems, and critically evaluate the issues to generate solutions that are feasible. In the learning experience that is applicable to the community, a generic skill that stands out was the ability to apply knowledge into practice and problem-solving. Also, students reported focusing on their capacity to learn actively.

Applicability to one's own interest

Learning that was for one's own personal interest was significant for students. The personal interests of students varied from establishing one's own teaching philosophy as a future educator to language learning for practical reasons. One student was learning a new language to go abroad and to communicate with foreign friends and found the experience to be significant because she had a clear goal for learning and she could monitor her own progress.

> As I visit France often, I naturally developed an interest in French. . . I had a desire to communicate more actively, so I started learning (French). . . French grammar was difficult, but the more I learned, the more I felt like I was putting together a puzzle, and every time I learned a new vocabulary, more French words appeared on street signs and trademarks. Also, I even felt euphoric as my clumsy pronunciation gradually began to sound plausible. It was a novel experience in which the joy of learning the French language was motivating in itself. . . I was experiencing first-hand the concepts of intrinsic and extrinsic motivation learned in the educational psychology class (Student 16).

Another student expressed how much he enjoyed the Human Anatomy course, which was quite difficult and boring for most other students because the course materials were closely related to his own interest.

> I enjoyed taking this class because I was interested in the functioning of the body. The reason I was interested in the functions of the body was that I have a herniated disk and I like to exercise and do weight training as a hobby. Through the anatomy class, I was able to learn about the human body from a functional point of view, such as how the muscles of the body work, what specific movements these muscles are attached to, and what sensations they induce. I was able to take the lectures without difficulty and in a fun way as I studied the theory and practice while thinking about my own experience of exercise (Student 1).

Since this experience of significant learning was for one's own interest, students gained the capacity to learn actively as well as the ability to apply knowledge into practice. Moreover, students found the learning experiences to be meaningful because they were engaging in self-evaluation, monitoring their progress based on the personal learning goals that they set for themselves.

Applicability to one's career

Students found that learning relevant to one's own career path or helpful to one's career decision-making was significant. When they learned knowledge that is practical to the career field of their choice, they experienced enhanced

interest and enjoyment. With practical experience of real-life assignments such as engaging in program development, students had a meaningful experience of understanding how their previous knowledge can be utilized and applied.

> There was a time when I had the goal of becoming a "Korean language teacher." The lectures on Korean Language and Literature. . . which I took at that time, were memorable. In both lectures, the instructor was an actual Korean language teacher, so I was able to acquire a lot of knowledge related to the field. . . In particular, in both lectures, there was an assignment to do a mock lecture and upload a video recording of the demonstration. It was very difficult, but I feel that the experience of actually planning an educational program and writing a lesson plan was special (Student 13).

Students also reported that they were able to reflect back on their career choices and establish their sense of value regarding their career. The generic skill most identifiable in this theme was the ability to apply knowledge into practice and to establish a sense of value.

Venturing into advanced learning

Students felt that they were engaging in significant learning when they recognized themselves advancing in the learning process. This theme includes gaining experiences, especially *via* innovative instruction strategies, learning *via* assignments, engaging in non-evaluative learning, and accumulating knowledge through systematic learning.

Experiencing innovative instruction strategies

Students described their experiences of engaging in problem-based learning (PBL) and flipped learning (FL) as being significant. Through these innovative instruction strategies, students could have indirect experience of the world while collaborating with colleagues to solve problems. Participating in PBL and FL led students to "achieve growth in short period of time" as well as "produce outcome through synthesis." Students also reported that they were able to reflect on themselves, face their own prejudice, and think more flexibly. One student reported that FL was a new experience that changed her attitude to learning and made her feel like "an intellectual." These experiences led to heightened motivation, interest, creativity, and critical thinking. As the experiences illustrate, students were engaging in the capacity to learn actively. Also, comprehensive thinking skills, such as synthesis and problem-solving skills were integrated into the experiences of PBL and FL.

> The most memorable moment was the Cultural Immersion Project, which was a project to experience a culture that I had never experienced before and write a report. I decided to experience the culture of a sexual minority group and had various experiences by participating in the queer parade. I had negative thoughts about sexual minorities for religious reasons, but through the project, I was able to experience the culture and look back on myself, challenging my prejudices. I think it was a significant learning experience because it was a time to break the biases I had (Student 28).

> The first significant learning experience I experienced when I came to university was FL. It was a change from the learning method in which knowledge was delivered in the classroom, reviewed at home, and prepped for the class, to a new learning method of watching a video at home, putting the knowledge in mind, and then having a discussion in the classroom. At the time, I felt like a great intellectual. There was significance from novelty itself to have discussions together under the premise that everyone had already learned the material (Student 6).

Having a chance to dwell on what one has learned

Working on assignments was reported to be an effective tool that advances learning because it gave students an opportunity to review and reorganize what they have learned. Engaging in assignments during or after class, whether as an individual or in groups, helped students to "think deeply," "gain insight," and "express their thoughts freely." Assignments allowed students to expand their horizons and engage in critical thinking. One student reported that engaging on experiments allowed him to advance into deeper learning by helping him understand the theories better, make creative implications, and contain knowledge for a longer time. One student stated that her assignments required a lot of time and effort but it was meaningful because it required new ways of thinking. Another student said she was able to establish her value system and her own identity while engaging on reflective assignments.

> Personally, the part of the course that made me experience significant learning at university was the result of a task I worked on by myself, and not a test or one-way lecture. The task had more specific goals and details compared to other assignments, it was a topic I had never thought about, and it required a whole new way of thinking. The task was not easy and I had to invest a lot of time and effort to contemplate and complete it, and I think that was what made the experience significant (Student 7).

Through assignments that focused on the content material, students utilized comprehensive thinking skills such as critical thinking and creative thinking. Assignments that focused on self-reflection of the students, they engaged in the establishment of a sense of value.

Engaging in non-evaluative learning

Students found that learning that was not for exams meant more to them. Since their learning was not geared toward evaluation, they were "free to learn what (they) preferred" and found the motivation and enjoyment for learning. One student stated that he had found the meaning of learning and was able to self-evaluate his own progress. Also, another student reported that non-evaluative learning allowed him to express his thoughts rather than deliver information, which was a very meaningful experience. The key skill identified in this category was self-evaluation and autonomy.

> I did have experience of having discussion classes in middle school and high school. However, the difference here was that the class back then already had the right answer, and it was more like testing how well you talked about it. I had an experience of discussing climate change in the science class, and the arguments and the objections were already too obvious, and even before the discussions started, the process was already set and we just had to follow it. As long as you say the rights things in the right way, it was guaranteed that you would receive a perfect score on the performance evaluation. But when I came to college, the discussions in the classes were different... There were opportunities and time for me to say my own thoughts. Unlike high school, where I had been studying to receive good performance evaluation or test scores, these classes really gave me a chance to freely express my thoughts on given topic... I think having opportunity to expand my own thoughts is an experience of true learning that truly develops students' abilities (Student 30).

Accumulating knowledge *via* systematic learning

Learning through systematically designed courses and curricula allowed students to have significant experience because they could accumulate knowledge in a meaningful way. They were not simply inputting information, but were engaging in continuous inquiries and exploration. Since they had ample opportunities and course materials to fully grasp the knowledge, they maintained their intellectual curiosity, discovering new ways to utilize the knowledge. Hence the generic skills within these learning experiences include intellectual curiosity and information management.

> In order to answer these kinds of questions, it was necessary to first learn and understand the theory sufficiently, so rather than simply attending a class, I had to review it repeatedly and continuously explore how such a theoretical conclusion was reached. In addition, I investigated the exceptions to the theory, whether these theories are still valid, and if they are valid, how they can be used in our daily life... I started to become more interested and felt less vague about the theory through the process of learning the theory first, applying it to various situations, and rethinking the theory... In addition, when learning various theories, I first sequentially build up knowledge about one theory, and while learning the next theory, learn about the similarities and differences between the theories... I felt that the concepts were systematically accumulated, and I was able to experience significant learning (Student 22).

Experiencing respectful learning atmosphere

Seven students identified their significant learning experience as feeling respected in their learning environment. Such a respectful climate was sometimes created by the instructor intentionally and effortfully, sometimes by all the members of a class, and sometimes even by a title by which the instructor addressed each student. No matter how the atmosphere was created, students experiencing it reported feeling warm and welcomed. The learning climate allowed students to feel safe, leading the students to stir up the courage to express their opinion more actively in class.

> ... I was able to freely express my thoughts without pressure. I like to share my experiences and talk to others, but in most classes, I had to do presentations or discussions in order to get a score, and even when I have discussions, they were perfunctory most of the time, so I really liked the atmosphere of the class where I was able to freely share myself (Student 25).

Students also felt accepted which lessened their burden to perform. One student reported "being free to speak" in such an atmosphere, and another student wrote that the environment made it possible for him to ask questions and offer his ideas in class. Because students were able to take risk and challenge themselves in a respectful environment, they were more passionate and more inquisitive. They communicated more with one another as well. One student explicitly stated that she gained self-confidence and courage in the respectful atmosphere.

> I think that it was the role and attitude of the instructor that made the learner-centered teaching method even more possible. As a facilitator, the professor gave feedback on each stage of development and actively supported me to come up with better ideas. And although the knowledge that each team had formed was different, she respectfully accepted the diversity and complexity of everyone. From my point of view, as I received encouragement and praise, the class participation time became more interesting and I became more curious about the activities in each step of the course. Also, since I knew that all of my suggestions would be respected, I gained the confidence to offer my opinions more courageously (Student 29).

A typical generic skill identified by students, whose significant learning experience was experiencing a respectful learning atmosphere, was self-confidence and passionate attitude toward learning. Intellectual curiosity was also identified within the experiences.

Discussion

Findings and implications

This study examined the significant learning experiences identified by university students and analyzed generic skills entailed in those experiences. For this analysis, we collected 33 essays written by students studying at a university located in Seoul, South Korea, and applied a document analysis approach. In the current study, five themes about significant learning experiences were derived and systematized. The themes were interacting with others, learning by oneself and about oneself, realizing applicability in real-life, venturing into advanced learning, and experiencing a respectful learning atmosphere. These themes were reflected in the educational experiences of Korean students and were embedded in the Korean educational culture.

First, most of the students identified learning experiences in which they were engaged with one another as being significant. In the findings, students pointed out that meaningful learning for them involved freely speaking about their thoughts and opinions in discussions or debates, listening to other students' opinions and receiving feedback from lecturers and peers. This finding is in alignment with what Fink (2013) described as *human dimension*, one of the elements for significant learning experience indicating learning through actively interacting with others. However, this finding can be discussed further in the context of Korean education. Demanding high school curricula and a high level of standards in Korean high school systems might affect the way high school classes are taught and learned, and teachers and students cannot spend much class time having discussions. In addition, instead of education that broadens various thoughts, students are used to learning to get answers and cram knowledge (Lee, 2008; Kim and Cho, 2014). Therefore, for Korean university students, significant learning experiences are the experiences of constructing knowledge in their own way and freely generated ideas. Interaction with lecturers and peers led students to be continuously provided with information about their learning process, which means disclosing their own ideas and receiving others' opinions and feedbacks (Lee and Kim, 2008).

Second, students perceived the experiences of self-directed and self-reflective learning as significant learning experiences. This is because university students are adult learners. Previous studies on adult learners (Knowles, 1999; Merrian and Cafarella, 1999) assumed that they have different characteristics from elementary and secondary school students. These studies consistently argue that adult learners have diverse and rich experiences and various levels of self-directed tendencies. Self-reflective learning is a significant learning experience for students to explore themselves and gain self-esteem, confidence, and understanding of themselves. In the Korean high school education, there are not many opportunities to fully reflect on oneself as the education mainly focuses on preparing for the university entrance examination (Yoon, 2013). Therefore, it can be inferred that students attached significant meaning to these experiences of self-understanding and reflection.

Third, learning experiences related to the real-life situations are considered powerful learning experiences. Students' learning experiences are continuing to influence each other in succession, and at the time of experience, they are affected by interactions with the environment surrounding individuals (Dewey, 1916). Students perceived the learning as significant when they did activities to solve problems in the community through club activities. Students also recognized that learning experiences related to their careers and learning experiences that have confidence in their careers were significant. In South Korea, students often choose to go to a college based on the scores they receive on their college entrance exam without deeply considering their preference for a major, their aptitude, and career trajectory (Kim and Moon, 2005). Therefore, when students engage in an experience during their university years that transforms this uncertainty into conviction, students can discover the meaning of learning (Han et al., 2021).

Fourth, students described their experiences in courses using innovative instructional strategies such as FL and PBL as being meaningful. It is consistent with the results of previous studies that FL classes have higher learning effects and higher class satisfaction compared to traditional lecture-style classes (Moravec et al., 2010; Deslauriers et al., 2011; Touchton, 2015) and that students and teachers can become better problem-solvers *via* PBL (Mettas and Constantinou, 2008). Students perceived that the class experience of receiving intellectual stimulation and feeling a sense of accomplishment through a little tricky and challenging task was meaningful. Ramsden

(1991) suggested that clear goals and intellectual challenges in the class as being principles of good instruction. Brophy (2000) also emphasized that students should be provided with sufficient opportunities to learn higher-order thinking.

Fifth, students identified their significant learning experience as involving a respectful learning environment. Students pointed out that they could express their opinions more actively and freely in a comfortable atmosphere, and they also felt recognized. It is supported by previous studies (Lee and Kim, 2008; Han et al., 2021) that the instructor's passion or attitude affects the students' learning experiences.

From the aforementioned significant learning experiences, the current study also identified 18 generic skills categorized into four clusters: comprehensive thinking skills, information utilization skills, interpersonal skills, and personal attributes. First cluster was comprehensive thinking skills. Comprehensive thinking skills is one of the core competencies measured by K-CESA and is defined as the ability to engage in higher-order thinking to recognize and solve problems, make sound judgments, and present plausible solutions (Hwang et al., 2016). In the present study, skills such as critical thinking, creative thinking, problem-solving skills, recognizing and reducing prejudice, synthesis, and intellectual curiosity were grouped into the first cluster. Comprehensive thinking skills were involved with all five themes of significant learning experiences. That is, when students were engaged in learning experience that they identify as being significant, they were engaging in various higher-order thinking skills, indicating that generic thinking skills may be fostered in various contexts. As Kember et al. (2007) posited, generic skills can be developed through various learning experiences. Thus, while it is important to design a curriculum specifically to enhance certain generic skills, it is also important to build learning activities and atmosphere that students would recognize as being meaningful to them in order to facilitate various thinking skills.

Second cluster was information utilization skills, which include information management skills and ability to apply knowledge into practice. This cluster was formed based on the category of K-CESA called resource information technology utilization competency, defined as collecting, analyzing, and applying various resources, information, and technology (Hwang et al., 2016). In the current study, students only mentioned accessing and applying information and did not address other resources and technology, so the cluster only referred to information. Information utilization skills were related to learning by interacting with others, realizing applicability to real-life, and venturing into advanced learning. Kang et al. (2010) found information management and knowledge utilization, among others, as competencies in the cognitive domain that are necessary for learners to face the complex challenges of the modern world. In order to assist students to manage various information they are obtaining and to foster abilities to apply their knowledge, linking learning material to real-life situations, especially through innovative instructions such as PBL would be helpful. Also, it should be noted that the experience of asserting one's own ideas in discussions and debates involved processing and organizing knowledge and opinion. To enhance such information management skills, it would be important to provide opportunities for students to express themselves.

Third, interpersonal skills included openness to others, collaboration skills, and communication skills. K-CESA differentiates communication competency and interpersonal competency, the former including traditional reading, writing, listening, speaking skills and the latter including emotional relatedness, collaboration, mediation, leadership, and understanding of organization (Hwang et al., 2016). However, Braun and Leidner (2009) referred to social competence as encompassing cooperative competence and communication competence, and communication competence was referred to as verbally sharing ideas in discussions. In the current study, Braun and Leidner's (2009) concept of social competence is adopted, and the name interpersonal skills was used as an umbrella term to include collaboration skills and communication skills as well as openness to others, which was another skill addressed in the essays. In the findings, collaboration skills and communication skills were mostly reported along with the experience of engaging in team activities. It is interesting to note that students who reported having discussions as being significant did not report collaborative or communication skills as being associated with the experience while identifying openness to others as an integral part. This could be because, in team activities, the teammates strive for a common goal and the sense of fellowship could be the key factor in relation to the generic skills. This could be a meaningful finding for instruction design. From the perspective of significant learning, learning together with others is an important part of promoting significant learning (Fink, 2013; Evans et al., 2016), and team activities and discussions may both be used in the instruction. However, with a closer look, these two activities are associated with different generic skills; team activity with collaborative and communication skills and discussion with openness to others. Although generic skills are fostered in various learning contexts (Kember et al., 2007), cultivating specific generic skills may require tailored learning activities or environments.

Finally, the students' reports derived various personal attributes, such as self-confidence, autonomy, and passionate attitude as part of their significant learning experiences. Such personal attributes are in alignment with Braun and Leidner's (2009) notion of personal competence, indicating "positive attitude toward learning and development of the self" (p. 301). In the findings of the present study, all five themes of significant learning experiences were related to personal attributes. These personal attributes may be understood in relation to what Fink (2013) referred to as *caring* and *learning how to learn*. Caring refers to changing feelings and interests related to learning and increasing motivation, and learning how to learn is related to

being more self-directed (Fink, 2013). When these elements are integrated to create significant learning experiences, generic skills related to personal attributes may also be affected. Most of the excerpts also indicated that the experiences were identified as being significant because students had the opportunity to foster personal attributes. Therefore, learning activities should not only focus on content materials but also on developing personal attributes of students.

Previous studies found that generic skills are not effectively fostered through specific courses designed for generic skills (Hattie et al., 1996) or through a specific pedagogical practice (Virtanen and Tynjälä, 2019), but through various learning experiences and teaching methods (Kember et al., 2007; Virtanen and Tynjälä, 2019). The findings of the current study also illustrate that various generic skills are identified across different types of significant learning experiences. Nonetheless, certain generic skills are addressed more often with specific learning experiences, indicating that more research is needed to identify relationships between significant learning experiences and generic skills.

Limitations and direction for future studies

There are some limitations to this study. First, the essay data were collected mainly from students who are majoring in education. Students from different majors could offer varying perspectives on significant learning experiences from which different generic skills may be induced. Thus, future study is required to encompass students from different majors. Second, the data were collected from a 4-year university in South Korea, but there are also 2 and 3-year vocational universities in South Korea that focus on practical, field-specific knowledge and skills. Studies are needed to investigate the experiences of students from vocational universities to discover similarities and differences in students' experiences among different types of universities. Third, the current research focused on gaining insight *via* a qualitative approach and found certain generic skills to be associated with significant learning experiences. However, in order to fully understand the mechanism between significant learning experience and generic skills, an experimental or longitudinal approach may be useful to identify any directional and causal relationships between the two. Despite these limitations, the current study contributes to the understanding of students' significant learning experiences and generic skills by providing themes derived from the descriptions of students' direct learning experiences.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethics review and approval was not required as per local legislation and institutional requirements. The participants provided their written informed consent to participate in the study.

Author contributions

AL and SJL contributed to conception and design of the study. SJL organized the database and method, and wrote sections of the manuscript. AL wrote the first draft of the manuscript. Both authors contributed to manuscript revision, read, performed the document analysis, and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Agra, G., Formiga, N. S., de Oliveira, P. S., Costa, M. M. L., Fernandes, M. G. M., and Nobrega, M. M. L. (2019). Analysis of the concept of meaningful learning in light of the Ausubel's theory. *Rev. Bras. Enferm.* 72, 258–265. doi: 10.1590/0034-7167-2017-0691

Ausubel, D. P. (1963). *The Psychology of Meaningful Verbal Learning: An Introduction to School Learning*. New York, NY: Grune and Stratton.

Ausubel, D. P. (2000). *The Acquisition and Retention of Knowledge: A Cognitive View*. Boston: Kluwer Academic Publishers. doi: 10.1007/978-94-015-9454-7

Badcock, P. B. T., Pattison, P. E., and Harris, K. L. (2010). Developing generic skills through university study: A study of arts, science and engineering in Australia. *High. Educ.* 60, 441–458. doi: 10.1007/s10734-010-9308-8

Bae, S. H., and Hwang, S. J. (2021). Student engagement and institutional commitment and predictors in online learning during the COVID-19 pandemic: A latent profile analysis. *Korean J. Educ. Res.* 59, 279–308. doi: 10.30916/KERA.59.1.279

Ban, S. J. (2016). The decline of school age population, government sponsored work expenditure for universities, and the regional disparity of higher education. *J. Educ. Res.* 14, 213–242.

Barrie, S. (2006). Understanding what we mean by the generic attributes of graduates. *High. Educ.* 51, 215–241. doi: 10.1007/s10734-004-6384-7

Bath, D., Smith, C., Stein, S., and Swann, R. (2004). Beyond mapping and embedding graduate attributes: Bringing together quality assurance and action learning to create a validated and living curriculum. *High. Educ. Res. Dev.* 23, 313–328. doi: 10.1080/0729436042000235427

Blustein, D. L. (2019). *The Importance of Work in an Age of Uncertainty: The Eroding Work Experience in America*. New York, NY: Oxford University Press. doi: 10.1093/oso/9780190213701.001.0001

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qual. Res. J.* 9, 27–40. doi: 10.3316/QRJ0902027

Bratianu, C., and Vatamanescu, E.-M. (2017). Students' perception on developing conceptual generic skills for business: A knowledge-based approach. *VINE J. Inform. Knowl. Manag. Syst.* 47, 490–505. doi: 10.1108/VJIKMS-11-2016-0065

Braun, E., and Leidner, B. (2009). Academic course evaluation: Theoretical and empirical distinctions between self-rated gain in competences and satisfaction with teaching behavior. *Europ. Psychol.* 14, 297–306. doi: 10.1027/1016-9040.14.4.297

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Brophy, J. (2000). *Teaching. Educational Practices Series I*. Switzerland: International Bureau of Education.

Byoun, S. Y. (2018). An exploration on the notion of excellence in the Korean undergraduate education: Based on the performance indicators of ACE universities. *J. Polit. Educ.* 25, 77–106. doi: 10.52183/KSPE.2018.25.3.77

Choi, J. H. (2020). Improvements through extracurricular core competency mapping analysis: Focusing on K university. *J. Learn. Centered Curr. Instruc.* 20, 1071–1100. doi: 10.22251/jlcci.2020.20.18.1071

Deslauriers, L., Schelew, E., and Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science* 332, 862–864. doi: 10.1126/science.1201783

Dewey, J. (1916). *Democracy and Education*. New York, NY: The Macmillan Company.

Drummond, I., Nixon, I., and Wiltshire, J. (1998). Personal transferable skills in higher education: The problems of implementing good practice. *Qual. Assur. Educ.* 6, 19–27. doi: 10.1108/09684889810200359

Evans, H. G., Heyl, D. L., and Liggit, P. (2016). Team-based learning, faculty research, and grant writing bring significant learning experiences to an undergraduate biochemistry laboratory course. *J. Chem. Educ.* 93, 1027–1033. doi: 10.1021/acs.jchemed.5b00854

Fink, L. D. (2003). *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses*. San Francisco, CA: Jossey-Bass.

Fink, L. D. (2013). *Creating Significant Learning Experiences, Revised and Updated: An Integrated Approach to Designing College Courses*. San Francisco, CA: John Wiley and Sons.

Guba, E. G., and Lincoln, Y. S. (1985). *Naturalistic Inquiry*. Beverly Hills, CA: Sage Publication.

Han, S., Bae, S. H., and Hwang, S. J. (2021). "Exploring the perception of college students' significant learning experiences," in *2021 Summer Conference of the Korean Educational Administration Association*, (New York, NY: The Macmillan Company), 16–31.

Han, S., and Hwang, S. (2021). A study on the perception of 'significant learning experiences' in A university. *Cult. Converg.* 43, 1273–1288. doi: 10.33645/cnc.2021.12.43.12.1273

Hattie, J., Biggs, J., and Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Rev. Educ. Res.* 66, 99–136. doi: 10.3102/00346543066002099

Hwang, J., Kim, H., and Song, O. (2016). The relationship between course grades and the K-CESA core competencies for engineering students. *J. Engin. Educ. Res.* 19, 35–46. doi: 10.18108/jeer.2016.19.4.35

Jin, M., Sohn, Y., and Chu, H. (2011). A study on development plan of K-CESA for college education assessment. *J. Educ. Administr.* 29, 461–486.

Joyce, B., Weil, M., and Calhoun, E. (2000). *Models of Teaching*. United States, MA: Allyn and Bacon.

Kang, M., Heo, H., Jo, I., Shin, J., and Seo, J. (2010). Developing an educational performance indicator for new millennium learners. *J. Res. Technol. Educ.* 43, 157–170. doi: 10.1080/15391523.2010.10782567

Karalis, T., and Raikou, N. (2020). Teaching at the times of COVID-19: Inferences and implications for higher education pedagogy. *Int. J. Acad. Res. Bus. Soc. Sci.* 10, 479–493. doi: 10.6007/IJARBSS/v10-i5/7219

Kember, D., and Leung, D. Y. P. (2005). The influence of active learning experiences on the development of graduate capabilities. *Stud. High. Educ.* 30, 155–170. doi: 10.1080/03075070500043127

Kember, D., Leung, D. Y. P., and Ma, R. S. F. (2007). Characterizing learning environments capable of nurturing generic capabilities in higher education. *Res. High. Educ.* 48, 609–632. doi: 10.1007/s11162-006-9037-0

Kim, H. S. (2022). Effect of declining school age population on university finances: Focusing on financial indicators and models for judging business performance. *J. Bus. Educ.* 36, 1–21.

Kim, S. J., and Moon, H. K. (2005). Survey on the actual conditions and consciousness of freshmen in 2004. *Creativ. Devel. Stud.* 8, 1–48.

Kim, Y., and Cho, Y. H. (2014). The second leap toward "world class" education in Korea. *Asia Pacif. Educ. Res.* 23, 783–794. doi: 10.1007/s40299-013-0144-3

Knowles, M. (1999). *Definite Classics on Adult Education*. New York, NY: Association Press.

Lee, C. (2008). "Education in the Republic of Korea: Approaches, achievements, and current challenges," in *An African exploration of the East Asian Education Experience*, eds B. Fredriksen and T. J. Peng (Paris: World Bank), 155–217.

Lee, E., and Kim, H. (2008). Defining a "good instruction": The qualitative study of undergraduate students in Korea. *J. Korean Educ. Idea* 22, 123–146. doi: 10.17283/jkedi.2008.22.1.123

Lee, M. (2017). The meaning of education in the core competencies and practice strategies of the universities participating in the ACE project. *Asian J. Educ.* 18, 339–364. doi: 10.15753/aje.2017.06.18.2.339

Leung, D. Y. P., and Kember, D. (2005). The influence of the part time study experience on the development of generic capabilities. *J. Further High. Educ.* 29, 91–101. doi: 10.1080/03098770500103101

Levine, L. E., Fallahi, C. R., Nicoll-Senft, J. M., Tessier, J. T., Watson, C. L., and Wood, R. M. (2008). Creating significant learning experiences across disciplines. *College Teach.* 56, 247–254. doi: 10.3200/CTCH.56.4.247-254

Merrian, S. B., and Cafarella, R. S. (1999). *Learning in Adulthood: A Comprehensive Guide*. San Francisco, CA: Jossey-Bass.

Mettas, A. C., and Constantinou, C. C. (2008). The technology fair: A project-based learning approach for enhancing problem solving skills and interest in design and technology education. *Int. J. Technol. Design Educ.* 18, 79–100.

Moravec, M., Williams, A., Aguilar-Roca, N., and O'Dowd, D. K. (2010). Learn before lecture: A strategy that improves learning outcomes in a large introductory biology class. *CBE Life Sci. Educ.* 9, 473–481. doi: 10.1187/cbe.10-04-0063

OECD (2005). *The Definition and Selection of Key Competencies: Executive Summary*. Paris: OECD.

OECD (2018). *The Future of Education and Skills: Education 2030*. Paris: OECD.

Park, C. N., and Chung, W. H. (2017). Study on the effects of non-subject educational programs influencing the core competency of university students: Focusing on the case of K university. *Korean J. General Educ.* 11, 39–71.

Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The course experience questionnaire. *Stud. High. Educ.* 16, 129–150. doi: 10.1080/03075079112331382944

Sanchez, S., Park, N., and Fedorek, B. (2020). Creating significant learning outcomes in criminal justice courses: A classroom activity to encourage reflexivity and empathetic thinking. *J. Crim. Justice Educ.* 31, 267–282. doi: 10.1080/10511253.2019.1705364

Saulnier, B. M. (2003). Creating significant learning experiences in systems analysis and design: Towards a service learning paradigm. *Inform. Syst. Educ. J.* 4, 1–7.

Shin, M., Choi, J., and Lee, H. (2021). Characteristics based on the level of their core competency: A case of local national K university. *Cult. Converg.* 43, 95–114.

Strange, C. C., and Banning, J. H. (2015). *Designing for Learning: Creating Campus Environments for Student Success*. San Francisco, CA: John Wiley and Sons.

Touchton, M. (2015). Flipping the classroom and student performance in advanced statistics: Evidence from a quasi-experiment. *J. Polit. Sci. Educ.* 11, 28–44. doi: 10.1080/15512169.2014.985105

Trudeau, D., and Kruse, T. P. (2014). Creating significant learning experiences through civic engagement: Practical strategies for community-engaged pedagogy. *J. Publ. Scholarsh. High. Educ.* 4, 12–30.

Virtanen, A., and Tynjälä, P. (2019). Factors explaining the learning of generic skills: A study of university students' experiences. *Teach. High. Educ.* 24, 880–894. doi: 10.1080/13562517.2018.1515195

Yoon, Y. M. (2013). The relationship among extracurricular club activities with adaptation to campus life and career preparation behavior of college students. *J. Commun. Devel. Assoc.* 38, 21–30.

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 13 September 2022
DOI 10.3389/feduc.2022.909968

OPEN ACCESS

EDITED BY
Heidi Hyytinen,
University of Helsinki Centre for University Teaching and Learning, Finland

REVIEWED BY
Svanborg Rannveig Jónsdóttir,
University of Iceland, Iceland
Milla Räisänen,
University of Helsinki, Finland

*CORRESPONDENCE
Agnese Slišāne
agnese.slisane@lu.lv

SPECIALTY SECTION
This article was submitted to Higher Education, a section of the journal Frontiers in Education

RECEIVED 31 March 2022
ACCEPTED 04 July 2022
PUBLISHED 13 September 2022

CITATION
Slišāne A, Lāma G and Rubene Z (2022) How is entrepreneurship as generic and professional competences diverse? Some reflections on the evaluations of university students' generic competences (students of education and bioeconomics). *Front. Educ.* 7:909968. doi: 10.3389/feduc.2022.909968

How is entrepreneurship as generic and professional competences diverse? Some reflections on the evaluations of university students' generic competences (students of education and bioeconomics)

Agnese Slišāne*, Gatis Lāma and Zanda Rubene

Faculty of Education, Psychology and Art, University of Latvia, Riga, Latvia

Generic competences have an interdisciplinary nature, which indicates their usability in different disciplines, situations, and contexts in the performance of different tasks. Generic competencies are thus considered from two perspectives, daily life and professional activity, that are equally important, implying that generic competences are necessary for individuals to successfully adapt to change and live meaningful and productive lives. Entrepreneurship competences can be observed from two perspectives: generic competencies viewed from the perspective of the individual's personal experience and professional competencies viewed from the perspective of the individual's professional experience. In this article, it will be observed from both perspectives to see its performance in diverse contexts and to clarify distinctions between these contexts. The present study aimed to shed light on how specific university study disciplines with a professional focus (educational sciences and bioeconomics) support the development of a specific generic competence (entrepreneurship competencies). The Specific Research Questions of This Article Are: (1) What Entrepreneurship Competences Emerge Among Latvian Bioeconomics and Educational Science Students? (2) How Do Entrepreneurship Competences Differ Between Bioeconomics and Educational Science Students? (3) How Are Entrepreneurship Competences Correlated With Each Other? Data for the study were gathered by using the online survey platform QuestionPro. The questionnaire was filled in by 135 students, of whom 82 were from the field of educational sciences and 53 from the field of bioeconomics. The study presents a comparison of entrepreneurship competence's self-assessments of bachelor's, master's, and doctoral students of bioeconomics and educational sciences. Despite the fact that entrepreneurship is more linked to economics, the results show that, in two out of three main areas of entrepreneurship competences, students of educational sciences self-assessed their entrepreneurship competences as higher than students of bioeconomics.

KEYWORDS

entrepreneurial skill, generic skills and competences, professional skills, education, pedagogy

Introduction

Globalization can be defined as an economic, social, political, cultural, and territorial integration process (Arrighi, 2005), resulting in a change in the knowledge, skills, and attitudes needed to carry out entrepreneurial work, such as the ability to handle digital technologies, knowledge of global processes, and an open attitude toward the cultures of other nations. Differences between the requirements of education and the labor market were the main reason for the development of competences (Grant et al., 1979), contributing to the development of the competence profiles of professional associations, which included requirements applicable to candidates in a particular profession. In this study, entrepreneurial skills were researched from the perspectives of both generic and professional competences to see their performance in diverse contexts and to clarify distinctions between the contexts.

Generic competences have an interdisciplinary nature, which indicates their usability in different disciplines, situations, and contexts in the performance of different tasks (Florea, 2014; Pârvu et al., 2014; Economou, 2016). Generic competences are considered from two perspectives, daily life and professional activity, both of which are equally important (Direito et al., 2014; Larraz et al., 2017; Sá and Serpa, 2018) and indicate that generic skills are necessary for individuals to successfully adapt to change and live meaningful and productive lives (UNESCO, 2016). In the European Higher Education Area, generic competences are described as the skills, knowledge, and attitudes acquired in one situation or field that can be used in other situations, areas, or types of occupations and include communication skills, self-control skills, and problem-solving skills (Akadēmiskās Informācijas Centrs, 2017). (UNESCO, 2016) divides generic competences into six areas:

1. Critical and innovative thinking;
2. Interpersonal skills (e.g., the ability to present, communicate, organize, work in a team, etc.);
3. Intrapersonal skills (e.g., self-discipline, enthusiasm, perseverance, self-motivation, etc.);
4. Global citizenship (e.g., tolerance, openness, respect for diversity, intercultural understanding);
5. Media and information literacy (e.g., the ability to find and access information, analyse and evaluate media content, etc.); and
6. Other skills (this field was created so that researchers could include competences such as physical health or religious values that may not fall into any of the other areas).

Another way to allocate generic skills was suggested by the project "Assessment of competences of students in higher education and dynamics of their development during the study period," where the following generic competences were listed: research, entrepreneurial skills, innovation, global competences, civic competences, and digital competences (Rubene et al., 2021). These competences emphasize critical thinking, creativity, initiative-taking, problem-solving, risk assessment, decision-making, and the constructive management of emotions (Pepper, 2011).

Professional competences are related to motivation, intelligence, professional performance, and vocational education, which are characterized as skills to interact effectively with one's (social and intellectual) environment and as a result of intensive and continuous learning, which is impossible to implement without the desire to acquire a certain level of professional skills.

It is a general, integrated, and internationalized skill to ensure sustainable, effective performance in a particular professional field, job, organizational context, or task-related situation. It must be also stressed that professional competences are a coordinated set of knowledge, skills, and attitudes that can be used to address real professional situations (Mulder, 2014). Given the changing environment, professional competences are inherently unsustainable and need to be developed consistently in the context in which they should be applied (Epstein and Hundert, 2002).

Entrepreneurship competences have normally been researched from the business perspective since traditionally they come from the business area. However, since 2006, entrepreneurship competences have been highlighted as generic skills that are needed in all areas of life (Bacigalupo et al., 2016). Although research has been done on how students of educational sciences self-assess their entrepreneurial skills (Slišāne et al., 2021b), a discipline like bioeconomics that is relatively related to entrepreneurship has not received the attention it deserves. This is despite the fact that entrepreneurship competences have been recognized as an essential part of bioeconomics students' professional development as the related skills are directly used in a professional context (Kuckertz et al., 2020).

Theoretical framework

Entrepreneurship as a generic competence

In 2015, an extensive overview of entrepreneurship competences was created, identifying and comparing different theoretical approaches from both academic and non-academic backgrounds. From the study, it can be understood that although entrepreneurship competences were originally an economic phenomenon and its conceptualization was strongly dependent on the economic aspects of entrepreneurship, the concepts of entrepreneurship and entrepreneurial activities have since developed beyond their original economic domain

(Komarkova et al., 2015). The authors of *EntreComp: The European Entrepreneurship Competence Framework* (Bacigalupo et al., 2016) reflect the dimensions of entrepreneurial skills that foster innovation, creativity, and self-determination. Entrepreneurship as generic competences is seen as distinct to turn research and education data into economic value and, more broadly, to create social value (Slišāne et al., 2021a) in a personal or a professional context.

Based on extensive baseline analysis (reviews and case studies), *EntreComp* defines entrepreneurial skills as generic competences as it covers all areas of life, from promoting personal development to active participation in society and (re-)entering the labor market as an employee or self-employed person, as well as start-ups (cultural, social, or commercial; Bacigalupo et al., 2016). Within the framework presented in *EntreComp*, entrepreneurial skills are described as basic generic competences applicable to individuals and groups, which include three competence areas and 15 dimensions (Bacigalupo et al., 2016).

The three competence areas presented in *EntreComp* are interconnected:

1. Ideas and opportunities: Problem-solving skills and creativity describe the ability to spot opportunities and critically assess them, find a solution that has added value to society/the market, and make strategic, ethical, long-term decisions based on a vision. This area includes five dimensions: spotting opportunities, creativity, vision, evaluation of ideas, and ethical and sustainable thinking.
2. Resources: The identification, mobilization, and efficient use of internal and external resources describe the ability to use one's strengths and opportunities to overcome failures and challenges and to mobilize financial and human resources to achieve goals and/or create value. This area includes five dimensions: the assessment of one's abilities, motivation and perseverance; mobilizing resource; financial and economic competences; communication; and human resources mobilization.
3. Into action: Initiative and action orientation describe the ability to show initiative, set goals, plan their achievement, evaluate risks, work and manage a team, evaluate results, and make improvements to achieve the highest possible result. This area includes five dimensions: initiative, planning, action in times of uncertainty, teamwork, and learning from experience.

Entrepreneurial skills are recognized as the key to the development and fulfillment of the individual, active citizenship, social inclusion, and employability in the knowledge society (European Parliament Council, 2006). The concept of the "new economy," which emphasizes the transition from "manual work" to "knowledge work," i.e., the need to work with information, can be defined very differently, but the role of information and communication technology (ICT) and the information field in economic processes is constantly emphasized (Neumark and Reed, 2004). Individuals should therefore develop competences to help them successfully enter the labor market, where competitiveness is determined by the ability to apply knowledge (Moretti, 2004; Abel and Gabe, 2011; Kalleberg, 2011; Rubin, 2012). These changes show a growing demand in the labor market for competent individuals who have entrepreneurial skills, as these are important for organizations/companies and are in demand in different positions in the labor market (Szafranski et al., 2017). Furthermore, the structure of entrepreneurial competences indicates skills that are useful not only in the labor market but also in other aspects of life (Komarkova et al., 2015).

Entrepreneurship as a professional competence

The European Union (European Parliament Council, 2006) defines entrepreneurial skills as an individual's ability to translate ideas into action, which includes creativity, innovation, and risk-taking, as well as planning and managing projects to achieve goals. Entrepreneurship competences promote individuals not only in their daily lives at home and in society but also at work, contributing to social or commercial activities. It involves an awareness of ethical values and that entrepreneurial skills are not only about the formation of a company but are also generic and professional competences that help an individual to be proactive, independent, and innovative in his or her personal life, as well as in the workplace (Luppi et al., 2019). Almost every classification of entrepreneurial skills features generic skills (Komarkova et al., 2015), which confirm the generic nature of these skills.

Professional competences include the knowledge and skills necessary for the performance of specific and general work in a particular profession or sector (Mulder, 2014). Professional competences also include one's attitude, which is the desire and motivation to achieve a specific result. Professional competences related to entrepreneurial skills can be classified into four groups: work-related knowledge; skills for work-related tasks; personal qualities that contribute to the achievement of work tasks; and sets of characteristics of the individual that help to achieve meta competences (sets of light skills and other individual qualities that tend to be associated with excellent performance in situations of difficulty, including flexibility, tolerance for ambiguity, ability to learn, reasoning and intuition, creativity, and analytical and problem-solving abilities) (Cheetham and Chivers, 1996, 1998).

According to the Dutch scholar Martin Mulder, and based on the research undertaken by international organizations, professional competences are formed of three complementary components: knowledge, skills, and

TABLE 1 Distinction—entrepreneurship as generic and professional competences.

	No.	Elements of entrepreneurial skills	Generic competences (viewed from the perspective of the individual's personal experience)	Professional competences (viewed from the perspective of the individual's professional experience)
Ideas and opportunities	1	Ability to spot opportunities	Competence to notice opportunities to achieve the goals of the social community, promote wellbeing, and realize offers beneficial to the personal budget.	Competence to notice opportunities in the workplace—prepare projects, attract funding, spot opportunities for professional development, etc.
	2	Creativity	Competence to innovate to improve personal life—housing, education, social environment, etc. Competence to solve problems creatively.	Competence to innovate in the process of work, organization, and product, providing added value in development.
	3	Vision	Competence to create and explain a long-term vision for life plans.	Competence to see professional development in the long-term—one's professional role in the organization and beyond.
	4	Ability to critically evaluate ideas	Competence to critically evaluate ideas that are important for personal life, making decisions, and implementing them as intended.	Competence to evaluate ideas in the context of work tasks, evaluating the profitability of the idea—the work invested against the possible benefit.
	5	Ethical and sustainable thinking	Competence in ethical decision-making about the environment in which one lives for equal treatment of diversity in society.	Competence to perform tasks and improve professionally, observing ethical principles, and considering sustainability aspects—environment, social equality, and cooperative management—when making decisions.
Recourses	6	Awareness and self-efficacy of skills	Competence to perform tasks effectively, organize daily life, and perform tasks for personal wellbeing.	Competence to evaluate strengths and capabilities in the performance of specific tasks to achieve a higher result, identify bottlenecks, and, if necessary, improve them for professional development.
	7	Motivation and perseverance	Competence to achieve the set goals/tasks, demonstrating motivation and perseverance.	Competence to achieve professional goals/tasks, looking for ways to achieve the set goals without giving up in case of failure.
	8	Ability to mobilize the necessary resources	Competence to mobilize resources to ensure domestic wellbeing and leisure facilities.	Competence to mobilize capital—product/service provision or financial resources—for the performance of professional duties.
	9	Financial and economic expertise	Competence to plan resources, be aware of income and expenditure flow, know different types of income, and have the ability to use financial literacy to improve living conditions.	Competence to predict changes in income level (workplace and/or sector), considering economic conditions in the country. Competence to plan finances and demonstrate knowledge about economic cycles within the organization—for example, organizing projects and attracting funding.
	10	Ability to mobilize and motivate human resources	Competence to mobilize and motivate human resources when help is needed—solving domestic problems, organizing personal events, etc.	Competence to mobilize human resources, evaluating the individual competences necessary for the performance of the work task. Competence to analyse the potential contribution of human resources to the performance of the work task and motivate resources to perform tasks.
Into action	11	Initiative	Competence to take the initiative and propose ideas.	Competence to propose, improve, and implement ideas (based on professional knowledge and work experience).
	12	Planning and management	Competence to plan personal time and activities and manage domestic events.	Competence to plan working hours and professional tasks, managing behavior to achieve a result.

(Continued)

TABLE 1 (Continued)

No.	Elements of entrepreneurial skills	Generic competences (viewed from the perspective of the individual's personal experience)	Professional competences (viewed from the perspective of the individual's professional experience)
13	Ability to cope with uncertainty	Competence to adapt to changing socio-economic, political, and personal living conditions.	Competence to adapt to changing situations in the professional field—performance of work tasks, changes in the sector or the labor market, etc.
14	Ability to work in a team	Competence to cooperate with family, friends, various social groups, local commune, etc.	Competence to participate in the work team and cooperate with management, colleagues, clients, partners, etc.
15	Learning from experience	Competence to evaluate one's activities—positive and negative aspects—and the possibility to change activities, situations, and attitudes in order to improve.	Competence to reflect on professional activities—to evaluate successes, identify necessary improvements, and learn from the experience for further professional development.

attitude and values (Mulder, 2014). It is considered both in a narrow context of specific professional activities and in the broader context of common professional standards. Professional competences are contextual, variable, and need to be developed along with changing labor market requirements, which leads to the conclusion that different professions will require different knowledge and skills but could also have complementary competences, such as values.

Entrepreneurial skills from two perspectives—generic and professional competences

To understand the distinction between the performance of entrepreneurial skills as generic and professional competences, and after analyzing the students' self-assessments of their entrepreneurial skills and evaluating the difference between students of bioeconomics and education, the authors created Table 1, where the performance of entrepreneurial skills from the two perspectives can be seen. This was based on the *EntreComp* conceptual model of entrepreneurial skills, which consists of 15 fundamental elements (Bacigalupo et al., 2016).

Self-monitoring, a skill necessary for effective self-assessment, involves paying focused attention to some aspects of behavior or thinking and actual doing, often in relation to external standards. Thus, self-monitoring concerns awareness of thinking and progress as it occurs, and as such, it helps to identify parts of what students do when they self-assess (McMillan and Hearn, 2008). The second component of self-assessment, self-judgement, involves identifying progress toward targeted performance. Made in relation to established standards and criteria, these judgements give students a meaningful idea of what they know and what they still need to learn (Bruce, 2001). Students find it difficult to manage self-assessment, which leads to data from students' self-assessments not always coinciding with their actual level; however, it should be considered that students' assessment skills constantly improve in the learning process (Slišāne et al., 2021b).

Given that professional competences are a part of the generic competences and overlap with the field of work, the authors assumed that professions where concrete skills are needed more will be more advanced and students would naturally assess it higher. However, it must also be taken into account that different professional fields have higher expectations regarding the level of development, and it might be that self-assessment is higher because of lower expectations.

The specific research questions of this article are thus as follows:

1. What entrepreneurship competences emerge among Latvian bioeconomics and educational science students?
2. How do entrepreneurship competences differ between bioeconomics and educational science students?
3. How are entrepreneurship competences correlated with each other?

The study aims to shed light on how specific university study disciplines with a professional focus (educational sciences and bioeconomics) support the development of a specific generic competence (entrepreneurship competences).

Methodology

In this study, entrepreneurship competences were assessed and compared for students of educational sciences and bioeconomics. Data were gathered by using the online survey platform QuestionPro. The questionnaire was filled in by 135 students from five Latvian universities (Rezekne Academy

of Technology, University of Latvia, Daugavpils University, Liepaja University, and Riga Technical University), of whom 82 were from the field of educational sciences and 53 from the field of bioeconomics. The study field of bioeconomics was chosen as a result of the fact that entrepreneurship competences should be improved as both generic and professional competences in this area, while in the field of educational sciences, entrepreneurship competences should only be regarded as generic competences.

The study participants filled out the questionnaire as part of a module in different study programmes. The questionnaire was proposed to students as an alternative to another study. The participants were selected on an accessibility basis. Of the participants, 77% were women and 23% were men, and their average age was 30 years (SD = 8.09, Mo = 24, Me = 28). Of the participants, 18% were bachelor's students, 70% were master's students, and 12% were doctoral students. Students were asked to assess their entrepreneurship competences with 47 statements (Appendix 1) on a 7-point Likert scale (where 1 = not characteristic of me at all and 7 = completely characteristic of me). Their entrepreneurship competences were evaluated through 3 sub-competences that were further divided into 15 dimensions and 47 criteria. The value of each dimension was defined as the mean value of the corresponding statements' self-assessment values and was rounded to 2 decimal places. The sub-competence value was defined as the mean value of all corresponding dimensions' self-assessments rounded to two decimal places. To determine the questionnaire's internal consistency, Cronbach's alpha values were calculated for entrepreneurship competences as well as for each sub-competence separately to make sure that the criteria set for each sub-competence also had internal consistency. Correlations between entrepreneurial dimensions were explored for each study field separately. The exploratory factor analysis was chosen to examine how the questionnaire functions among Latvian bioeconomics and educational science students and to determine the number of factors that could be identified in the data. To determine whether there were statistically significant differences between each sub-competence, an independent sample *t*-test was carried out on the mean values of the self-assessments of students of educational sciences and students of bioeconomics.

The study used an assessment tool for students' transversal competences developed in the ESF project 8.3.6.2: "Development and Implementation of the Education Quality Monitoring System" 8.3.6.2/17/I/001 (Miltuze et al., 2021; Dimdiņš et al., 2022). One of the six transversal competences and two out of eight study fields were analyzed.

The questionnaire was available for completion from 26 November 2020 to 13 March 2021, and the data were analyzed using SPSS and Microsoft Excel. The study considered all ethical research standards in accordance with the General Data Protection Regulation (GDPR). The questionnaire was completed anonymously and participation in it was completely voluntary.

TABLE 2 Cronbach's alpha values for each entrepreneurship sub-competence.

Sub-competence	Cronbach's alpha
Problem-solving skills and creativity	0.954
Identification, mobilization, and efficient use of internal and external resources	0.894
Initiative and action orientation	0.922

Results

To determine the internal consistency of the Likert scale, the value of Cronbach's alpha was calculated for entrepreneurship competences ($\alpha = 0.962$) and each sub-competence separately (Table 2). The value of Cronbach's alpha for entrepreneurship competences as a whole and all sub-competences is >0.89 and is therefore considered to be high. Therefore, the Likert scale is reliable.

The exploratory factor analysis was chosen to examine how the questionnaire functions among Latvian bioeconomics and educational science students and to determine the number of factors that could be identified in the data. The KMO value (0.882) is >0.8; therefore, the correlation matrix is "meritorious" (Kaiser and Rice, 1974). To reduce the number of factors, the parallel analysis engine was used (Patil et al., 2017). The number of factors to retain will be the number of eigenvalues (generated from the researcher's dataset) that are larger than the corresponding random eigenvalues (Horn, 1965). Therefore, five factors were retained. For interpretation, the Kaiser–Varimax rotation matrix was used (Appendix 1). The results indicate that the statements that measure problem-solving skills and creativity sub-competences are mostly part of the first factor; statements that measure identification, mobilization, and efficient use of internal and external resources sub-competences are mostly part of the third and fifth factors; and statements that measure initiative and action orientation are mostly part of second and fourth factors.

By analysing the self-assessments of entrepreneurship competences in each of its sub-competences and comparing the mean values of the students of educational sciences' and students of bioeconomics' self assessments, it can be concluded that the results are similar. In two out of three sub-competences, students of educational sciences assessed their entrepreneurship skills higher than students of bioeconomics (Table 3).

Bioeconomics students' self-assessments' mean values are higher than the self-assessments of educational sciences students in the sub-competence of problem-solving skills and creativity. However, students of educational sciences assessed their identification, mobilization, and efficient use of internal and external resource sub-competences and initiative and action orientation sub-competences to be higher than those of bioeconomics students.

When analyzing students' entrepreneurship competences from both a professional perspective and a generic perspective,

TABLE 3 Mean values of educational sciences and bioeconomics students' self-assessments of entrepreneurship sub-competences.

	Problem-solving skills and creativity	Identification, mobilization, and efficient use of internal and external resources	Initiative and action orientation
Bioeconomics students	4.61 (SD = 0.82)	4.91 (SD = 0.81)	4.88 (SD = 0.74)
Educational sciences students	4.58 (SD = 1.30)	5.12 (SD = 0.85)	4.98 (SD = 0.91)

TABLE 4 Results of students' self-assessment of the problem-solving skills and creativity sub-competence.

Dimension	Bioeconomics students			Education students			Difference between mean values	Independent *t*-test *p*-value
	Mean	Median	Standard deviation	Mean	Median	Standard deviation		
Ability to spot opportunities	4.42	4.50	1.00	4.24	4.25	1.35	0.18	0.408
Creativity	4.36	4.60	1.05	4.64	4.80	1.44	−0.28	0.226
Vision	4.56	4.75	1.17	4.69	5.00	1.31	−0.13	0.575
Ability to critically evaluate ideas	4.91	5.00	0.99	4.68	5.00	1.51	0.23	0.337
Ethical and sustainable thinking	4.82	5.00	1.12	4.66	5.00	1.59	0.16	0.522

the results showed that the mean value is higher for educational sciences students than for bioeconomics students according to their own self-assessment. This might not be in line with anyone's expectations considering the essential role and necessity of entrepreneurial capacity in the further professional activities of bioeconomics students. Therefore, it is important to analyse and compare the results in each dimension of the entrepreneurship competences to find answers to the possible reasons for students' self-assessments in each field of study.

Problem-solving skills and creativity

This sub-competence of problem-solving skills and creativity contains five dimensions, three of which have a higher mean value in the self-assessments of bioeconomics students (Table 4).

The mean values show that bioeconomics students evaluated the sub-competences of the ability to spot opportunities, the ability to critically evaluate ideas, and ethical and sustainable thinking higher than students of educational sciences. However, the only median value that is higher for bioeconomics students is their ability to spot opportunities, while those for the ability to critically evaluate ideas and ethical and sustainable thinking are exactly the same for students from both study fields. By comparing the mean self-assessment values in the dimensions of creativity (Bioec. st. mean = 4.36, Ed. st. mean = 4.64) and vision (Bioec. st. mean = 4.56, Ed. st. mean = 4.69), it can be seen that higher mean values have been reported by students of educational sciences.

However, with a p-value >0.05 for each sub-competence, the results were not considered statistically significant. An analysis of students' self-assessments standard deviation leads to the conclusion that, in all five dimensions of the problem-solving skills and creativity sub-competence, educational students have significantly higher data dispersion. Further, while the standard deviations for bioeconomics students ranged from 0.99 to 1.17, those of students of educational sciences ranged between 1.31 and 1.59. This points to a polarization of education students' evaluations.

Although creativity and vision are essential parts of bioeconomics and students should therefore develop these competences from a professional perspective, we must keep in mind that they are also essential competences for educators. In the context of entrepreneurship competences, creativity and vision are characterized by the ability to create added value, and the use of external resources is required from a monetary perspective for bioeconomics students, while education students are associated with the ability to create added intellectual value for their pupils. The vision dimension is characterized by the development of future scenarios and the capacity for strategic decision-making, which is necessary as a professional competence both in the context of education and bioeconomics.

Following an analysis of the Spearman's rank correlations, we can conclude that there are significant differences in the number of dimensions between which a strong correlation (higher than or equal to 0.7) exists in each field of study (Table 5).

Strong correlations exist between educational sciences students' ability to spot opportunities, creativity, vision, and ability to critically evaluate ideas in all possible

TABLE 5 Spearman's rank correlations between all dimensions of the problem-solving skills and creativity sub-competence.

Dimension	Ability to spot opportunities		Creativity		Vision		Ability to critically evaluate ideas		Ethical and sustainable thinking	
	Ed. st.	Bi. st.	Ed. st.	Bi. st.	Ed. st.	Bi. st.	Ed. st.	Bi. st.	Ed. st.	Bi. st.
Ability to spot opportunities	1.00	1.00	0.78**	0.60**	0.73**	0.54**	0.79**	0.52**	0.63**	0.262
Creativity	0.78**	0.60**	1.00	1.00	0.83**	0.57**	0.84**	0.54**	0.68**	0.37**
Vision	0.73**	0.54**	0.83**	0.57**	1.00	1.00	0.89**	0.62**	0.68**	0.44**
Ability to critically evaluate ideas	0.79**	0.52**	0.84**	0.54**	0.89**	0.62**	1.00	1.00	0.75**	0.33*
Ethical and sustainable thinking	0.63**	0.26	0.68**	0.36**	0.68**	0.44**	0.76**	0.33*	1.00	1.00

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

TABLE 6 Results of students' self-assessment of the identification, mobilization, and efficient use of internal and external resource sub-competence.

Dimension	Bioeconomic students			Education students			Difference between mean values	Independent *t*-test *p*-value
	Mean	Med.	St. dev.	Mean	Med.	St. dev.		
Awareness and self-efficacy of your skills	5.31	5.33	0.96	5.46	5.67	1.05	−0.15	0.421
Motivation and perseverance	5.39	5.50	1.05	5.49	5.50	1.02	−0.10	0.579
Ability to mobilize the necessary resources	4.80	5.00	1.24	5.15	5.17	1.06	−0.35	0.077
Financial and economic expertise	4.22	4.33	1.33	4.33	4.33	1.37	−0.11	0.66
Ability to mobilize and motivate human resources	4.87	5.00	0.99	5.21	5.38	1.09	−0.34	0.07

combinations, and there are moderate correlations (between 0.4 and 0.7) between ethical and sustainable thinking and the other four dimensions. For bioeconomics students, there does not exist a strong correlation between any of the problem-solving skills and creativity sub-competences' dimensions. Although 7 out of 10 possible combinations of dimension pairings have a moderate correlation, we can conclude that the relationship between dimensions is significantly weaker in the self-assessments of bioeconomics students.

Consequently, it can be concluded that, within the dimensions of problem-solving skills and creativity, bioeconomics students in the study process most likely need to focus on the ability to spot opportunities, the ability to critically evaluate ideas, and the ability to focus on ethical and sustainable thinking as professional competences. For students of educational sciences, creativity and vision are better developed according to their self-assessments. This could be related to the specific nature of the teacher's work, where it is necessary to focus on both the creative use of different teaching methods in the learning process and the long-term planning of the process to achieve the learning objectives.

The identification, mobilization, and efficient use of internal and external resources

By analyzing the mean values of identification, mobilization, and efficient use of internal and external resource sub-competence, it can be concluded that education students have evaluated their competences as higher in all five dimensions (Table 6).

Awareness and self-efficacy of your skills, motivation and perseverance, and the ability to mobilize and motivate human resources are important competences for future educators, and, therefore, the results are to some degree in line with professional necessities. However, the ability to mobilize the necessary resources and financial and economic expertise are dimensions that are closely related to economics. In both dimensions, the mean value for educational sciences students is higher compared to bioeconomics students' self-assessments. This could indicate that bioeconomics students had higher expectations for the level of their development, and thus, it might be that educational sciences students' higher self-assessment relates to their lower expectations. Educational sciences students' high self-assessments in these two dimensions

TABLE 7 Spearman's rank correlation between all dimensions of the identification, and efficient use of internal and external resources sub-competence.

Dimension	Awareness and self-efficacy of your skills		Motivation and perseverance		Ability to mobilize the necessary resources		Financial and economic expertise		Ability to mobilize and motivate human resources	
	Ed. st.	Bi. st.	Ed. st.	Bi. st.	Ed. st.	Bi. st.	Ed. st.	Bi. st.	Ed. st.	Bi. st.
Awareness and self-efficacy of your skills	1.00	1.00	0.71**	0.73**	0.55**	0.51**	0.25*	0.16	0.54**	0.29*
Motivation and perseverance	0.71**	0.73**	1.00	1.00	0.64**	0.51**	0.23*	0.17	0.54**	0.29*
Ability to mobilize the necessary resources	0.55**	0.51**	0.64**	0.51**	1.00	1.00	0.47**	0.42**	0.43**	0.42**
Financial and economic expertise	0.25*	0.16	0.23*	0.17	0.47**	0.42**	1.00	1.00	0.45**	0.56**
Ability to mobilize and motivate human resources	0.54**	0.29*	0.54**	0.29*	0.43**	0.42**	0.45**	0.56**	1.00	1.00

**Correlation is significant at the 0.01 level (2-tailed); *Correlation is significant at the 0.05 level (2-tailed).

TABLE 8 Results of students' self-assessment of the initiative and action orientation sub-competence.

Dimension	Bioeconomics students			Education students			Difference between mean values	Independent *t*-test *p*-value
	Mean	Median	Standard deviation	Mean	Median	Standard deviation		
Initiative	5.25	5.33	1.05	5.44	5.67	1.07	−0.19	0.318
Planning and management	5.14	5.33	1.12	5.25	5.33	1.09	−0.11	0.58
Ability to cope with uncertainty	4.87	5.00	1.00	5.06	5.17	1.20	−0.19	0.349
Ability to work in a team	4.61	4.67	1.28	4.78	5.00	1.19	−0.17	0.443
Learning from experience	5.38	5.33	1.01	5.36	5.33	1.02	0.02	0.931

need to be studied in more detail in future research. However, with a p-value >0.05 for each sub-competence, the results were not considered statistically significant.

By comparing the median of students' self-assessment in educational sciences and bioeconomics, it can be seen that the median in each of the five dimensions is also higher for education students. Further, by analyzing the correlations between the different dimensions of the identification, mobilization, and efficient use of internal and external resources sub-competence, we can conclude that, in the self-assessments of educational sciences students and bioeconomics students, the dimensions between which strong or moderate correlations exist are similar (Table 7).

The only strong correlation that exists is between awareness and self-efficacy of your skills and motivation and perseverance (Bioec. st. = 0.73, Ed. st. = 0.71) for students from both study fields. There is a moderate correlation between 5 out of 10 possible dimension pairings for both study fields, and the correlation coefficient values are similar. This could point to the fact that both sets of students have a similar understanding, and the manifestation of these competences from a professional perspective is similar. However, for dimension pairings like the ability to mobilize and motivate human resources and the ability to spot opportunities (Bioec. st. = 0.29, Ed. st. = 0.54) or mobilize and motivate human resources and motivation and perseverance (Bioec. st. = 0.29, Ed. st. = 0.54), only a moderate correlation exists for educational sciences students, while for bioeconomics students, the correlation between these dimensions is considered to be weak.

Consequently, it can be concluded that students of educational sciences have a higher opinion of their identification, mobilization, and efficient use of internal and external resources sub-competence than bioeconomics students. However, the limitations of the self-assessment should be taken into account.

Initiative and action orientation

By comparing the mean values of the self-assessments in all dimensions of the initiative and action orientation sub-competence, it can be concluded that, in four out of five dimensions, the mean value is higher for educational sciences students (Table 8).

The only mean value that is not higher for educational sciences students in this sub-competence is learning from

TABLE 9 Spearman's rank correlation between all dimensions of the initiative and action orientation sub-competence.

Dimension	Initiative		Planning and management		Ability to cope with uncertainty		Ability to work in a team		Learning from experience	
	Ed. st.	Bi. st.	Ed. st.	Bi. st.	Ed. st.	Bi. st.	Ed. st.	Bi. st.	Ed. st.	Bi. st.
Initiative	1.00	1.00	0.67**	0.76**	0.52**	0.51**	0.54**	0.43**	0.55**	0.42**
Planning and management	0.67**	0.76**	1.00	1.00	0.66**	0.52**	0.51**	0.42**	0.58**	0.59**
Ability to cope with uncertainty	0.52**	0.51**	0.66**	0.52**	1.00	1.00	0.76**	0.55**	0.67**	0.49**
Ability to work in a team	0.54**	0.43**	0.51**	0.42**	0.76**	0.55**	1.00	1.00	0.56**	0.50**
Learning from experience	0.55**	0.42**	0.58**	0.59**	0.67**	0.49**	0.56**	0.50**	1.00	1.00

**Correlation is significant at the 0.01 level (2-tailed); *Correlation is significant at the 0.05 level (2-tailed).

experience (Bioec. st. = 5.38, Ed. st. = 5.36). However, the content of the other dimensions should be taken into account. The initiative includes taking responsibility and demonstrating initiative when tackling problems. Planning and management include job planning, goal setting, and time management. The ability to cope with uncertainty includes risk assessment and decision-making despite uncertainty. The ability to work in a team includes cooperation with both interested and uninterested parties. All these dimensions are essential parts of the day-to-day work of the educator. Consequently, the high results presented could indicate that these dimensions are necessary for education students to fully prepare for future professional challenges. However, with a p-value >0.05 for each sub-competence, the results were not considered statistically significant. By analyzing the correlation between all dimensions of the initiative and action orientation sub-competence, it can be concluded that there is a moderate or strong correlation between all dimensions for students from both fields (Table 9).

Although there is only one dimension pair in each of the fields of study between which there is a strong correlation, the results point to the consistency of the interrelationship between the dimensions in both fields of study. This could point to the fact that the manifestation of the initiative and action orientation sub-competence, both from a professional and a generic individual perspective, is similar in different individual and working contexts for students from both fields of study.

Discussion/conclusion

Entrepreneurship competences consist of two perspectives: generic competences viewed from the perspective of the individual's personal experience and professional competences viewed from the perspective of the individual's professional experience. The present study compared the self-assessments of bioeconomics students' and students of educational sciences' entrepreneurship competences. Despite the fact that entrepreneurship is more linked to economics, the results showed that, in two out of the three sub-competences, students of educational sciences assessed their entrepreneurship competences higher than students of bioeconomics. In the identification, mobilization, and efficient use of internal and external resources (Bioec. st. mean = 4.91, Ed. st. mean = 5.12), and initiative and action orientation sub-competences (Bioec. st. mean = 4.88, Ed. st. mean = 4.98), students of educational sciences self-assessed themselves higher than bioeconomics students, and the mean values for the problem-solving skills and creativity sub-competence are very similar (Bioec. st. mean = 4.61, Ed. st. mean = 4.58).

There are several potential reasons that might have determined the results of the study. First, educational sciences cover a wider spectrum of generic competences needed for everyday work. It is important for the educator not only to be an expert in a specific field of science but, more importantly, to be able to teach others, which includes being able to organize, manage, set objectives, cooperate, communicate, and various other generic competences (Jamil et al., 2015; Osman, 2011), while historically, in Latvian higher education, particularly in STEM sciences, the focus is on knowledge in the learning field (Namsone et al., 2021; Dudareva et al., 2021). Therefore, generic competences are neglected. Further, one of the limitations of the study is the evaluation method used. The accuracy of the self-assessment survey, which is related to the assessment form, is lower compared to objective ability tests or behavioral observations because respondents' responses can be affected by their limited ability to remember specific examples of their behavior, distorted memories of their past behavior, and a general tendency to assess themselves, their skills, and their abilities higher than they actually are (Rubene et al., 2021; Miltuze et al., 2021; Dimdiņš et al., 2022). It is possible that specific professional knowledge and understanding of the complexity of the highest levels of competences led bioeconomics students to assess their professional skills more objectively and therefore lower than educational sciences students. The assessment of the dimension of financial and economic expertise also shows this, where students of educational sciences (mean = 4.33) assessed their expertise higher than bioeconomics students (mean = 4.22).

It is also important to highlight the situation in Latvia, where students of educational sciences, due to a lack of teachers, start working shortly after starting their studies (OECD, 2020; Koroleva et al., 2017). This allows students to take on a lot of responsibility and to develop the necessary generic competences even further.

However, future studies should focus more closely on the reasons for the differences in the self-assessments. It is necessary to understand whether educational sciences students' high self-assessment of their entrepreneurship competences is linked only to the limitations mentioned above or whether there are specific teaching and learning methods used in educational sciences studies that can serve as examples of good practice for the development of entrepreneurship competences in other fields of study.

Limitations

The self-assessments have a high risk of not being representative because individuals' perceptions of their level are their own valuations from their own perspectives. Data have been taken from a pilot study; thereby, the sample was not representative as there was no random sample with a certain number and the data that contain students representing different study years and study levels need to be taken into account.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Author contributions

Conceptualization: AS and GL. Methodology, software, formal analysis, resources, writing—original draft preparation, and writing—review and editing: AS. Validation: AS, GL, and ZR. Investigation, data curation, and visualization: GL. Supervision, project administration, and funding acquisition: ZR. All authors have read and agreed to the published version of the manuscript.

Funding

This research was supported by the projects Strengthening the capacity of doctoral studies at the University of Latvia within the framework of the new doctoral model (Identification no. 8.2.2.0/20/I/006) and Assessment of Competences of Higher Education Students and Dynamics of Their Development in the Study Process (ESF project 8.3.6.2: Development and Implementation of the Education Quality Monitoring System) (Project agreement no. ESS2022/422).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abel, J. R., and Gabe, T. M. (2011). Human capital and economic activity in urban America. *Reg. Stud.* 45:8, 1079–1090. doi: 10.1080/00343401003713431

Akadēmiskās Informācijas Centrs (2017). *Augstākās izglitibas kvalitātes monitoringa sistēmas koncepcija*. Available online at: https://aic.lv/content/files/Augst%c4%81k%c4%81s%20izgl%c4%abt%c4%abbas%20kvalit%20%c4%81tes%20monitoringa%20sist%c4%93mas%20koncepcija.pdf accessed May 30, 1985.

Arrighi, G. (2005). "Globalization in world-systems perspective," in *Critical Globalization Studies*, eds. R. Appelbaum and W. I. Robinson (New York, NY: Routledge), 33–44.

Bacigalupo, M., Kampylis, P., Punie, Y., and Van den Brande, G. (2016). *EntreComp: The Entrepreneurship Competence Framework* (Luxembourg, Europe: Publication Office of the European Union).

Bruce, L. B. (2001). Student Self-Assessment: Making Standards Come Alive. *Classroom Leadership* 1:5, 1–6.

Cheetham, G., and Chivers, G. (1996). Towards a Holistic Model of Professional Competence. *J. Eur. Ind. Train.* 20:5, 20–30. doi: 10.1108/03090599610119692

Cheetham, G., and Chivers, G. (1998). The reflective (and competent) practitioner: a model of professional competence which seeks to harmonise the reflective practitioner and competence-based approaches. *J. Eur. Ind. Train.* 22:7, 267–276. doi: 10.1108/03090599810230678

Dimdiņš, G., Miltuze, A., and Olesika, A. (2022). "Development and initial validation of an assessment tool for student transversal competences," in *Proceedings of Scientific Papers on Human, Technologies and Quality of Education* (Riga, University of Latvia).

Direito, I., Duarte, A. M., and Pereira, A. (2014). The development of skills in the ICT sector: analysis of engineering students' perceptions about transversal skills. *Int. J. Eng. Ed.* 20, 6B. 1556–1561. https://www.it.pt/Publications/PaperJournal/9821

Dudareva, I., Namsone, D., Butkēviča, A., and Cakane, L. (2021). Teacher competence gap identification by using an online test. *INTED*2021 *Proceedings* (IATED) 9787–9791.

Economou, A. (2016). *Research Report on Transversal Skills Frameworks.* Available online at: http://www.ats2020.eu/images/deliverables/D1.1_TransversalSkillsFrameworks_CP.pdf Accessed July 05, 2022.

Epstein, R. M., and Hundert, E. M. (2002). Defining and assessing professional competence. *JAMA* 287:2, 226–235. doi: 10.1001/jama.287.2.226

European Parliament and Council (2006). Recommendation of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning. *Off. J. Eur. Union*, (2006/962/EC).

Florea, N. (2014). Contribution to gender studies for competences achievement stipulated by national qualifications framework in higher education (NQFHE). *J. Res. Gend. Stud.* 4:2, 741–750. https://www.ceeol.com/search/article-detail?id=2974

Grant, G., Elbow, P., Ewens, T., Gamson, Z., Kohli, W., Neumann, W., et al. (1979). *On Competence: A Critical Analysis of Competence-Based Reforms in Higher Education* (San Francisco: Jossey-Bass).

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika* 30, 179–185. doi: 10.1007/BF02289447

Jamil, F. M., Sabol, T. J., Hamre, B. K., and Pianta, R. C. (2015). Assessing teachers' skills in detecting and identifying effective interactions in the classroom. *Elementary School J.* 115, 407–432. doi: 10.1086/680353

Kaiser, H. F., and Rice, J. (1974). Little Jiffy, Mark Iv. *Educ. Psychol. Measure.* 34, 111–117. doi: 10.1177/001316447403400115

Kalleberg, A. L. (2011). *Good Jobs, Bad Jobs: The Rise of Polarized and Precarious Employment Systems in the United States, 1970s-2000s* (New York: Russell Sage Press).

Komarkova, I., Gagliardi, D., Conrads, J., and Collado, A. (2015). *Entrepreneurship Competence: An Overview of Existing Concepts, Policies and Initiatives—Final Report*. Joint Research Centre Technical Report.

Koroleva, I., Trapenciere, I., Aleksandrovs, A., and Kaša, R. (2017) *Studentu Sociālie un Ekonomiskie Dzives Apstākli Latvijā.* Available online at: https://www.izm.gov.lv/lv/media/3943/download Accessed August 06, 2022.

Kuckertz, A., Berger, E. S. C., and Brändle, L. (2020). Entrepreneurship and the sustainable bioeconomy transformation. *Environ. Innov. Soc. Transit.* 37, 332–344. doi: 10.1016/j.eist.2020.10.003

Larraz, N., Vázquez, S., and Liesa, M. (2017). Transversal skills development through cooperative learning. Training teachers for the future. *On the Horizon* 25:2, 85–95. doi: 10.1108/OTH-02-2016-0004

Luppi, E., Bolzani, D., and Terzieva, L. (2019). Assessment of transversal competencies in entrepreneurial education: a literature review and a pilot study. *Form@re—Open J. Per La Formazione in Rete* 19:2, 251–268. doi: 10.13128/formare-25114

McMillan, J. H., and Hearn, J. (2008). Student Self-Assessment: The Key to Stronger Student Motivation and Higher Achievement. *Educ. Horiz.* 87:1, 40–49. https://eric.ed.gov/?id=EJ815370

Miltuze, A., Dimdinš, G., Olesika, A. Ābolina, A., Lāma, G., Medne, D., Rubene, Z., et al. (2021). *Augstākajā izglitibā studējošo caurviju kompetenču novērtēšanas instrumenta (CKNI) lietošanas rokasgrāmata* (Latvijas Universitāte, Riga).

Moretti, E. (2004). Workers' education, spillovers, and productivity: evidence from plant-level production functions. *Am. Econ. Rev.* 94:3, 656–690. doi: 10.1257/0002828041464623

Mulder, M. (2014). "Conceptions of professional competence," in *International Handbook of Research in Professional and Practice-based Learning*, eds. S. Billett, C. Harteis, and H. Gruber (Dordrecht: Springer), 107–137.

Namsone, D., Cakane, L., and Erina, D. (2021). Theoretical framework for teachers self- assessment to teach 21st century skills. society. integration. education. *Proc. Int. Sci. Conf.* 2, 402–429. doi: 10.17770/sie2021vol2.6437

Neumark, D., and Reed, D. (2004). Employment relationships in the new economy. *Labour Econ.* 11:1, 1–31. doi: 10.1016/S0927-5371(03)00053-8

OECD (2020). *PISA 2018 Results (Volume V): Effective Policies, Successful Schools, PISA* (Paris, OECD Publishing).

Osman, K. (2011). The Inculcation of Generic Skills through Service Learning Experience among Science Student Teachers. *Procedia - Social and Behavioral Sciences*, 18, 148–153. doi: 10.1016/j.sbspro.2011.05.022

Pârvu, I., Ipate, D. M., and Mitran, P. C. (2014). Identification of employability skills—starting point for the curriculum design process. *Econ. Manag. Financ. Mark.* 9:1, 237–246.

Patil, V., and Surendra, H. N. S., Sanjay, M., and Donavan, D., T. (2017). *Parallel Analysis Engine to Aid in Determining Number of Factors to Retain using R [Computer software]*. Available from: https://analytics.gonzaga.edu/parallelengine/ Accessed July 07, 2022.

Pepper, D. (2011). Assessing key competences across the curriculum—and Europe. *Eur. J. Educ.* 46:3, 335–353. doi: 10.1111/j.1465-3435.2011.01484.x

Rubene, Z., Dimdinš, G., Miltuze, A., Baranova, S., Medne, D., Jansone-Ratinika, N., et al. (2021). Augstākajā izglitibā studējošo kompetenču novērtējums un to attistibas dinamika studiju periodā. 1. *kārtas noslēguma zinojums* (Riga: Latvia University).

Rubin, M. (2012). Working-class students need more friends at university: a cautionary note for australia's higher education equity initiative. *Higher Educ. Res. Dev.* 31:3, 431–433. doi: 10.1080/07294360.2012.689246

Sá, M. J., and Serpa, S. (2018). Transversal competences: their importance and learning processes by higher education students. *Educ. Sci.* 8:3, 126. doi: 10.3390/educsci8030126

Slišāne, A., Lāma, G., and Bernande, M. (2021a). Knowledge valorisation in doctoral studies in latvia: entrepreneurship and the development of research competencies in the study process. *Acta Paedagogica Vilnensia* 47, 193–210. doi: 10.15388/ActPaed.2021.47.13

Slišāne, A., Lāma, G., and Rubene, Z. (2021b). Self-assessment of the entrepreneurial competence of teacher education students in the remote study process. *Sustainability* 13, 11. 6424. doi: 10.3390/su13116424

Szafranski, M., Golinski, M., and Simi, H. (2017). *The Acceleration of Development of Transversal Competences* (Kokkola: Centria University of Applied Sciences).

UNESCO. (2016). *School and Teaching Practices for Twenty-first Century Challenges: Lessons from the Asia-Pacific Region. Regional Synthesis Report* (Paris: UNESCO).

Appendix

TABLE A1 Rotated factor matrix (Kaiser–Varimax rotation) for each entrepreneurship competences' assessment statement (values below 0.400 are suppressed).

Nr.	Statement	F1	F2	F3	F4	F5
1.Problem-solving skills and creativity						
Ability to spot opportunities						
1.	Recognise the need on the market and, on the basis of existing solutions and knowledge, offer a solution that creates value for society/market	0.654				
2.	From the market and competitor research, see the needs of the market, which does not have an effective/or no solution at all	0.641				
3.	Forecasts market trends and needs	0.626				
4.	Sees opportunities to commercialise knowledge and create added-value products	0.716				
Creativity						
5.	Trying to create ideas that can differ from the more common	0.669				
6.	Understands that a specific task or problem may have different alternative solutions, looks for alternative solutions to problems	0.487	0.426			
7.	Converts an idea into a prototype or a finished product	0.732				
8.	Try to get feedback and develop ideas that create value for others	0.663				
9.	Create new, revolutionary ideas for the market, which differ significantly from existing products and/or services	0.710				
Vision						
10.	Based on an assessment of the current status, future goals, and needed resources, define a vision for the future (preferred location)	0.614				
11.	Builds an inspiring vision that involves others. Defines vision, justifying the importance of the outcome to be achieved (e.g., by creating a solution for a part of society, commercialising the idea, patrolling the discovery, etc., which also involves other people - creating a team)	0.694				
12.	The implementation of the vision is based on strategic decision-making, where the benefits, risks, and good practices are assessed using acquired knowledge and experience.	0.561				
13.	Using the long-term strategy established, plan action steps to achieve this, which includes the necessary tasks, resources, time, people	0.568				
Ability to critically evaluate ideas						
14.	Analyse and compare the added value of different ideas in dealing with similar situations	0.744				
15.	Assess the value of different ideas, by analysing the profitability of the idea by comparing contribution to benefits	0.701				
16.	When assessing the idea, take into account possible future scenarios for its disposal	0.745				
Ethical and sustainable thinking						
17.	Recognise the ethical and sustainability aspects of business and related decision-making	0.732				
18.	When establishing and taking the relevant decisions, are guided by ethical and sustainability principles	0.734				
2. The Identification, Mobilisation, and Efficient Use of Internal and External Resources						
Awareness and self-efficacy of your skills						
19.	Are aware of their strengths and knows how to use them to create value for others			0.714		
20.	Are aware not only of the strengths but also of the weaknesses and specialises accordingly in maximising their potential			0.724		

(Continued)

TABLE A1 (Continued)

Nr.	Statement	F1	F2	F3	F4	F5
21.	Compensates weaknesses by cooperating with others and continues to develop their strengths			0.502		
Motivation and perseverance						
22.	Make efforts and resources to follow own interests and create value for others		0.422	0.641		
23.	Keeps focus on their own interests and goals for a long time, despite failures and difficulties			0.578		
Ability to mobilise the necessary resources						
24.	Find and use existing resources responsibly and effectively					0.437
25.	Compile and manage different types of resources (e.g., human resources, time, finance, natural resources) to create value for others					0.477
26.	Implement and use a strategy to leverage new resources needed to create value for others					0.455
Financial and economic expertise						
27.	Drawing up and managing the budget (e.g., balancing income and expenditure)					0.497
28.	Finds funding opportunities and manages the budget (e.g., drawing up an estimate, raising funds)	0.401				0.602
29.	Not only raises funds to realise the idea but also draws up long-term plans for sustainable financial existence and development (e.g., by creating cash stocks and depreciation deductions)				0.427	0.535
Ability to mobilise and motivate human resources						
30.	Persuades, engages, and inspires others to realise their ideas (e.g., creating a team)					
31.	Motivates and directs human resources to achieve business goals (e.g., managing a team, motivating team members)		0.524			
32.	Persuading other players about the value of their ideas and the development of products					
Initiative and Action Orientation						
Initiative						
33.	Demonstrate awareness of the challenges and is prepared to engage in the development of solutions		0.741			
34.	Follow along, analyse and critically assess problems and propose the development of their solutions		0.789			
35.	Create high-quality solutions for problems, take responsibility for the solution created and its impact on the target group		0.767			
Planning and management						
36.	Defines goals and takes simple steps to partially or fully achieve the set result		0.787			
37.	Drawing up an action plan and working in line with a plan setting out priorities and milestones to achieve its objectives		0.693			
38.	Clarifies priorities and plans to adapt to changing conditions		0.668			
Ability to cope with uncertainty						
39.	Is prepared to make a mistake by testing new things			0.408		
40.	Systematically assess the benefits and risks of action alternatives, choose action with higher value		0.573			
41.	Capable of prolonged action in times of uncertainty and risk when making decisions				0.638	

(Continued)

TABLE A1 (Continued)

Nr.	Statement	F1	F2	F3	F4	F5
Ability to work in a team						
42.	Collaborate and create values while working on a small team		0.598			
43.	Collaborate and create values when working with a broad range of people and groups				0.686	
44.	Form and manage a large team, develop a network of cooperation contacts, and take responsibility for the decisions taken in order to realise the problem				0.665	
Learning from experience						
45.	Take into account other criticisms of the solution or product		0.437	0.431		−0.415
46.	Critically evaluates shortcomings and strengths of a solution or product, identifying things that should be otherwise done		0.694			
47.	Integrate their and other experience in the relevant scope to avoid errors and improve the solution		0.680			

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 23 September 2022
DOI 10.3389/feduc.2022.886307

OPEN ACCESS

EDITED BY
Heidi Hyytinen,
University of Helsinki, Finland

REVIEWED BY
Jani Ursin,
University of Jyväskylä, Finland
Nina Katajavuori,
University of Helsinki, Finland

*CORRESPONDENCE
Katharina Lohberger
katharina.lohberger@erziehung.uni-giessen.de

SPECIALTY SECTION
This article was submitted to Higher Education, a section of the journal Frontiers in Education

RECEIVED 28 February 2022
ACCEPTED 26 August 2022
PUBLISHED 23 September 2022

CITATION
Lohberger K and Braun E (2022) Comparing learning opportunities of generic skills in higher education to the requirements of the labour market. *Front. Educ.* 7:886307. doi: 10.3389/feduc.2022.886307

Comparing learning opportunities of generic skills in higher education to the requirements of the labour market

Katharina Lohberger* and Edith Braun

Institute for Educational Science, Professorship for Teaching and Learning in Higher Education, Justus-Liebig-University Giessen, Germany

In research on higher education, the link between education and future professional success is well-documented. Little research, however, has been done about existing learning opportunities at institutions of higher education that help students acquire generic skills and on the fit between such learning opportunities and labor market demands. To address these questions, we adapted an existing scale for assessing areas of generic skills, which originated in research on job requirements, and transferred it to a survey of students (N = 4,258). We also implemented a comparable questionnaire, assessing the same set of generic skills, in a graduate survey (N = 378). The results of our study show that by using a theoretical model such as this, it is possible to connect student and graduate surveys related to generic skills. Factor analysis provides evidence for the theoretical expected areas for students. Cluster analysis of student data suggests that learning opportunities for generic skills differ according to field of study. We conclude by discussing our study's limitations and implications.

KEYWORDS

higher education, generic skills assessment, generic skills development, learning opportunities, requirements, labor market, higher education graduates, higher education students

Theory

Impact of education for employability

Higher education fosters the development of skills that contribute to the individual and society's prosperity. By examining the relationship between attendance at an institution of higher education and post-graduation working life, research on higher education also contributes to the development of individual educational and career paths. The importance of matching education and professional career is well-documented in the literature on higher education and the labor market (McGuinness, 2006; OECD, 2011; Quintini, 2011; International Labour Organization, 2014). Labor

market match exists when an employee's qualification (level of education) and the skills that they acquired through their education correspond with those required for their chosen profession after graduation (Morgado et al., 2016; Li et al., 2018; Kracke and Rodrigues, 2020). Adequate employment is often linked to higher income (Bauer, 2002; McGuinness, 2006; Sattinger and Hartog, 2013; Levels et al., 2014) and greater satisfaction personally and professionally (McGuinness and Sloane, 2011; Berlingieri and Erdsiek, 2012). However, less is known about learning opportunities for a broad range of generic skills in higher education. Furthermore, there is a lack of research that connects student learning opportunities at institutions of higher education of those generic skills, which graduates report as currently required on the labor market.

A student's field of study at an institution of higher education seems to impact their employment status after graduation (Franzen, 2002; McGuinness and Sloane, 2011; Altonji et al., 2016; Verhaest et al., 2017). The current shortage of skilled workers indicates, however, that it is difficult for employers to find suitable job candidates to provide the appropriate professional qualifications for specific positions (Berlingieri and Erdsiek, 2012). Especially graduates from arts or humanities programs look longer for employment or are at a higher risk for being inadequately employed. This may result from their low level of specialization. It may also be related to the fact that these subjects do not provide training for a specific career path or occupation (Leuze and Strauß, 2008; Verhaest et al., 2017).

Most of the studies have analyzed discipline-specific knowledge and skills (McGuinness, 2006; Baert et al., 2013). Such fields as economics, science, technology, engineering, mathematics, and medicine are known to provide a highly structured curriculum along with discipline-specific knowledge (Bligh, 2000; Johnson et al., 2002). Consequently, there is a link between study programs with a highly structured curriculum and future access to the labor market. Since these fields are also known to have a strong vocational field connection (Falk et al., 2009), their greater access to the labor market could be related to the fact that employers are aware of what knowledge they can expect graduates of such programs to bring with them. In contrast to the labor market usability of those highly structured fields of study, the labor market usability of humanities and arts is discussed in the literature (Robst, 2007; Leuze and Strauß, 2008).

Institutions of higher education function both as autonomous institutions and as integral components of the societies in which they operate. As autonomous institutions, they are "organization centers of the science system" (Wissenschaftsrat, 2013) and social places, where education is linked to research, knowledge transfer, and cultural self-perception (Wissenschaftsrat, 2013). Institutions of higher education also play an essential role in preparing citizens for the current and future labor market. The current labor market appears to require employees who are equipped with a broad range of skills, like communication or problem solving skills (Suarta et al., 2017). In all likelihood, this will continue. As Handel (2020) puts it: Even policymakers and pedagogues, who believe that technological change will not take place as rapidly or as radically as some others have suggested, they nonetheless share the concern that job skill requirements are increasing so quickly, or are poised to do so, that many people are at risk of being shut out of the workforce all together. Institutions of higher education may help counter such risks. In addition to facilitating the learning of domain-specific knowledge and skills, higher education also promotes the acquisition of generic skills. Such skills include an ability: to act independently when confronted with challenging tasks; to communicate effectively; to work cooperatively; and to promote other individuals (Barrie, 2004; Braun and Brachem, 2015; Ursin et al., 2021). Moreover, in addition to improving a student's chance for future professional success, the acquisition of generic skills also appears to facilitate their participation in civic life.

So far, little research has been done on whether there is a fit between existing learning opportunities for generic skills at institutions of higher education and the generic skills required in the labor market. This gap in the literature may be due to the fact that generic skills are difficult to assess.

Job requirements approach for the assessment of generic skills

The job requirements approach is a methodological concept, which aims to assess broad generic skills, and originates in international labor market research (Felstead et al., 2007; OECD, 2011; Autor and Handel, 2013; International Labour Organization, 2014; Handel, 2020). The job requirements approach is based on a number of theoretical considerations:

1. Employees perform tasks that require certain skills. These tasks and skills vary between occupations.
2. Since a certain fit can be assumed between activities and the skills needed to carry them out, the activities can be understood as a proxy of their skills.
3. The employed individual is equipped to provide information about their profession and best able to report on activities in their everyday work life.
4. Individuals are more likely to report with more accuracy about the frequency of their performed activities than about the importance of those activities.

The job requirements approach therefore investigates the employee's activities and the frequency with which they are carried out. Employees are thus asked how often they carry out certain activities, e.g., reading long texts. This information can then be used to identify generic skills that

certain occupations require (Felstead et al., 2007; Braun and Brachem, 2015). The job requirements approach is used internationally for the assessment of job related skills, for example in the Generalized Work Activities Questionnaire (GWA) (Peterson et al., 2001; Handel, 2020; O*NET, 2021), in the International Assessment of Adult Competencies (PIAAC) (Klaukien et al., 2013; OECD, 2013a,b, 2019), the UK Skills Survey (BMRB Social Research, 2006; Felstead et al., 2007), and in the German National Educational Panel Study (NEPS) (Matthes and Christoph, 2011; Matthes et al., 2014). Since the job requirements approach assumes a certain fit between work-related activities and the frequency in which they are carried out and an employee's skills, it provides insights into assessing labor market demands for generic skills (Felstead et al., 2007; OECD, 2011; Autor and Handel, 2013; International Labour Organization, 2014; Handel, 2020).

Our study builds on previous findings from Braun and Brachem (2015), in which they transferred the skills found in international labor market research to a graduate survey. Based on the theoretical foundations and the empirical results, Braun and Brachem (2015) suggest certain areas of generic skills of graduates that can be assessed with their instrument. Braun and Brachem (2017) identified nine areas of generic skills that graduates must possess for their daily work life:

1. planning and organizing of work processes
2. promoting others
3. leading and management
4. dealing autonomously with challenging tasks
5. information processing
6. number processing
7. communication and cooperation, including competence in a foreign language and intercultural communication
8. using information and communication technologies
9. physical performance

These nine areas are the starting point for this present research.

The term "generic skills" can be interpreted in a variety of ways. For the purposes of our study, we draw on the same definition employed by Braun and Brachem (2015, 2017), since our instrument is based on their previous studies. We therefore define generic competences as the ability to successfully master complex situations. Performance-based competences consist of a skillset which can be applied to different disciplines and which is needed in variety of situations (Heijke et al., 2003; Rychen and Salganik, 2003; Green, 2009; Braun and Brachem, 2017).

Other scholars have concentrated on examining skills or competences like critical thinking, interpersonal understanding, problem solving and (written) communication (Jones, 2009b; Badcock et al., 2010; Hyytinen et al., 2015). Analogous to Braun and Brachem (2017), such research has also referred to skills like "information processing"; "dealing autonomously with challenging tasks"; and "communication and cooperation." Current research has thus made inquiries into ascertaining which generic skills tend to help employees acquire and maintain a job, in general, and what is expected of graduates in the labor market, specifically.

We have used the previous findings from labor market research and adopted the instrument developed by Braun and Brachem (2017), which has already been validated for higher education graduates, to test whether there are learning opportunities for students in these, theoretically and empirically established generic skills. Another advantage of this approach is that it makes it possible to examine and compare a broad range of generic competences.

Differences between fields of study regarding generic skills

So far, little research has been done on whether these relevant generic skills are taught in the study programs and on whether there are differences between the study programs in terms of what learning opportunities they offer students as part of their curriculum. Martin et al. (2005) reported that the majority of academic engineering programs do not teach generic skills and therefore recommend that such programs improve their curriculum to provide students with better training in, for example, communication skills (Martin et al., 2005; Paviotti, 2020). Jones (2009a) examined how lecturers perceived differences in how some generic skills were taught in five fields: economics, physics, history, medicine, and law. One of their key findings was that when generic skills are an inherent part of the field of study or the subsequent profession—like problem solving and communication in medicine or writing and critical thinking in history—they tend to be taught as part of the curriculum, either intentionally or not (Jones, 2009a). In a study conducted at an Australian university, Badcock et al. (2010) found that programs in the arts, engineering, and science had significant differences in learning opportunities for certain generic skills. They found that art students, for instance, attained higher scores in critical thinking and interpersonal understanding than those in other programs. Other studies have ascertained that students in the social sciences tend to have stronger skills in written communication. Engineering students, by contrast, appear to score higher in problem solving but lower for generic skills on all other tested scales. Research has also suggested that in fields of study that are based mainly on lectures and in which students are assessed mostly by written exams, students develop fewer generic skills, since they often have merely to recall the content (Bligh, 2000; Johnson et al., 2002).

There is a significant body of research available that examines and measures specific skills, such as communication or critical thinking (Jones, 2009a; Badcock et al., 2010). Studies

with a broader view base their claims either on surveys of students (Kember et al., 2007; Virtanen and Tynjälä, 2019) or on data gathered about graduates (Martin et al., 2005; Suarta et al., 2017). There are few studies that combine data from students and graduates. Furthermore, there are hardly any studies that examine a wide range of generic skills and whose selection of specific skills is conceptually and empirically based.

Comparison of students and graduates

There is abundant research supporting the claim that education plays an important role in the acquisition of generic skills (Crebert et al., 2004; Smith and Bath, 2006; Huber and Kuncel, 2016). Yet, little research has been done that establishes a connection to the labor market. We are not aware of any study that links student assessments with the demands of the labor market. Previous studies have either examined study conditions by surveying students while they are still enrolled in a program or retrospectively by graduates after they have completed them. This results in a problematic phenomenon: Graduates evaluate entry into the labor market and selected criteria for career success, the results of which are then linked retrospectively to assessments of their prior study conditions. In other words, graduates are surveyed *after* they have completed their studies; they are thus asked to evaluate events that took place between three and 6°years prior. During this period, study regulations and study conditions may have changed, making it difficult to attribute the results of the graduate survey to current study conditions. Accordingly, the results of the graduate survey generate very little concrete knowledge that could help policymakers, administrators, and instructors develop curricula and improve the management of study programs that address the current labor market.

To address these gaps in the literature on the relationship between the acquisition of generic skills in higher education and the labor market, research thus needs to be conducted that can directly compare student and graduate data. The aim of this paper therefore is to examine whether such a comparison is feasible by using the same types of assessments for generic skills for students and graduates.

Research questions

Based on previous findings and the proposed theoretical framework, we address the following research questions:

1. Can the proposed conceptual areas of generic skills, developed using the job requirements approach, be transferred to a student survey?
2. a) Can profiles for learning opportunities of generic skills be found in the student data?
3. How are the fields of study distributed among these profiles?
4. How do students rate the learning opportunities for generic skills during their studies and how do graduates perceive the requirements of the labor market for generic skills?

Materials and methods

Sample

To answer these research questions, two online surveys, one among students and one among graduates, were conducted at a large university in Germany between December 2020 and March 2021.

Student sample

For the student survey, all matriculated students were invited *via* e-mail, to participate. Overall the response rate was about 25%. The items regarding our study were filled out by 4,258 students, of which 1,007 were male, 2,931 female, and 25 non-binary. Most of the students were enrolled in programs related to the humanities (N = 761), STEM (N = 586), and educational science (N = 546).

Graduate sample

The graduates were contacted *via* mail (postal, not electronic) and invited to participate in an online survey. Graduates were surveyed about 1.5 years after they had left the university. A total of 378 graduates participated, of which 118 were male, 251 female, and two non-binary; seven did not provide an answer regarding their gender. Most of the respondents graduated in educational science (N = 58), economics (N = 51), and humanities (N = 46).

In both samples, arts (N = 89 students, N = 6 graduates) and social science (N = 114 students, N = 17 graduates) were the least represented. The programs were grouped based on the system of subject classification used by the German Federal Statistical Office (Statistisches Bundesamt, 2021). According to this system, there are ten fields of study (see **Table 1**).

Table 1 shows the distribution of the students and graduates among the fields of study.

Instruments

The surveys included over 100 questions each. The instruments discussed below were included within them. However, the following descriptions only refer to the part, which is relevant for the present study.

TABLE 1 Distribution of fields of study.

Field of study	*N* students	*N* graduates
Humanities	761	46
Sports	133	18
Economics	232	51
Social sciences	114	17
STEM (science, technology, engineering, and math)	586	35
Agriculture and forestry	309	25
Nutritional science	314	31
Veterinary medicine	212	36
Arts	89	6
Psychology	191	32
Educational science	546	58
Total	**3,487**	**355**

N, numbers of students and graduates.

To explore the requirements as well as the learning opportunities, we built on the work of Braun and Brachem (2017). We used the instruments that they developed to assess the generic skills on requirements that graduates encounter on the current labor market and adopted to ask students about available learning opportunities within their fields of study for acquiring generic skills. We divided one scale from Braun and Brachem's previous study into "communication and cooperation" and "foreign language and intercultural communication," resulting in ten areas of generic skills (see Table 2).

Instruments for measuring learning opportunities for students

The ten scales consisted of a total of 27 items, and were framed to survey learning opportunities available to students in their programs. Each item started with phrases such as "In my studies . . .," and then followed a statement of what the students did in their program. We had to reword a few items from the original instrument developed by Braun and Brachem (2017) in order to make them apply to students. The students rated each item on a five-point Likert scale, in which each point was labeled with such time clauses as: (1) "never"; (2) "less than once a month"; (3) "at least once a month, but less than once a week"; (4) "at least once a week, but not daily"; (5) "daily." Cronbach's alpha of the scales ranged between alpha = 0.64 and alpha = 0.84, which was appropriate considering the small numbers, i.e., a maximum of three items. In the area of "physical work," students were asked only one question, so that no scale value was calculated here. All scales and example items are listed in detail in Table 2.

Instruments for measuring generic skills requirements for graduates

The same 27 items were used in the graduate survey. They were worded in such a way that they always began with "In my main occupation . . .," and were then followed by the same statements as in the student survey, using the same response categories. Cronbach's alpha of the scales ranged between alpha = 0.66 and alpha = 0.86, which was appropriate considering the small numbers, i.e., a maximum of three items. The rationale behind this procedure was to use the same wording in order to compare student and graduate answers, and to prevent a situation in which graduates were asked to assess after graduation, the generic skills that they had acquired during their studies. All scales and example items are described in detail in Table 2.

Procedure

In order to answer the first research question, we examined whether the theoretically assumed factor structure of the graduate survey (Braun and Brachem, 2017) can also be confirmed in the student survey. A confirmatory factor analysis with nine factors was carried out for this purpose. The area "physical performance" was not considered as a latent factor, as it only consists of one item and therefore did not allow latent modeling. To evaluate the model fit for the learning opportunities, we used the following cut-off criteria: Considering the Root Mean Squared Error of Approximation (RMSEA) the model fit was considered as close when RMSEA $\leq$ 0.05, RMSEA $\leq$ 0.08 the model fit was considered as reasonable and RMSEA $\leq$ 0.1 was considered as acceptable. For the Standardized Root Mean Squared Residual (SRMR) the model fit was considered as reasonable when SRMR $\leq$ 0.08 and SRMR $\leq$ 0.1 showed an acceptable fit. A Comparative Fit Index (CFI) $\geq$ 0.95 show good fit and CFI $\geq$ 0.9 show acceptable fit (Browne and Cudeck, 1992; Beauducel and Wittmann, 2005; Backhaus et al., 2015). The smaller the RMSEA and SRMR, the better the estimated model fit, while a larger CFI, by contrast, indicated a better model fit.

To analyze the second research question, about possible profiles in learning opportunities in the various fields of study, we conducted a cluster analysis. To be able to conduct the final cluster analyses, we carried out several pre-tests. We first performed a single-linkage cluster analyses to identify the breakout cases. After identifying the breakout cases, we were able to perform a Ward hierarchical cluster analyses with 3,394 cases for learning opportunities for generic skills. The proposed preliminary cluster solutions were tested both graphically, with a dendrogram, and statistically. Based on the hierarchical cluster analyses, we were then able to conduct k-mean clusters, which we then used first to evaluate the different profiles for learning opportunities and second to evaluate the distribution of the fields of study among them. We tested the significance of the distribution using χ^2 and Cramér's V for effect sizes. Even though the area of "physical performance" was only measured with one item, and therefore not included in the confirmatory

TABLE 2 Dimensions for learning opportunities and labor market requirements.

		Student survey	Graduate survey	Student survey	Graduate survey
Dimension	Number of Items	Cronbach's alpha for learning opportunities	Cronbach's alpha for labor market requirements	Example item for learning opportunities ("In my studies . . .")	Example item for labor market requirements ("In my primary profession . . .")
Planning and organizing of work processes	3	0.76	0.75	I organize work processes.	I organize work processes.
Dealing autonomously with challenging tasks	3	0.66	0.80	I assess possible consequences and outcomes for other areas or people.	I assess possible consequences and outcomes for other areas or people.
Promoting others	3	0.80	0.86	I train, teach or educate other people.	I train, teach or educate other people.
Leading	3	0.84	0.81	I learn to set goals or strategies for other areas or people.	I set goals or strategies for other areas or people.
Information processing	3	0.75	0.71	I apply scientific methods, procedures or techniques to solve problems.	I apply scientific methods, procedures or techniques to solve problems.
Number processing	3	0.82	0.85	I specifically analyze information or data.	I specifically analyze information or data.
Communication and cooperation	3	0.81	0.66	I create a joint product as part of a team (reports, presentations, projects, etc.).	I create a joint product as part of a team (reports, presentations, projects, etc.).
Foreign language and intercultural communication	2	0.64	0.67	I communicate in a language other than my mother tongue.	I communicate in a language other than my mother tongue.
Using information and communication technology	3	0.71	0.83	I use internet-based applications to exchange or work out work-related issues with other people.	I use internet-based applications to exchange or work out work-related issue with other people.
Physical performance	1			I carry out tasks, which require physical competences (e.g., manual labor).	I carry out tasks, which require physical competences (e.g., manual labor).

factor analyses, we decided to include it in the evaluation of the profiles of learning opportunities, as this area has been identified as significant in research that applies the job requirement approach (Morgeson and Humphrey, 2006; Felstead et al., 2007; Autor and Handel, 2013).

To enable a comparison between students and graduates, in order to address the third research question, we turned to a visual presentation of mean values. We deliberately refrained from inferential statistical analyses: firstly, because we made no assumptions about existing differences and secondly, because we could make numerous comparisons (between the ten fields of study and ten scales), so that the procedure would be richly explorative and its possible significance random.

All empirical analyses were conducted using Stata 16.1. An overview of used samples and analysis can be seen in **Table 3**.

Results

Confirming the assumed theoretical structure of the constructs for students

A confirmatory factor analysis was conducted to check the structure of nine factors related to learning opportunities for students. The confirmatory factor analysis with the nine constructs showed acceptable to reasonable fits. The RMSEA showed a reasonable to close fit (0.054), the SRMR showed a reasonable fit (0.053), and the CFI showed an acceptable fit (0.925). **Table 4** shows the standardized factor loadings for the areas of generic learning opportunities. In general, the results of the confirmatory factor analyses showed acceptable to high-factor loadings for all items. Only in the construct "planning and

TABLE 3 Methods of analysis and samples used according to the three research questions.

Research question	Sample	Method for analysis
1	Students (N = 3,487)	Confirmatory factor analysis
2	Students (N = 3,487)	Cluster analysis
3	Students (N = 3,487) Graduates (N = 355)	Descriptive statistics (mean values and standard deviations)

organizing of work processes" the item "In my studies I evaluate the performance or quality of people, objects, or processes" showed a barely acceptable factor loading (0.487). Since the content of the item is seen as important for the latent factor, and because the internal consistency of the scale (α = 0.76) was acceptable, we kept the item for further analyses.

Overall, the expected structure could be firmly established in the student survey.

Profiles of learning opportunities in the fields of study

In this section, we examine the student responses to the ten areas related to learning opportunities. The focus here was on the question of whether the fields of study differ in terms of which learning opportunities students reported.

To analyze profiles of learning opportunities, we conducted cluster analyses, and then tested the distribution of the fields of study among those clusters.

The dendrogram of the Ward hierarchical cluster analyses showed three reasonable clusters of profiles for learning opportunities. Therefore, we refined the results with a k-means cluster analyses on those three clusters. Figure 1 shows the mean values of the ten learning opportunities within the three different clusters. Cluster 1 is characterized by overall high number of learning opportunities for generic skills in nine areas. It is particularly noticeable that the area of physical activities was only marginally represented in this cluster. The areas "using information and communication technology"; "planning and organizing of work processes"; and "dealing autonomously with challenging tasks" were particularly high.

In cluster 2, learning opportunities in all ten areas of generic skills were also high and comparable to cluster 1. Cluster 2, in contrast to the other two clusters, showed a particularly high level for physical activity. Again, "using information and communication technology"; "planning and organizing of work processes"; and "dealing autonomously with challenging tasks" were also pronounced. The area "number processing" was the least pronounced. Nevertheless it was similarly high to how it appeared in cluster 1.

Cluster 3 showed overall the fewest learning opportunities for all areas of generic skills. As well as in the other twoclusters, the areas "using information and communication technology"; "planning and organizing of work processes"; and "dealing autonomously with challenging tasks" showed the most learning opportunities in this cluster. The areas "promoting others" and "leading" were particularly low.

In general, it was noticeable that the areas "using information and communication technology"; "planning and organizing of work processes"; and "dealing autonomously with challenging tasks" showed the greatest number of learning opportunities in all three clusters.

In the next step, we looked at the distribution of the fields of study across the three found clusters for learning opportunities. Table 5 shows the assignment of the fields of study to the three profiles of learning opportunities; it also shows absolute frequencies and relative frequencies per row. The analyses revealed that the fields of humanities, social sciences, educational sciences, and psychology could be assigned to the first cluster; sports and arts to the second cluster; and economics, STEM, agriculture and forestry, nutritional science, and veterinary medicine to the third cluster.

This distribution showed a tendency, that the fields of humanities, social sciences, educational science, psychology, sports, and arts provide good opportunities for the acquisition of generic skills, but have partially few learning opportunities in the area of "physical performance"–something they share with economics, STEM, agriculture and forestry, nutritional science, and veterinary medicine. Sports and arts seem to support more learning opportunities for the acquisition of physical skills. For economics, STEM, agriculture and forestry, nutritional science, and veterinary medicine, which could be best assigned to the third cluster, we saw overall the fewest learning opportunities for acquiring generic skills. As Table 5 shows, these are tendencies only and all field of studies show strengths in all profiles.

The relation between the field of study and the profiles formed is statistically significant, but is not high (Cramér's V = 0.25). The assignment to the clusters could be made with varying degrees of clarity. While a few fields, such as arts and psychology, could be clearly assigned to one cluster, some were distributed across several clusters (i.e., sports). Overall, however, an assignment was possible for all ten fields of study. The results confirmed our first assumptions: Cluster analysis provided some evidence that learning opportunities for generic skills differ between the fields of study.

Descriptive comparison of student and graduate assessments

In this section, we make descriptive comparisons between the responses from students on learning opportunities within their fields of study and those from graduates on labor market

TABLE 4 Factor loadings for areas of generic learning opportunities (student data).

Tasks	Item ("In my studies.")	Factor loadings
Planning and organizing of work processes	I organize work processes.	0.87
	I plan timelines.	0.83
	I evaluate the performance or quality of people, objects or processes.	0.49
Dealing autonomously with challenging tasks	I have to react to unexpected situations.	0.63
	I assess possible consequences and outcomes for other areas or people.	0.67
	I show initiative.	0.60
Promoting others	I train, teach or educate other people.	0.72
	I lead groups in a structured way.	0.79
	I support and motivate others.	0.76
Leading	I learn to set goals or strategies for other areas or people.	0.72
	I learn to persuade others.	0.88
	I learn to negotiate with others.	0.82
Information processing	I assess the quality of professional articles.	0.66
	I document complex facts.	0.75
	I apply scientific methods, procedures or techniques to solve problems.	0.70
Number processing	I create number-based diagrams or tables.	0.83
	I carry out complex calculations.	0.76
	I specifically analyze information or data.	0.73
Communication and cooperation	I create a joint product as part of a team (reports, presentations, projects, etc.)	0.72
	I stick to agreements made in a work group.	0.77
	I negotiate compromises with other people.	0.82
Foreign language and intercultural communication	I communicate in a language other than my mother tongue.	0.56
	I maintain contact with people from other cultures or social groups.	0.82
Using information and communication technology	I use internet-based applications to exchange or work out work-related issues with other people.	0.67
	I process content digitally.	0.67
	I deal with questions concerning the digitalization of work processes.	0.68

Criteria of the confirmatory factor analysis: $\chi^2(263) = 2001.65$, $p \leq 0.01$, RSMEA = 0.054, SRMR = 0.053, CFI = 0.925.

requirements. As stated above, we did not conduct any inference statistics, since we did not have a directed hypothesis. For this purpose, we have graphically contrasted the mean values and standard deviations for all students with those for all graduates in bar charts.

The graphic comparisons of the mean values for each of the ten different fields of study are also worth noting. These can be found in the **Appendix**.

Figure 2 shows student perceptions of current learning opportunities within their programs of study and graduate perceptions of current labor market requirements within the bar charts. It also shows the particular standard deviations for every dimension for learning opportunities as well as for labor market requirements. As discussed above, only descriptive analyses were conducted to avoid producing findings that were random.

Based on our study, we can draw some preliminary conclusions. There are areas in which student and graduate perceptions appear to overlap. In categories, such as "dealing autonomously with challenging tasks"; "using information and communication technology"; and "planning and organizing of work processes"; students report learning opportunities for these skills as widely available and graduates report them as greatly in demand. The absolute differences between learning opportunities and labor market requirements are quite big in the areas of "dealing autonomously with challenging tasks" and "promoting others." We also observed that there are areas, such as "number processing" or "leading," where the labor market requirements as perceived by graduates nearly coincided with the learning opportunities reported by students within their fields of study.

Regardless of the field of study, students reported slightly fewer learning opportunities than graduates reported labor market requirements. The exception was in the areas "using information and communication technology"; "information processing"; and "planning and organizing of work processes."

The high standard deviations suggest the high variance in learning opportunities, as well as in the requirements, which is not surprising due to the fact, that **Figure 2** shows mean values and standard deviations aggregated across all fields of study.

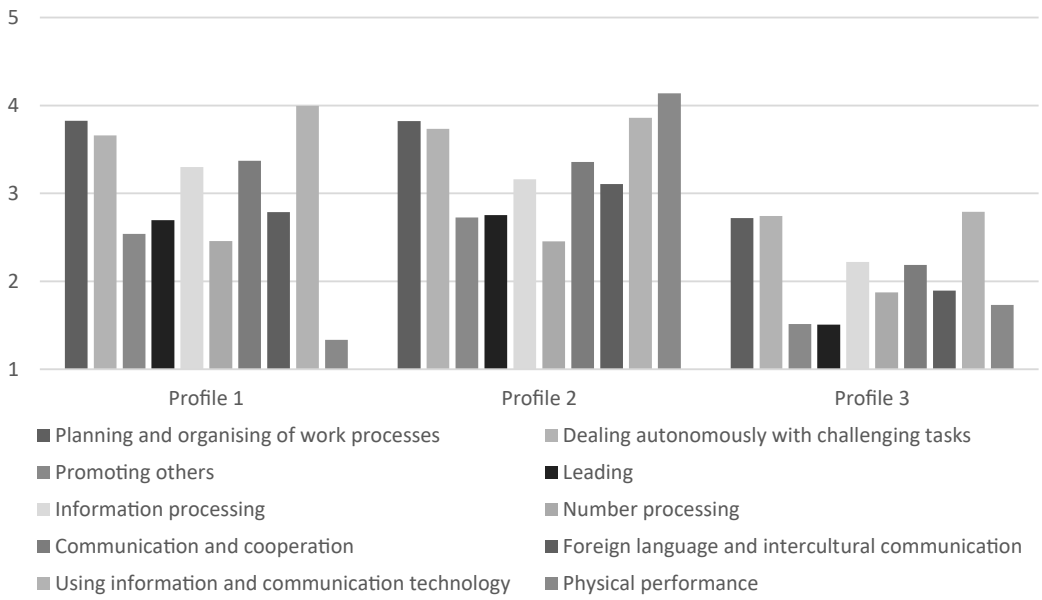


FIGURE 1
Profiles for learning opportunities.

Examining the requirements and learning opportunities at the level of fields of study (see **Appendix**), a more differentiated picture emerged. Especially for veterinary medicine or nutritional science, we observed that learning opportunities and requirements did not align very well with each other. For STEM and social sciences, however, we observed that the reported learning opportunities and labor market requirements nearly coincided.

As noted above, we refrained from inferential statistical comparisons because we did not have an explicit hypothesis directing our study and, moreover, the multiple comparisons could also have generated random findings.

Discussion

This study offers some innovative but nonetheless preliminary results. To be able to capture the acquisition of generic skills in higher education, which has been shown in prior research to be important on the labor market, we transferred ten areas of generic skills that were based on theory from the job requirements approach to a student survey. A confirmatory factor analysis verified the structure of different areas of generic skills, which were adopted from graduate surveys, and therefore also allow the appropriate use in the student survey.

We also analyzed the student data for learning opportunity profiles, structured by the ten areas of generic skills. Cluster analyses showed three different types of learning opportunity profiles, which differed mainly in the frequency of the specific areas of generic skills. While the first cluster contained frequent learning opportunities in almost all areas of generic skills except in the area of "physical performance," the second cluster showed frequent learning opportunities especially in "physical performance," as well as in the other areas. Only the third cluster stood out from the others in that it consistently contained fewer learning opportunities in all areas. If we now look at how students from different disciplines were distributed among the three clusters, some tendencies can be observed: the first cluster was made up mainly of humanities and educational sciences, while the second cluster was made up primarily of arts and sports. The third cluster was made up mostly of STEM and economics. The results are consistent with the previous feedback on perceptions of fields of study. While fields like humanities and educational science are often perceived as fields in which students tend to acquire more general skills, in fields like sports and arts, it is very common to engage in a lot of physical activities. The curriculum for students of STEM and economics, by contrast, is mostly characterized by very stringently prescribed curricula with a high level of discipline-specific skills (Bligh, 2000; Johnson et al., 2002). The results of cluster analyses suggest that the areas of "using information and communication technology," "planning and organizing of work processes," and "dealing autonomously with challenging tasks" are the dimensions in which students seem to perceive the highest learning opportunities for generic skills, regardless of their field of study.

Finally, addressing our third research question, we contrasted the data from students with that of graduates. Here we made use of graphic representations and believe,

TABLE 5 Distribution of the fields of study across the profiles for learning opportunities.

Field of study	*N* profile 1	*N* profile 2	*N* profile 3	Total
Humanities	260 42.76%	183 30.10%	165 27.14%	608
Sports	19 16.38%	50 43.10%	47 40.52%	116
Economics	65 38.69%	35 20.83%	68 40.48%	168
Social sciences	51 51%	22 22%	27 27%	100
STEM	145 31.52%	144 31.30%	171 37.17%	460
Agriculture and forestry	72 30.25%	48 20.17%	118 49.58%	238
Nutritional science	84 32.81%	25 9.77%	147 57.42%	256
Veterinary medicine	32 16.08%	50 25.13%	117 58.79%	199
Arts	10 13.33%	56 74.67%	9 12%	171
Psychology	111 64.91%	20 11.70%	40 23.39%	171
Educational science	159 38.95%	128 30.99%	126 30.51%	413
Total	**1,008** **35.95%**	**761** **27.14%**	**1,035** **36.91%**	**2,804** **100%**

$\chi^2(20) = 347.76$, $p \leq 0.01$, Cramér's $V = 0.25$. The biggest groups are shaded gray.

despite the limited nature of the empirical analysis, that such a comparison offers insights into the differences between learning opportunities in higher education and generic skill requirements in the labor market. Our findings also underscore the importance of generic skills in the areas of "planning and organizing of work processes"; "dealing autonomously with challenging tasks;" and "using information and communication technology" for both students and graduates, which aligns with previous findings (Tynjälä et al., 2006; Suarta et al., 2017).

The preliminary results regarding the relationship between different learning opportunities and the field of study fit in well with previous findings from labor market research. The labor market usability of humanities and arts are discussed in the literature, and there are some regions in the world where arts and humanities programs are being either reduced in size or dismantled all together (Cassity and Ang, 2006; Jenkins, 2015; Olmos-Peñuela et al., 2015; Preston, 2015). Especially for humanities, it is often argued that it is a field in which little discipline-specific knowledge is taught (Leuze and Strauß, 2008; Falk et al., 2009). However, our findings suggest that fields of study within the arts, humanities, and educational science in particular, offer learning opportunities for the acquisition of generic skills. Previous studies have shown that learning environments, which promote collaborative learning or where students have to deal with authentic problems foster the development of generic skills (Kember et al., 2007; Virtanen and Tynjälä, 2019). This could be linked to the fact that fields like social sciences, humanities, and psychology often deal with current developments and frequently implement collaborative methods. By bringing together student surveys and labor market research, our study was thus able to fill a gap in knowledge about the benefits of arts and humanities programs for the labor market in general. The benefits of STEM programs for an individual's economic well-being are well-documented. However, our study shows that the contribution to employability in the arts and humanities programs seems to lie in the learning opportunities for the acquisition of generic skills. In the light of the need to foster "21st-century skills" (Germaine et al., 2016; Suarta et al., 2017; Rios et al., 2020), our findings thus suggest that fields like arts, humanities, and educational science seem to offer important learning opportunities for students but are not limited to these fields of study.

Although our study enables a comparative evaluation of student and graduate data, it is also subject to limitations. So far, the scales have been applied at only one university. Therefore, we cannot yet make any statements about the broader application of the results to other contexts. Hopefully, similar studies of linked student and graduate surveys will be carried out in other countries, making it possible to compare the results of such studies with ours. Since the initial instrument (Braun and Brachem, 2015, 2017) was already tested and used in a Germany-wide graduate survey, we did not pretest our instruments. We are aware that there is a certain risk that the statements will be perceived differently by students than by graduates. Nevertheless, by using similar phrasing of the statements, cross-comparisons can be made between learning opportunities and labor market requirements for generic skills.

Another limitation is connected with the area of "physical performance," since it was captured with only one item. Despite this limitation, this area is nonetheless significant because the fields of study differ in how much physical labor is carried out by students during their studies. Future research might extend this scale by including more items. In addition, future research could use more advanced and person-centered methods to test our findings.

Even if the distribution of fields of study among the three profiles is significant, the distribution itself is not that evident. All ten fields of study show up in all three profiles, which is an indicator, that all study programs contribute to the acquisition of generic skills.

The results leave open the question of whether generic skills are embedded learning goals within higher education and whether the fields of study seek to promote the acquisition of generic skills, or if they are a mere by-product of the content of individual fields of study. To answer this question, it would be necessary to examine module manuals of the subjects in reference to their learning opportunities for generic skills.

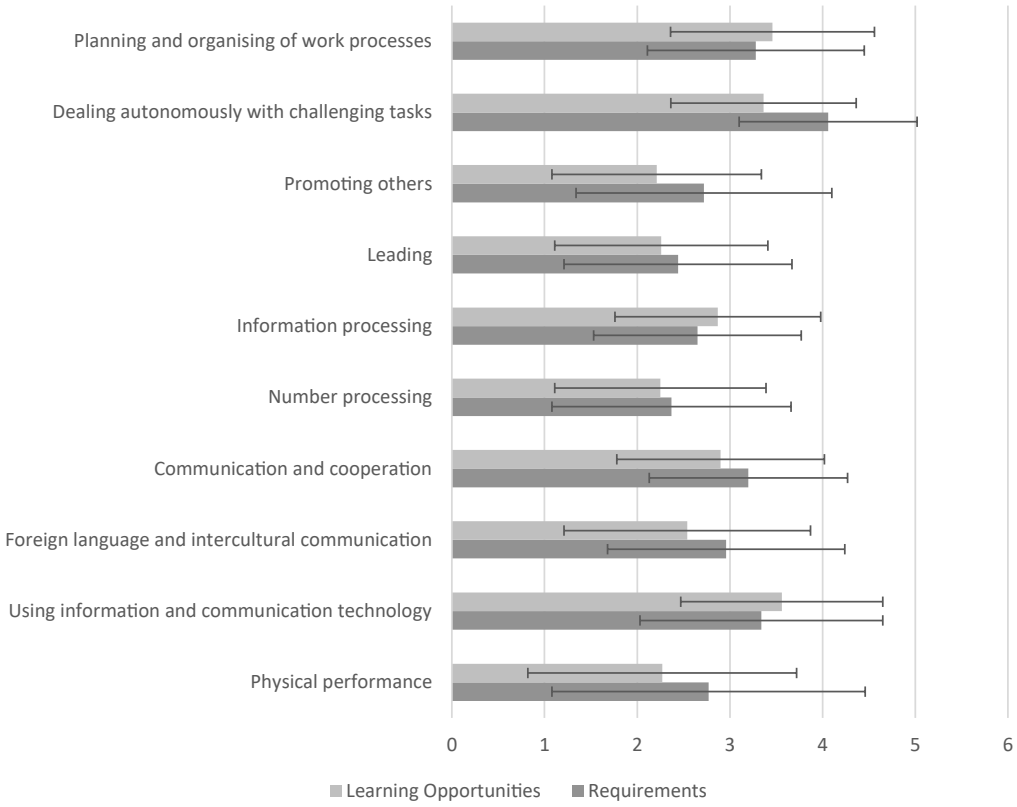


FIGURE 2
Perception of requirements in the view of graduates and learning opportunities in the view of students.

The findings regarding the descriptive comparison between the student and graduate assessments are quite restricted because of their explorative nature. Further research might build on these preliminary results to conduct more hypotheses-driven research. Nevertheless, the descriptive results provide the first indications about the fit between learning opportunities and labor market requirements for generic skills and can be used in the context of (re)accreditation for the individual subject groups.

The data were collected as part of the quality management at the university where we conducted our study. We are grateful for the permission that they granted us to use the data for our publications. Universities can draw on the gathered data and on our findings to assist in the development and administration of their programs; students and graduates can draw on them for their personal development. As part of our work, we always consider the ethical implications of our work and proceed accordingly, seeking to work to the best of our conscience.

The scales presented here are based on the job requirements approach, and therefore firmly grounded on a theoretical basis. They are nonetheless self-reported. Although this allowed an assessment of perceived learning opportunities, it did not offer a "hard" measurement of competences. Moreover, in this paper we examined only generic skills; we therefore cannot make any statements about the promotion of subject-specific knowledge, which can certainly be seen as a primary learning outcome of higher education.

Despite these limitations, the use of scales provides an empirical basis, especially for the area of quality management of study programs and teaching at universities. So far, mainly graduate data and retrospective assessments have been used to draw conclusions about the quality of the fit between study programs and the labor market. As noted above, study regulations and conditions may have changed several times in the interval between when the graduates pursued their studies and when they were surveyed about them. This time lag makes it difficult for those responsible for the organization and planning of the various study programs to draw on research to improve their programs. By using the scales presented here and applying them both to the student and the graduate surveys, a direct reference can be made between learning opportunities for generic skills that currently exist at institutions of higher education and those currently in demand on the labor market. Our experience shows that this empirical data can be used in a variety of ways,

both for the accreditation of study programs and for evidence-based curriculum development. We are not suggesting that the requirements reported by graduates should be integrated one-to-one into the curriculum. We do believe, however, that the scales facilitate planning and development. They make it possible to carry out an informed discussion about which generic skills are already being promoted and those which should be given greater consideration in the future.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

KL wrote the first draft of the manuscript and performed the statistical analysis. EB wrote sections of the manuscript. Both authors equally contributed to the design and the conception of this study, contributed to the article, and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Altonji, J. G., Arcidiacono, P., and Maurel, A. (2016). "The analysis of field choice in college and graduate school," in *Handbook of the economics of education*, 5th Edn, Vol. 5, eds E. A. Hanushek, S. Machin, and L. Woessmann (Amsterdam: Elsevier), 305–396.

Autor, D. H., and Handel, M. J. (2013). Putting tasks to the test: Human. *J. Labor Econ.* 31, 59–96. doi: 10.1086/669332

Backhaus, K., Erichson, B., and Weiber, R. (2015). *Fortgeschrittene multivariate analysemethoden. Eine Anwendungsorientierte Einführung*, 3rd Edn. Berlin: Springer.

Badcock, P. B. T., Pattison, P. E., and Harris, K. (2010). Developing generic skills through university study: A study of arts, science and engineering in Australia. *High Educ.* 60, 441–458. doi: 10.1007/s10734-010-9308-8

Baert, S., Cockx, B., and Verhaest, D. (2013). Overeducation at the start of the career: Stepping stone or trap? *Labour Econ.* 25, 123–140. doi: 10.1016/j.labeco.2013.04.013

Barrie, S. C. (2004). A research-based approach to generic graduate attributes policy. *High. Educ. Res. Dev.* 23, 261–275. doi: 10.1080/0729436042000235391

Bauer, T. K. (2002). Educational mismatch and wages: A panel analysis. *Econ. Educ. Rev.* 21, 221–229. doi: 10.1016/S0272-7757(01)00004-8

Beauducel, A., and Wittmann, W. W. (2005). Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. *Struct. Equ. Model.* 12, 41–75.

Berlingieri, F., and Erdsiek, D. (2012). *How relevant is job mismatch for German graduates? (ZEW discussion papers, 12-075)*. Mannheim: Universitätsbibliothek Mannheim.

Bligh, D. A. (2000). *What's the use of lectures?*, 2nd Edn. San Francisco, CA: Jossey-Bass Publishers.

BMRB Social Research (2006). *2006 skills survey technical report*. Available online at: https://sp.ukdataservice.ac.uk/doc/6004/mrdoc/pdf/6004userguide.pdf (accessed June 8, 2022).

Braun, E., and Brachem, J. (2017). "The labour market's requirement profiles for higher education graduates," in *Higher education transitions*, eds E. Kyndt, V. Donche, K. Trigwell, and S. Lindblom-Ylänne (London: Routledge), 219–237.

Braun, E. M. P., and Brachem, J. (2015). Requirements higher education graduates meet on the labor market. *Peabody J. Educ.* 90, 574–595.

Browne, M. W., and Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociol. Methods Res.* 21, 230–258.

Cassity, E., and Ang, I. (2006). Humanities–industry partnerships and the 'Knowledge Society': The Australian experience. *Minerva* 44, 47–63.

Crebert, G., Bates, M., Bell, B., Patrick, C., and Cragnolini, V. (2004). Developing generic skills at university, during work placement and in employment: Graduates' perceptions. *High. Educ. Res. Dev.* 23, 147–165. doi: 10.1080/0729436042000206636

Falk, S., Reimer, M., and Sarcletti, A. (2009). *Studienqualität, Kompetenzen und Berufseinstieg in Bayern: der Absolventenjahrgang 2004. Hochschulforschung und Hochschulplanung (Studien zur Hochschulforschung, 76)*. München: Bayerisches Staatsinst.

Felstead, A., Gallie, D., Green, F., and Zhou, Y. (2007). *Skills at work, 1986 to 2006*. Oxford: ESRC Centre on Skills, Knowledge and Organisational Performance.

Franzen, A. (2002). *Der Einstieg in den Arbeitsmarkt von Schweizer Hochschulabsolvent/innen. Eine empirische Analyse der Absolventenbefragungen 1981 bis 2001*. Neuchâtel: Bundesamt für Statistik (Statistik der Schweiz 15, Bildung und Wissenschaft).

Germaine, R., Richards, J., Koeller, M., and Schubert-Irastorza, C. (2016). Purposeful use of 21st century skills in higher education. *J. Res. Innov. Teach.* 9, 19–29.

Green, F. (2009). *The growing importance of generic skills*. London: UK Department for Children, Schools and Families.

Handel, M. J. (2020). *Job skill requirements: Levels and trends (MIT work of the future working paper series, 02)*. Cambridge: MIT Industrial Performance Center.

Heijke, H., Meng, C., and Ris, C. (2003). Fitting to the job: The role of generic and vocational competencies in adjustment and performance. *Labour Econ.* 10, 215–229. doi: 10.1016/S0927-5371(03)00013-7

Huber, C. R., and Kuncel, N. R. (2016). Does college teach critical thinking? A meta-analysis. *Rev. Educ. Res.* 86, 431–468. doi: 10.3102/0034654315605917

Hyytinen, H., Nissinen, K., Ursin, J., Toom, A., and Lindblom-Ylänne, S. (2015). Problematising the equivalence of the test results of performance-based critical thinking tests for undergraduate students. *Stud. Educ. Eval.* 44, 1–8. doi: 10.1016/j.stueduc.2014.11.001

International Labour Organization (2014). *Global employment trends 2014. Risk of a jobless recovery?* Geneva: International Labour Office.

Jenkins, N. (2015). *Alarm over huge cuts to humanities and social sciences at Japanese universities*. TIME USA. Available online at: https://time.com/4035819/japan-university-liberal-arts-humanities-social-sciences-cuts/ (accessed September 16, 2015).

Johnson, E., Herd, S., and Tisdall, J. (2002). Encouraging generic skills in science courses. *Electron. J. Biontechnol.* 5, 22–23.

Jones, A. (2009b). Redisciplining generic attributes: The disciplinary context in focus. *Stud. High. Educ.* 34, 85–100. doi: 10.1080/03075070802602018

Jones, A. (2009a). Generic attributes as espoused theory: The importance of context. *High. Educ.* 58, 175–191.

Kember, D., Leung, D. Y. P., and Ma, R. S. F. (2007). Characterizing learning environments capable of nurturing generic capabilities in higher education. *Res. High. Educ.* 48, 609–632.

Klaukien, A., Ackermann-Piek, D., Helmschrott, S., Rammstedt, B., Solga, H., and Wößmann, L. (2013). "Grundlegende Kompetenzen auf dem Arbeitsmarkt," in *Grundlegende Kompetenzen Erwachsener im internationalen Vergleich*, ed. B. Rammstedt (Münster: Waxmann Verlag (Erwachsenenbildung 2013/14)), 127–166.

Kracke, N., and Rodrigues, M. (2020). A task-based indicator for labour market mismatch. *Soc. Indic. Res.* 149, 399–421. doi: 10.1007/s11205-019-02261-2

Leuze, K., and Strauß, S. (2008). "Berufliche Spezialisierung und Weiterbildung – Determinanten des Arbeitsmarkterfolgs von GeisteswissenschaftlerInnen," in *Findigkeit in unsicheren Zeiten*, eds H. Solga, D. Huschka, P. Eilsberger, and G. G. Wagner (Opladen: Verlag Barbara Budrich), 67–94.

Levels, M., van der Velden, R., and Di Stasio, V. (2014). From school to fitting work: How education-to-job matching of European school leavers is related to educational system characteristics. *Acta Sociol.* 57, 341–361. doi: 10.1177/0001699314552807

Li, I. W., Harris, M., and Sloane, P. J. (2018). Vertical, horizontal and residual skills mismatch in the Australian graduate labour market. *Econ. Rec.* 94, 301–315. doi: 10.1111/1475-4932.12413

Martin, R., Maytham, B., Case, J. M., and Fraser, D. M. (2005). Engineering graduates' perceptions of how well they were prepared for work in industry. *Eur. J. Eng. Educ.* 30, 167–180.

Matthes, B., and Christoph, B. (2011). *Nationales Bildungspanel. Großpilot E8 Feldversion. Version 1.02*. Bamberg: LIfBi.

Matthes, B., Christoph, B., Janik, F., and Ruland, M. (2014). Collecting information on job tasks—an instrument to measure tasks required at the workplace in a multi-topic survey. *J. Labour Mark. Res.* 47, 273–297. doi: 10.1007/s12651-014-0155-4

McGuinness, S. (2006). Overeducation in the labour market. *J. Econ. Surv.* 20, 387–418. doi: 10.1111/j.0950-0804.2006.00284.x

McGuinness, S., and Sloane, P. J. (2011). Labour market mismatch among UK graduates: An analysis using REFLEX data. *Econ. Educ. Rev.* 30, 130–145. doi: 10.1016/j.econedurev.2010.07.006

Morgado, A., Sequeira, T. N., Santos, M., Ferreira-Lopes, A., and Reis, A. B. (2016). Measuring labour mismatch in Europe. *Soc. Indic. Res.* 129, 161–179. doi: 10.1007/s11205-015-1097-0

Morgeson, F. P., and Humphrey, S. E. (2006). The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *J. Appl. Psychol.* 91, 1321–1339. doi: 10.1037/0021-9010.91.6.1321

O*NET (2021). *Work activities questionnaire*. Available online at: https://www.onetcenter.org/dl_files/MS_Word/Generalized_Work_Activities.pdf (accessed June 8, 2022).

OECD (2011). *Skills matter: Further results from the survey of adult skills (OECD skills studies)*. Paris. OECD Publishing.

OECD (2013a). *OECD skills outlook 2013: First results from the survey of adult skills*. Paris: OECD Publishing.

OECD (2013b). *Technical report of the survey of adult skills (PIAAC)*. Available online at: https://www.oecd.org/skills/piaac/_Technical%20Report_17OCT13.pdf (accessed June 8, 2022).

OECD (2019). *Skills matter: Additional results from the survey of adult skills*. Paris: OECD Publishing.

Olmos-Peñuela, J., Benneworth, P., and Castro-Martínez, E. (2015). Are sciences essential and humanities elective? Disentangling competing claims for humanities' research public value. *Arts Humanit. High. Educ.* 14, 61–78. doi: 10.1177/1474022214534081

Paviotti, G. (2020). *Regional universities' and pedagogy*. Cham: Springer International Publishing.

Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P. R., Fleishman, E. A., Levin, K. Y., et al. (2001). Understanding work using the Occupational Information Network (O* NET): Implications for practice and research. *Pers. Psychol.* 54, 451–492. doi: 10.1093/workar/wax016

Preston, A. (2015). *The war against humanities at Britain's universities*. The Guardian. Available online at https://www.theguardian.com/education/2015/mar/29/war-against-humanities-at-britains-universities (accessed March 02, 2022)

Quintini, G. (2011). *Right for the job. Over-qualified or under-skiled? OECD social employment and migration working papers (120)*. Paris: OECD Publishing.

Rios, J. A., Ling, G., Pugh, R., Becker, D., and Bacall, A. (2020). Identifying critical 21st-century skills for workplace success: A content analysis of job advertisements. *Educ. Res.* 49, 80–89. doi: 10.3102/0013189X19890600

Robst, J. (2007). Education and job match: The relatedness of college major and work. *Econ. Educ. Rev.* 26, 397–407. doi: 10.1016/j.econedurev.2006.08.003

Rychen, D. S., and Salganik, L. H. (2003). "A holistic model of competence," in *Key competencies. For a successful life and a well-functioning society*, eds D. S. Rychen and L. H. Salganik (Cambridge, MA: Hogrefe & Huber), 41–63.

Sattinger, M., and Hartog, J. (2013). Nash bargaining and the wage consequences of educational mismatches. *Labour Econ.* 23, 50–56. doi: 10.1016/j.labeco.2013.03.002

Smith, C., and Bath, D. (2006). The role of the learning community in the development of discipline knowledge and generic graduate outcomes. *High. Educ.* 51, 259–286. doi: 10.1007/s10734-004-6389-2

Statistisches Bundesamt (2021). *Studierende an hochschulen. Fächersystematik. Destatis (Fachserie 11, Reihe 4.1)*. Wiesbaden: Statistisches Bundesamt.

Suarta, I. M., Suwintana, I. K., Sudhana, I. G. P. F. P., and Hariyanti, N. K. D. (2017). "Employability skills required by the 21st century workplace: A literature review of labor market demand," in *Proceedings of the international conference on technology and vocational teachers (ICTVT 2017). Yogyakarta, Indonesia, 28.09.2017–28.09.2017* (Paris: Atlantis Press), 337–342.

Tynjälä, P., Slotte, V., Nieminen, J., Lonka, K., and Olkinuora, E. (2006). "From university to working life: graduates' workplace skills in practice," in *Higher education and working life: Collaborations, confrontations and challenges*, eds P. Tynjälä, J. Välimaa, and G. Boulton-Lewis (Bingley: Emerald Group Publishing Limited), 73–88. doi: 10.1111/ecc.13066

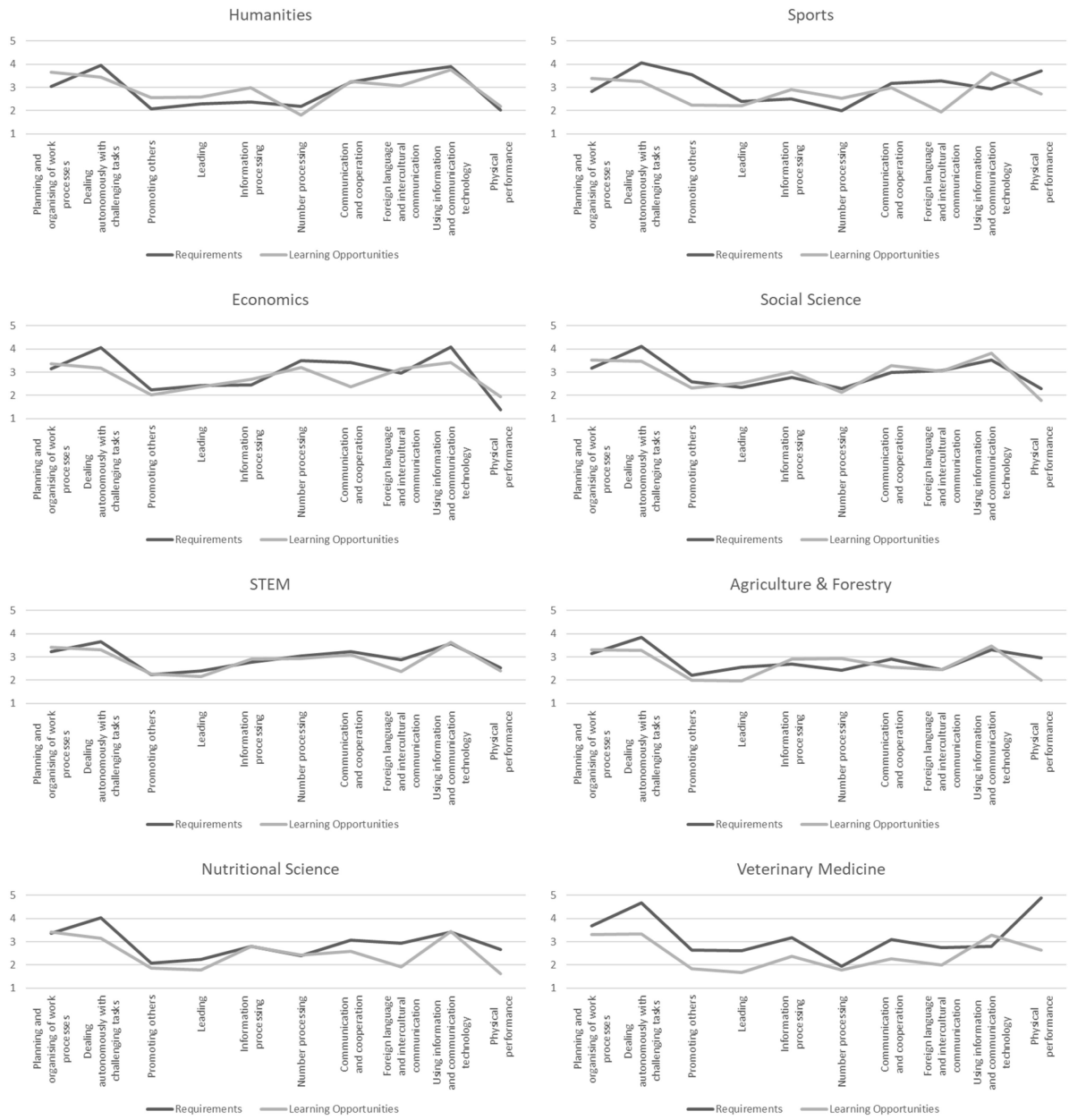
Ursin, J., Hyytinen, H., and Silvennoinen, K. (eds) (2021). *Assessment of undergraduate students' generic skills in Finland. Findings of the Kappas! Project*. Helsinki: Ministry of Education and Culture.

Verhaest, D., Sellami, S., and van der Velden, R. (2017). Differences in horizontal and vertical mismatches across countries and fields of study. *Int. Labour Rev.* 156, 1–23. doi: 10.1111/j.1564-913X.2015.00031.x

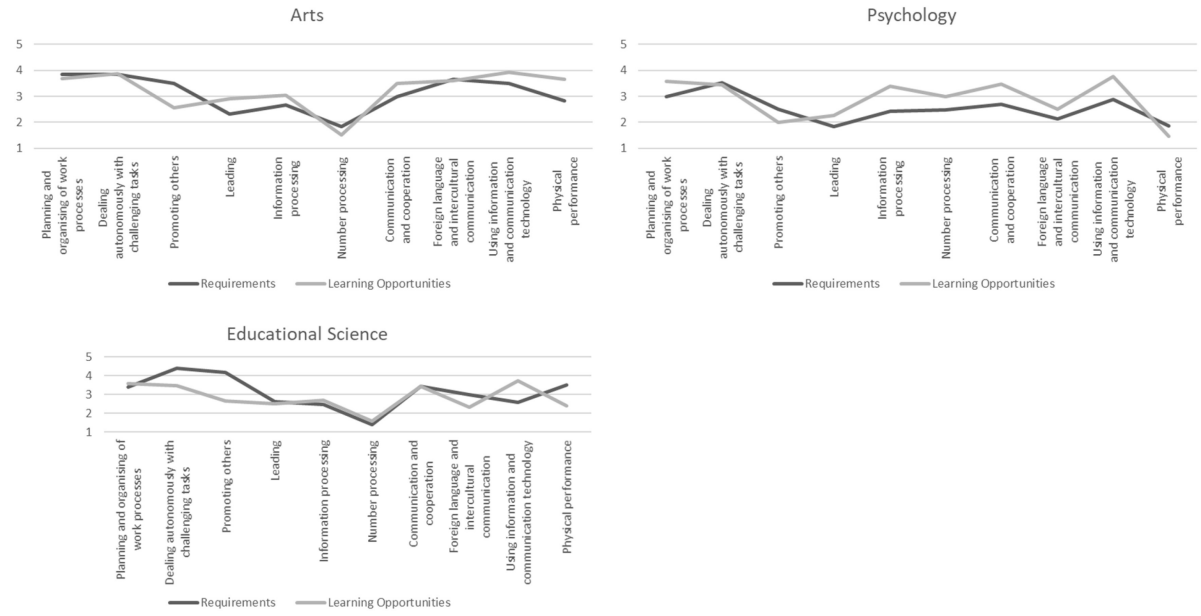
Virtanen, A., and Tynjälä, P. (2019). Factors explaining the learning of generic skills: A study of university students' experiences. *Teach. High. Educ.* 24, 880–894. doi: 10.1080/13562517.2018.1515195

Wissenschaftsrat (2013). *Perspektiven des deutschen Wissenschaftssystems*. Braunschweig: Wissenschaftsrat.

Appendix



APPENDIX FIGURE 1
(Continued)



APPENDIX FIGURE 1

Perception of requirements in the view of graduates and learning opportunities in the view of students by field of study.

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 07 November 2022
DOI 10.3389/feduc.2022.914857

OPEN ACCESS

EDITED BY
Edith Braun,
Justus-Liebig
Universität, Germany

REVIEWED BY
Tobias Jenert,
University of Paderborn,
Germany
Heidi Hyytinen,
University of Helsinki Center for University Teaching and Learning, Finland

*CORRESPONDENCE
Marie-Theres Nagel
marie.nagel@uni-mainz.de

SPECIALTY SECTION
This article was submitted to Higher Education, a section of the journal Frontiers in Education

RECEIVED 07 April 2022
ACCEPTED 05 October 2022
PUBLISHED 07 November 2022

CITATION
Nagel M-T, Zlatkin-Troitschanskaia O and Fischer J (2022) Validation of newly developed tasks for the assessment of generic Critical Online Reasoning (COR) of university students and graduates. *Front. Educ.* 7:914857. doi: 10.3389/feduc.2022.914857

Validation of newly developed tasks for the assessment of generic Critical Online Reasoning (COR) of university students and graduates

Marie-Theres Nagel*, Olga Zlatkin-Troitschanskaia and Jennifer Fischer

Department of Business and Economics Education, Johannes Gutenberg University, Mainz, Germany

In recent decades, the acquisition of information has evolved substantially and fundamentally affects students' use of information, so that the Internet has become one of the most important sources of information for learning. However, learning with freely accessible online resources also poses challenges, such as vast amounts of partially unstructured, untrustworthy, or biased information. To successfully learn by using the Internet, students therefore require specific skills for selecting, processing, and evaluating the online information, e.g., to distinguish trustworthy from distorted or biased information and for judging its relevance with regard to the topic and task at hand. Despite the central importance of these skills, their assessment in higher education is still an emerging field. In this paper, we present the newly defined theoretical-conceptual framework Critical Online Reasoning (COR). Based on this framework, a corresponding performance assessment, Critical Online Reasoning Assessment (CORA), was newly developed and underwent first steps of validation in accordance with the Standards for Educational and Psychological Testing. We first provide an overview of the previous validation results and then expand them by including further analyses of the validity aspects "internal test structure" and "relations with other variables". To investigate the internal test structure, we conducted variance component analyses based on the generalizability theory with a sample of 125 students and investigated the relations with other variables by means of correlation analyses. The results show correlations with external criteria as expected and confirm that the CORA scores reflect the different test performances of the participants and are not significantly biased by modalities of the assessment. With these new analyses, this study substantially contributes to previous research by providing comprehensive evidence for the validity of this new performance assessment that validly assesses the complex multifaceted construct of critical online reasoning among university students and graduates. CORA results provide unique insights into the interplay between features of online information acquisition and processing, learning environments, and the cognitive and metacognitive requirements for critically reasoning from online information in university students and young professionals.

KEYWORDS

critical online reasoning, online information, web search, validation, performance assessment, higher education

Introduction

The digital age has transformed learning in higher education as well as the learning materials accessible to students (Ali, 2020; Banerjee et al., 2020). The acquisition and use of information has evolved substantially in recent decades and also fundamentally affects students' learning (Boh Podgornik et al., 2016; Brooks, 2016; Maurer et al., 2020). University students nowadays prefer the Internet to traditional textbooks for information acquisition; moreover, in the recently increasingly prevalent digital teaching and learning contexts, students use not professionally produced learning resources, found by eclectically browsing the web, more often and ubiquitously than the recommended OER. The Internet has therefore become one of the most important sources of information for learning; not only for the preparation of papers or presentations but also when studying for exams (Brooks, 2016; Newman and Beetham, 2017; Maurer et al., 2020). The World Wide Web provides a flexible learning resource while also accelerating the dissemination and processing of information and knowledge (Braasch et al., 2018; Weber et al., 2019; Maurer et al., 2020). However, learning with freely accessible online resources also presents challenges (Qiu et al., 2017; Ciampaglia, 2018). Since content can be freely distributed on the Internet, vast amounts of unstructured, untrustworthy, inaccurate, or biased information are just as readily available to learners as credible, verified information (Walton et al., 2020). Dealing with the vast amount of information available online, on a platform characterized by low publication barriers and deficiently established quality standards, requires students to be critically evaluative (Liu et al., 2014; Tribukait et al., 2017). Thus, the ever-changing information and learning environment has profound consequences for the imparting of knowledge in higher education (Harrison and Luckett, 2019; Weber et al., 2019; Maurer et al., 2020). To competently use and successfully learn from the information and resources openly accessible on the Internet, students must be able to critically search, select, review, and evaluate online information and sources based on relevant quality criteria (Sendurur, 2018; Molerov et al., 2020; Nagel et al., 2020). In the context of increasingly digital and self-directed teaching and learning processes in higher education, the successful use of digital media and competent, critical use of online information constitutes one of the most important student skills for successful study (Harrison and Luckett, 2019; Molerov et al., 2020), as has also been emphasized by the most recent research review (Osborne et al., 2022). This classifies it as a so-termed generic skill, which college graduates are expected to develop to operate successfully as professionals and responsible citizens of democratic societies (Binkley et al., 2012; National Research Council, 2012; Shavelson et al., 2018; Virtanen and Tynjälä, 2018; Zlatkin-Troitschanskaia et al., 2021a). In addition to professional knowledge, such skills include quantitative reasoning, critical literacy and thinking, ethical and moral reasoning, and written and oral communication that college graduates can draw upon to address life's everyday judgments, decisions, and challenges. As a current literature review indicates, nowadays, searching, evaluating, selecting, and using high-quality online information have additionally become generic skills important for successfully studying in higher education (Zlatkin-Troitschanskaia et al., 2021b).

So far, the related subskills have been assessed based on various theoretical constructs, such as "multiple-source use" (MSU; Braasch et al., 2018; Hahnel et al., 2019), "information trust" (Johnson et al., 2016; Leeder, 2019), and "web credibility" (Flanagin and Metzger, 2017; Herrero-Diz et al., 2019). While providing important insights into the individual subskills, these approaches have not yet systematically focused on the interplay between features of online information acquisition and learning environments and the (cognitive) requirements for critical reasoning from online information (Goldman and Brand-Gruwel, 2018). Another relevant research strand focusses on the aspect of communicating the selected and critically evaluated information to answer an initial question, as such communication skills are particularly needed in later (professional) life (Chan et al., 2017; Braun, 2021). Lawyers or physicians, for example, not only have to compile various, reliable pieces of information on individual cases and draw conclusions from them, but also regularly exchange information with clients and patients in this process (e.g., Korn, 2004; Aspegren, and LØnberg-Madsen, P., 2005).

A recent review consolidating information problem-solving and multiple source use approaches highlights existing desiderata in examining how evaluated information is used in more advanced analytical reasoning processes and what role the characteristics of information play in reasoning (Goldman and Brand-Gruwel, 2018). For instance, while students may differ in their judgment of the credibility of a source, drawing invalid inferences is generally wrong epistemically and indicates poor (online) reasoning skills. In addition, most of the tests used so far to measure these subskills have a close-ended format, thus covering only limited aspects of dealing with online information use and, in particular, failing to measure the actual reasoning process, and underlying procedural skills (Ku, 2009; Desai and Reimers, 2019). In addition, these procedures no longer do justice to the current efforts of higher education institutions regarding the measurement of students' competencies, which increasingly focus on a holistic representation of students' capabilities to act (Shavelson et al., 2019).

To make these crucial student skills pertaining to the online information environment empirically measurable and to be able to specifically promote them, a new theoretical-conceptual framework of *Critical Online Reasoning* (COR) was developed (see section "Conceptual background"; for details, see Molerov et al., 2020). COR describes the abilities of searching, selecting, accessing, processing, and critically reasoning from online information, e.g., to solve a particular generic or domain-specific problem or task (for details, see Molerov et al., 2020). This involves critically distinguishing trustworthy from untrustworthy information and making argumentative and coherent judgments based on credible and relevant information from the online environment. Based on this conceptual framework, a COR performance Assessment (CORA) was newly developed and underwent initial validation (Molerov et al., 2020; Zlatkin-Troitschanskaia et al., 2021a). Based on the COR construct definition, CORA includes various authentic situational tasks in the online media environment, i.e., the real Internet, to objectively and validly assess students' COR skills in a realistic performance assessment. This holistic assessment measures all required skill (sub-)dimensions and their interplay instead of only individual facets as would be the case, for example, with closed-ended tests (Davey et al., 2015; for a CORA task example, see Figure 1).

When measuring students' COR skills through CORA, validity is one of the key quality criteria for the reliable interpretation of students' test results. The Standards for Educational and Psychological Testing (hereafter referred to as "AERA Standards") provide criteria for the reliable validation of educational tests (AERA, APA, and NCME, 2014). According to the AERA Standards, five aspects should be analyzed during validation and various sources of information should be used as evidence. The aspects to be analyzed are "test content," "task-and test-response processes," "internal structure of a test," "interrelationships with other variables," and "consequences of testing" (for details, see AERA, APA, and NCME, 2014). Therefore, the focus and central contribution of this paper is to present the comprehensive, multi-perspective and in-depth validation of the CORA as a novel performance-based test of generic student skills in higher education.

To validate the CORA tasks and interpret the test scores, initial validation steps have already been carried out:

1. Validity evidence regarding the CORA content was obtained through expert interviews and expert ratings of the CORA tasks (for details, see Molerov et al., 2020).
2. Validity evidence regarding the task response processes of the test takers was analyzed by Schmidt et al. (2020) on the basis of log files and eye-tracking data including gaze duration and fixations.
3. Initial validity evidence on the correlations with other variables was obtained by Nagel et al. (2020) through analyzing the extent to which participants' web search behavior—specifically, the number and type of web pages accessed as well as the quality of the content on the web pages—is related to better task performance and thus to a more critically-reflective use of online information.

In this paper, further validation of the CORA tasks focusing on the two criteria 'internal structure of the test' and 'correlations with other variables' is presented and critically discussed. In this way, further validity aspects not yet considered are systematically and thoroughly investigated according to the AERA standards to obtain a comprehensive overview of the validity of the CORA. The results of the analyses are combined with the validity evidence outlined above to provide a comprehensive validity assessment of the new COR Assessment.

In Chapter 2, the definition of the COR construct, which serves as a basis for an appropriate interpretation of the CORA test results (Molerov et al., 2020), is explained in more detail. In addition, the COR Assessment framework is presented, including a sample task. Chapter 3 explains the validation approach of CORA, which is based on the model of argumentation-based validation of test score interpretations. According to the

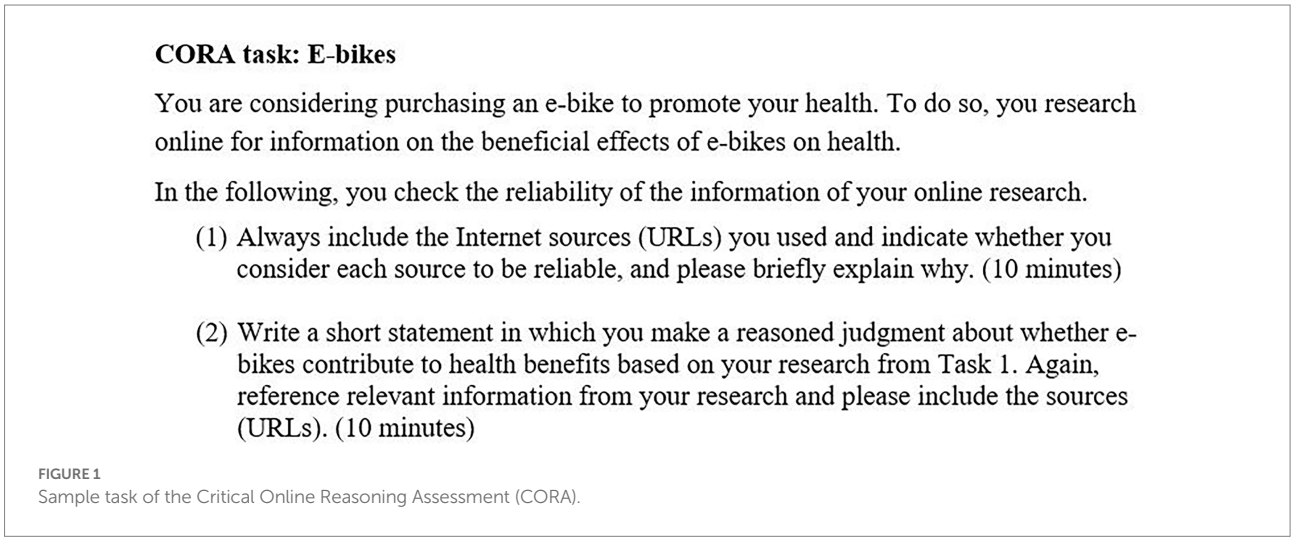


FIGURE 1
Sample task of the Critical Online Reasoning Assessment (CORA).

argumentation-based validation process (Mislevy et al., 2012), we briefly summarize the results of the previous validation studies on the content validity (section Content validity) and validity of task response processes (section Validity of task reponse processes) of CORA, before the newly obtained validity evidence is presented (sections Internal test structure and Relations with other variables) and integrated with the previous validation work for CORA. Chapter 4 provides a critical discussion of the results including the limitations of the work and an outlook on the further research.

Conceptual background

The COR construct definition

To harness the potential of the Internet for learning, students require a variety of information acquisition and processing skills, which have been previously summarized as such a broad literacy construct as digital literacy (Reddy et al., 2020; Park et al., 2021), which is also related to media literacy (Koltay, 2011), information literacy (Limberg et al., 2012; Sanders et al., 2015; Walton et al., 2020), and computer literacies [e.g., information and communication technology (ICT) literacy, computer and information literacy (CIL); Siddiq et al., 2016; Makhmudov et al., 2020; see also, e.g., studies on multimedia learning, Mayer, 2009]. Particularly for students of higher education, current research presumes basic computer knowledge (Rammstedt, 2013; Schlebusch, 2018) as well as multimedia (Naumann et al., 2001; Goldhammer et al., 2013) and general Internet skills, which are required for self-directed online learning, a given (Rammstedt, 2013). However, numerous studies outline substantial deficits in students' Internet-based learning in higher education that can hinder their study success. Based on prior research, we are going beyond such broad literacy and general ability concepts, and focus more specifically on modeling and validly assessing actual online information acquisition and processing skills, and in particular critical reasoning based on this online information. When modeling COR, we particularly draw on extended information problem-solving (IPS-I) models (Brand-Gruwel et al., 2009; Huang et al., 2019; Whitelock-Wainwright et al., 2020) to distinguish and describe the main processes involved in self-directed online learning. Thereby, we further expand these models by focusing on processes of argumentation as well as communication, which are not only important for students' academic success but also key requirements that higher education graduates encounter on the labor market (Braun and Brachem, 2018). These skills can be summarized under the REAS-facet: Reasoning based on Evidence, Argumentation and Synthesis. Therefore, the COR model describes students' key generic skills not only for searching, evaluating, and selecting—as in IPS-I models–but also additional processes including analyzing, synthesizing, and reasoning from (high-quality) online information, while self-directedly engaging with (more or less domain-specific) content or working toward course-related learning goals, e.g., outside of classrooms (e.g., preparing an essay at home). We differentiate between two main requirement areas for COR processes: generic and domain-specific, e.g., within particular study domains like Medicine or Law (for details, see Molerov et al., 2020). The focus of the analyses presented here is particularly on the generic COR skills required for researching more general topics that are not specifically related to a particular domain (for a differentiation between generic and domain-specific requirement areas for COR in higher education, see Zlatkin-Troitschanskaia et al., 2021a).

In our prior research, we theoretically analyzed the links and overlaps between the existing concepts and models for assessing students' skills related to COR (for more details on these specific concepts, underlying constructs, and particularly overlaps and distinctions, see our differentiated descriptions in Molerov et al., 2020). Going beyond established abovementioned "literacy" concepts and constructs like digital literacy and multiple source use, we especially draw on the triad model of critical alertness, reflection, and analysis (Oser and Biedermann, 2020). Thereby, we particularly focus on how students analytically reason from as well as justify and critically reflect on online information they used for their higher education studies and infer from and weight arguments and (covert) perspectives of (partly conflicting) sources and information pieces. Based on this theoretical rationale, we specify a set of skills assumed crucial for the acquisition and use of high-quality online information for learning in higher education, which we term Critical Online Reasoning (for details, see Molerov et al., 2020). Thereby, in addition to the abovementioned models and concepts, we also particularly draw on the U.S.-established concept of civic online reasoning. This concept describes the ability to successfully deal with online information and distinguish, for instance, reliable and trustworthy sources of information from biased and manipulative ones (Wineburg et al., 2016). While this concept focuses especially on the handling of online information on political and social topics in particular, our approach of COR has been expanded to encompass all cross-domain topics relevant for students' learning in higher education and beyond. In addition, we further substantially expanded the concept of civic online reasoning as well as the information problem-solving models by Brand-Gruwel et al. (2009), to cover the whole process of searching, evaluating, selecting, analyzing, synthesizing, and reasoning from online information. In doing so, we also specifically incorporated a new reasoning facet, described as Reasoning based on Evidence, Argumentation, and Synthesis (for details, see Molerov et al., 2020).

To sum up, the COR concept leans closely on previous process and phase models of (online) information search, selection, and evaluation, in particular the information problem-solving models (Brand-Gruwel et al., 2009; Huang et al., 2019; Whitelock-Wainwright et al., 2020). Thereby, we also consider insights from related "web credibility" research, especially on

multiple-source use and multiple-source comprehension (Braasch et al., 2018; Goldman and Brand-Gruwel, 2018; Hahnel et al., 2019). We expand the modeling of students' information use in self-directed learning by adding a new critical reasoning component, i.e., Reasoning based on Evaluation, Argumentation, and Synthesis (REAS). In addition, we also integrate a metacognitive regulative component, i.e., Metacognitive Activation (MCA) skills, that helps students decide when to employ COR skills (e.g., to initiate a critical evaluation; for more details, see Molerov et al., 2020).

Based on this conceptual work, to model and measure COR according to international testing standards by AERA, APA, and NCME (2014) in an evidence-centered design (Zieky, 2014; Mislevy, 2017), we specified its construct definition with three overarching and overlapping cognitive facets:

1. online information acquisition skills (OIA), e.g., selecting search engines or databases, specifying search queries;
2. critical information evaluation (CIE) skills, e.g., evaluating website credibility based on cues; and
3. reasoning skills, e.g., using evidence to generate and justify a valid argument based on a synthesis of accessed information (REAS), including accounting for common errors and biases as well as considering (contradictory) arguments and (covert) perspectives from (possibly conflicting) sources and information.

In addition, metacognitive (MCA) skills regulate the state-specific and situation-specific activation, continuation, and conclusion of COR process within the encompassing information acquisition context, e.g., recognizing the need to use COR in learning-related contexts.

Based on this definition, we established COR as an operationalizable, multifaceted construct of students' (meta) cognitive skills for goal-oriented and competent use of online information focusing on study-related contexts in higher education (for details, see Molerov et al., 2020).

The COR assessment framework

Methodologically, recent assessment research shows that tests with a closed-ended format are limited when it comes to validly measuring (meta)cognitive higher-order skills such as COR (e.g., Braun et al., 2020). In addition, they no longer do justice to the more recent efforts at universities to ensure the validity of testing procedures, which increasingly aim to holistically measure students' capabilities to act (Shavelson et al., 2019). Closed-ended tests generally have a limited ecological validity as they fail to measure the procedural skills underlying the processing of (online) information used for learning, and, evidentially, students struggle to transfer the measured skills to more authentic, real-life situations (Ku, 2009; Davey et al., 2015; Desai and Reimers, 2019). It is thus evident that such complex, higher-order skill construct as COR can be more validly measured through performance assessments (Shavelson et al., 2019) that simulate the online information environment and adequately reflect the formal and informal learning contexts and conditions students of higher education experience in real life. The focus on the online information environment is therefore, following the tradition of measuring higher-order cognitive skills by means of performance assessments (Braun and Brachem, 2018; Shavelson et al., 2019; Braun et al., 2020), reflected in task scenarios that employ real websites and Internet searches, including sources, platforms, and services that are typical for current online media.

Since designing and developing new performance assessment tasks is particularly resource-intensive and time-consuming, we first looked for existing assessments, which could be possibly adapted and used to validly measure COR skills. In the past, therefore, we tried to measure COR using an adaptation of an Internet-based assessment developed and validated in the United States by the Stanford History Education Group (SHEG) to assess the abovementioned recently established concept of "civic online reasoning" at the middle school, high school, and college level (Wineburg et al., 2018). It is an innovative holistic assessment of how students evaluate online information and sources, containing short evaluation prompts, real websites, and an open Internet search (Wineburg et al., 2016; Wineburg and McGrew, 2016). The Stanford History Education Group asked students, for example, to evaluate the credibility of information on political and social issues of mostly U.S.-centric civic interest and to justify their judgment, also citing web sources as evidence (Wineburg and McGrew, 2019).

Based on preliminary validation, however, we further developed and expanded the COR assessment framework. Since an adaptation of this US assessment for the German university context was not feasible due to fundamental differences between the systems of higher education in the two countries, the conceptual-theoretical framework was modified and expanded, resulting in the new construct definition of Critical Online Reasoning described above (Section "The COR construct definition"; for more details, see Molerov et al., 2020; Zlatkin-Troitschanskaia et al., 2021a). In this process, a corresponding test definition was developed that provided the basis for the design of new CORA tasks with new scenarios as well as corresponding scoring rubrics to rate students' responses to the new tasks (for the description of the assessment and the ratings, see Section "Method and design").

Our newly developed COR performance assessment allows for validly measuring all theoretically defined COR facets (see the section "The COR construct definition") as we seek to demonstrate with the comprehensive validation presented in this paper.

Validity results

When developing the new COR assessment, the evidence-centered design (ECD) approach of Mislevy (2017) and the

Standards for Educational and Psychological Testing of AERA, APA, and NCME (2014) were followed to ensure the development of a valid assessment from the very beginning (see section "Conceptual background"). Consequently, as part of the CORA development, we also developed a student model (based on the construct definition), as well as a task model and an interpretive model (based on the test definition), as—according to the evidence-centered design approach—the alignment of these models is necessary for designing valid assessments (Mislevy, 2017). We also followed the standards according to AERA, APA, and NCME (2014) with regard to test development, scoring, and test quality assurance, in particular by conducting initial validity tests during the development of CORA (Molerov et al., 2019, 2020). These were systematically complemented by analyses of the different types of validity outlined in the following.

Content validity

Molerov et al. (2020) conducted a qualitative evaluation of CORA according to the standards of AERA, APA, and NCME (2014), with focus on the task content, i.e., analyzing the coverage of the theoretically derived COR construct facets by the tasks and the suitability of the requirements and content of the newly developed assessment and corresponding scoring approach for higher education in Germany. For this purpose, they conducted an analysis of the task content by means of 12 semistructured interviews with experts in the fields of computer-based performance assessments in higher education, media studies (focusing on online source evaluation or media literacy), linguistics, and cultural studies, which were then analyzed by means of content analyses.

The experts (1) confirmed that the CORA tasks measure the generic COR ability, (2) supported the assumption that CORA measures test participants' personal construct-relevant abilities in terms of the defined construct definition, and (3) concurred that no specific domain knowledge is required to complete the tasks. The experts also recommended to expand the scope of the assessment, as it was observed that the tasks might be too difficult for first-year students. In addition, some experts referred to the problem that participants' prior knowledge, interest, beliefs, or (political) attitudes in terms of the task topic could influence their CORA performance.

The additional content analysis confirmed that the assessment and corresponding scoring scheme included two different types of CORA tasks, each prioritizing a different COR facet (online information acquisition and critical information evaluation; Molerov et al., 2020, p. 20). To implement the indications of these analyses, a task format focusing more explicitly on the reasoning skills facet should be included for future assessments (Molerov et al., 2020, p. 20). Consequently, the tasks were expanded by two subtasks each, with a processing time of 10 min per subtask (see section Conceptual background).

Validity of task response processes

In a second validation approach focusing on the validity of task response processes, Schmidt et al. (2020) investigated how test participants' cognitive processes during task-solving can be described and to what extent certain empirically distinct patterns exist in the participants' task-and test-solving processes in relation to COR abilities. Therefore, their test-taking process data were collected through verbalizations, eye movements, response times, and computer clicks during the processing of the CORA tasks. Subsequently, Schmidt et al. operationalized the COR construct in two dimensions: At the level of COR ability, which is represented by the score in the CORA tasks (task performance), and at the level of process performance, which is indicated by gaze fixations and response times in the log files (online information processing).

The results showed that better process performance is associated with significantly higher scores, indicating a relationship between participants' process performance and task performance. Through an analysis of test-taking processes, the two distinct patterns of *avoidance strategy* and *strategic information processing* were identified during CORA task-solving. Participants using the avoidance strategy exhibited both poorer process performance and poorer task performance, i.e., they spent most of their time on only one web page, resulting in many fixations that were all focused on one specific process step. In contrast, participants using strategic information processing showed better performance and more intensive processing of online information through a larger number of (total) process steps, which was in line with the theoretical assumption for CORA (for details, see Schmidt et al., 2020).

Internal test structure

Theoretical background

According to the argumentative validation process following AERA, APA, and NCME (2014), evidence for the validity of the CORA scores and their interpretation could already be shown regarding the CORA content and the test takers' task response processes; initial evidence could also be obtained for correlations with other variables. The assessment's internal structure is also an important validity aspect, since analyses thereof can "indicate the degree to which the relationships among test tasks and test components conform to the construct on which the proposed test score interpretations are based" (AERA, APA, and NCME, 2014, p. 13). A performance assessment such as CORA, which includes a free Internet search and open-ended written answers that are evaluated by raters, differs fundamentally from classical test procedures with regard to its structure. Therefore, analysis methods according to classical test theory such as task analyses (e.g., test–retest reliability or internal consistency coefficients) are not suitable for this assessment format as they do not comprehensively take into account the complexity of various

possible influencing factors that are incorporated in performance assessments in contrast to conventional closed-ended assessments (Cronbach et al., 1972); for more details, see also Shavelson and Webb (1981) and Shavelson et al. (1989). Following Shavelson and Webb (1981), within the framework of *Generalizability Theory,* it is possible to sufficiently take into account the specifics of performance assessments. *Generalizability Theory* distinguishes between different components of the assessment, so-called *facets*, which can exert an influence on the test scores both individually and in interaction (Cronbach et al., 1963). In CORA, such facets are, in addition to test takers' varying COR abilities and other individual characteristics, certain characteristics of the tasks used (e.g., task topic, format, formulation, or time limits) and effects by the raters, which can also exert a systematic influence and thus affect the test results (Goldman and Brand-Gruwel, 2018; Solano-Flores, 2020).

While certain influences on the test scores are desirable, in particular those of participants' differing COR abilities or intentional variation of task difficulty, (uncontrolled) influences, for example those of rater effects, should be minimized. In the context of validating the CORA tasks, it should therefore be determined which influences the individual facets of the assessment exert on the scores and how they may interact with each other. The variance decomposition method used in this study allows for the analysis of the influencing factors across different CORA tasks (Jiang, 2018).

Method and design

The process described in the section "Conceptual background" resulted in the new COR assessment framework, which is a computer-based holistic performance assessment that measures students' and young professionals' real-world information-processing, decision-making, and judgment skills. It contains criterion-sampled realistic situations that students may encounter in their public and private lives or when studying and working in professional domains (Davey et al., 2015; Shavelson et al., 2018, 2019). Each task consists of a short context description, an objective, and a request to conduct a free Internet search (for a task example, see Figure 1). The participants are prompted to evaluate the online information they found during their search and to write a short open-ended response based on the information found. As the tasks are characterized by an open-ended information environment, with test takers having unrestricted access to the Internet for COR task processing to holistically capture the process of Internet research, those taking the test have to perform a live, open web search, find relevant and credible information, identify and exclude untrustworthy information, and write a short, coherent statement to answer the task prompt. While a processing time of 10 min per task was originally specified, the format was further adapted after the initial validation and extended to 20 min to capture the three COR facets (see section "Conceptual background") more validly.

The rating scheme for the scoring of the answers was also accordingly extended and adapted to the new CORA task format, with a greater differentiation and weighting of the individual COR facets aligned with the construct definition. The resulting rating scheme thus distinguishes between six aspects: (1) formulating a clear answer regarding the question, (2) comprehending the task, (3) quality of sources used (for researching general topics as may be encountered in public and private life), (4) accurately evaluating sources, (5) correctly considering arguments of different quality, and (6) giving a reasoned explanation. Depending on the degree of fulfillment, 0–2 points per aspect can be awarded in increments of 0.5, with the respective degree of fulfillment for the point categories described in more detail by behavioral anchors. The different aspects are then included in the overall score with different percentage weightings, depending on their importance to the overall COR construct (for an excerpt of the scoring scheme, see Table 1). While the first part of the task specifically addresses the facets of Online Information Acquisition and Critical Information Evaluation, the second part requires the ability of Reasoning based on Evidence, Argumentation, and Synthesis (see Figure 1).

In addition to the written responses, participants' browsing histories are recorded during their web search for further analysis (Nagel et al., 2020; Schmidt et al., 2020). Subsequently, the participants' responses are evaluated by trained raters using the newly developed and validated rating scheme, which takes into account the quality of the sources they used, the correctness of their evaluation of the information found, and the quality of their statements. The collected log data are analyzed, for example, in terms of the number of online sources used and the quality and type of web pages accessed. For this analysis, a new media categorization scheme was developed based on established research approaches (Nagel et al., 2020).

To examine the extent to which different test facets contribute to the variance of the test scores, we analyzed their individual contributions to the total variance of the test scores with the method of variance component analysis (Jiang, 2018). To this end, we computed linear mixed-effect models using R (lme4-package; Bates et al., 2015), in which we differentiated the assessment facets *person*, i.e., influences specific to the individual participants, *rater*, i.e., influences of rater effects or the scoring method, and *task*, i.e., influences of task characteristics, as independent variables (Shavelson and Webb, 1981). The test score was used as the dependent variable. The data set was converted for the analyses so that there was an entry in the dataset for every possible combination of characteristics (see Figure 2; Jiang, 2018). Subsequently, we calculated the linear mixed-effect models by gradually adding the *person*, *rater*, and *task* facets as well as the respective interactions, and compared them on the basis of the residuals and the variance explained in each case.

The analyses were conducted with the data of 125 students of economics and economics education at a German university, who participated in the CORA study in 2019–2020. Participants were 61% female, reported a mean age of 22 years ($SD = 2.8$), and were on average in their second semester of study ($SD = 1.82$; see Table 2). Participation in the CORA study was voluntary and

10.3389/feduc.2022.914857

TABLE 1 Excerpt of the COR Scoring Scheme, OIA Facet, Subfacet "Quality of Researched Sources."

Evaluation facets	Explanation	COR construct	Rating (highest possible score)	Gradation: Scale from 0 to 2: 0: does not apply at all 0.5: rather not applies 1: partly applies 1.5: rather applies 2: fully applies					Weighting factor
				0 points	0.5 points	1 point	1.5 points	2 points	
				Does not apply at all	Rather not applies	Partly applies	Rather applies	Fully applies	
3. Quality of sources used	**Usage of further sources** Have further sources besides the given ones been used? Is the number of used sources appropriate?	Online Information Acquisition (OIA)	Total: 2 points	Usage of qualitative sources is missing	Predominant usage of sources with insufficient quality; lack of content suitability	Partly usage of sources with insufficient quality; However, content-related suitability can be partly recognized	Usage of sources with sufficient quality and timeliness; However, content-related suitability can just be partly recognized	Usage of sources with sufficient quality and timeliness; content-related suitability absolutely recognizable	**20.00%**

requested in obligatory introductory lectures. To ensure higher test motivation for their participation in the study, the students received credits for a study module.

The study was conducted via an online assessment platform, which the participants could access individually using access data sent to them in advance. Prior to the survey, the students were informed that their web history would be recorded and that their participation in the experiment was voluntary; all participants signed a declaration of consent to the use of their data for research purposes. Subsequently, the participants were given a standardized questionnaire (approx. 10 min) collecting sociodemographic data such as gender, age, and study semester and their general (self-reported) media use behavior using the validated scale by Maurer et al. (2020). They were also asked to rate the reliability of various media types on a scale of 1 (not at all trustworthy) to 6 (very trustworthy). Due to limited test time, we used a booklet test design. Thereby, students were then given randomly assigned 2–3 CORA tasks to answer (out of a total of six available tasks), which all shared the same structure as well as task description and only differed in topic (for more details on the tasks, see section "Conceptual background"). Participants were asked to enter their written responses to the open-ended questions in the assessment platform, from which they could subsequently log out by themselves. After the assessment, the answers were scored by two trained human raters each, using the newly developed rating scheme (for more details on the scoring process, see section "Conceptual background"), and the scores of all raters for each participant and for each task were averaged to obtain the CORA score[1]. Participants' scores between tasks varied (task 1: $m=0.71$, $SD=0.64$; task 2: $m=1.3$, $SD=0.59$; task 3: $m=0.53$, $SD=0.66$; task 4: $m=0.63$, $SD=0.54$; task 5: $m=0.77$, $SD=0.61$) with an average overall score of $m=0.84$ ($SD=0.51$).

Results

Table 3 shows the results of the model calculations of the linear mixed-effect models (Jiang, 2018). First, separate models were computed for the direct effects of the considered facets *person*, *task*, and *rater* and compared to each other, showing already that in comparison, most variance is explained by the *person* facet ($R^2=0.397$), followed by the *task* facet ($R^2=0.164$). In contrast, an influence on the part of the raters was hardly observable ($R^2=0.076$). Even when combining the facets in pairs (M4–M6), the model including *person* and *task* explains most of the variance ($R^2=0.435$). Adding the facet *rater* in M7 leads only to a slight increase in the explained variance ($R^2=0.451$). If, in addition to the direct effects, the interaction effects between the facets were also taken into account, the greatest variance explanation was seen in M8, in which the interactions *person x task* and *person x rater* were included in addition to the *person* facet ($R^2=0.713$). In this model, especially the *person* x *task*

1 For the overall CORA score, a sufficient interrater reliability of Cohens kappa =0.80 ($p=0.000$) was determined.

	r1	r2	r3	r4	r5	r6	r7	r8	r9
p	1-18	1-18	1-18	9-16	9-16	9-16	17-24	17-24	17-24
1	X	x	x	NA	NA	NA	NA	NA	NA
2	NA	NA	NA	x	x	x	NA	NA	NA
3	NA	NA	NA	NA	NA	NA	x	x	x
..	..	..	..						
20	NA	NA	NA	NA	NA	NA	NA	NA	NA

FIGURE 2
Exemplary representation of the dataset format for calculating the variance component decomposition (adapted from Jiang, 2018).

TABLE 2 Sample description.

N = 125	Mean	SD
Age	22.0	2.8
Semester (Bachelor)	2.0	1.82
University Entrance Qualification[1]	2.4	0.52
CORA Score[2]	0.82	0.51
Gender		
Women	75	61.0
Men	45	36.6
Economics education students	116	94.3
First language German	110	89.4
Completed vocational training	52	42.3

[1]UEQ grades range from 1 (best grade possible) to 6 (lowest grade possible).
[2]The scores could vary between min. 0 and max. 2.

interaction stands out, which can be interpreted in the sense that there are not only general differences between the performances of the individual participants (direct effect of the *person* facet), but also that the demonstrated performances of the individual test takers differ depending on the task in question (*person* × *task*).

In summary, the comparison of the individual facets and their interactions shows that the largest effect on the CORA score is that of the individual test takers' personal characteristics or their interaction with the different tasks, with the effects of tasks and raters being present but much less pronounced.

Interpretation

The examination of the internal structure of the CORA tasks by means of variance decomposition confirms that, overall, by far the largest part of the score variances is explained by the test takers, as intended in the assessment. Here, it is also important to distinguish between the direct person effects and interaction effects of the participants with the tasks, both of which have an important influence: While the direct effect suggests that interindividual differences (in COR ability) among participants lead to different CORA performance, the interaction effects indicate that participants also perform differently intraindividually depending on the task they are working on. This can possibly be explained by the fact that certain task characteristics (e.g., formulation or the topic of the task) interact with differently developed personal characteristics of the test takers (e.g., different levels of ability in the individual COR facets, certain sociodemographic characteristics, or other personality traits) during task processing. For instance, although the tasks cover general (to the extent that this is possible) social topics, it can be assumed that the participants have a different degree of prior knowledge in certain subject areas due to individual interests, which influences them in their task performance. Which correlations between personal characteristics and CORA performance actually exist, and how these possibly interact with certain task characteristics, must be analyzed in detail in further investigations and falls within the validity criterion of "relationship with other variables" (AERA, APA, and NCME, 2014; section "Relations with other variables").

While the largest effects can be explained by the test takers, the direct effects of the raters and the tasks turn out to be much smaller, which suggests that, in terms of assessment, there are rather small systematic influences caused by the task properties (e.g., different difficulty) or the rater effects. Nevertheless, it is also necessary to analyze these in further studies, for example, with regard to the task difficulty of individual topics, to ensure the comparability of the respective results. In addition, to be able to draw comparable conclusions about the performance, the tasks should not be used alone, but, as intended in the assessment, rather in combination if possible.

Overall, the analyses confirm that the CORA scores indeed reflect differences in the performance of the participants and are only marginally influenced by rater effects and task properties, which also speaks in favor of maintaining the methodological approach used (rating scheme, rater training, and standardized structure of the tasks).

TABLE 3 Results of variance decomposition according to G-Theory using mixed-effect models.

Model	M1	M2	M3	M4	M5	M6	M7	M8
Persons	0.1892			0.1720	0.1751		0.1601	0.1293
Tasks		0.0996		0.0867		0.0964	0.0862	
Raters			0.0464		0.0252	0.0408	0.0235	
Persons x Tasks								0.2746
Persons x Raters								0.0143
Raters x Tasks								
Residuals	0.3975	0.5068	0.5617	0.3364	0.3902	0.4862	0.3288	0.1685
Explained R^2	**0.3225**	**0.1642**	**0.0763**	**0.4347**	**0.3392**	**0.2202**	**0.4506**	**0.7128**

Nevertheless, it is important for the further development and interpretation of CORA to investigate the causes for the found rater effects more closely and, if necessary, to make adjustments with regard to the rating scheme, the training, and the selection of the raters. Even if the content validity of the tasks and the developed rating scheme has already been demonstrated by the findings of Molerov et al. (2020), it should also be ensured in further analyses and, if necessary, expert interviews that these actually cover only the COR skills and do not, for example, systematically disadvantage individual groups of people due to the task topics (e.g., men/women might have different preferences on health-or sport-related topics).

Relations with other variables

Theoretical background

The previous explanations have shown that (1) the content of the assessment covers the targeted COR skills as expected (section "Content validity"), (2) the difference in test scores results from the participants' performance in the tasks and not from other aspects of the assessment (section "Internal test structure"), and (3) the tasks trigger different task-solving processes in the participants as expected (section "Validity of task response processes"). Subsequently, it is necessary for the further interpretation and use of the test scores to consider them in the context of further variables with which, according to the underlying COR construct, there should theoretically be (no) correlations. The testing of these relationships is referred to as convergent and discriminant validity, respectively (see also Campbell and Fiske, 1959). According to AERA, APA, and NCME (2014), this type of validity evidence belongs to the category "Evidence based on Relations to other Variables" and provides information on the extent to which the relationships of the test scores with other variables are consistent with the underlying construct and the proposed test score interpretation.

Previous studies, in which the construct related to COR, Civic Online Reasoning, was examined for middle school, high school, and college students in the United States, showed a positive correlation between COR-related skills and study progress (McGrew et al., 2018), in that college students performed better than high school students and high school students performed better than middle school students. Recent studies also concluded that these skills improved with increasing expertise and higher grade level (e.g., Nygren and Guath, 2020; Breakstone et al., 2021; Guath and Nygren, 2022). Also related to the COR construct, which, according to the definition, can be enhanced by corresponding training, students' COR ability should improve due to increasing experience with online research and the writing of scientifically argumentative texts over the course of studies (Molerov et al., 2020). Thus, in terms of convergent validity, students who are further along in their studies should perform better in CORA than students at the beginning of their studies.

No differences in COR-related abilities were found in previous studies with respect to gender (Breakstone et al., 2021). Moreover, according to the construct definition of COR, gender effects are not expected to occur in research on general social topics. Thus, in terms of discriminant validity, there should be no correlations between the participants' gender and their CORA scores.

A central aspect of the COR construct is the critical selection, weighting, and use of suitable reliable sources for task-based research (Molerov et al., 2020). In this regard, studies showed that the selection and use of online sources depends to a large extent on their trustworthiness as perceived by users, so that sources perceived as trustworthy are preferred when searching for information (Wathen and Burkell, 2001; Harrison McKnight and Kacmar, 2007; Rowley et al., 2015). Accordingly, a correct assessment of the trustworthiness of (online) sources should also lead to an appropriate differentiation and use of trustworthy versus untrustworthy sources, and thus to better performance in terms of the COR construct (Molerov et al., 2020). Social media in particular, which include video platforms and online encyclopedias, are to be regarded critically in terms of their trustworthiness as they are considered less reliable in terms of their information content (Ciampaglia, 2018; Maurer et al., 2018). Consequently, using such sources may correlate with poorer CORA performance. The use of the Google search engine as an information platform should also be evaluated critically. Search engines such as Google are often the starting point for an Internet-based search and also constitute an important tool for professional fact-checkers when researching information (Speicher et al., 2015; McGrew et al., 2017). However, they display results from media with varying degrees of reliability (which is the reason they lend themselves to the abovementioned practices), and the first search

results in particular are often sponsored (Wineburg et al., 2016). Thus, a reasonably low level of confidence in these websites as sources of information should lead to reduced use of these websites and thus higher research quality and a better performance in CORA.

Method and design

The examination of the assumed correlations took place within the same study framework and sample described in the section "Method and design." Relationship analyses of the CORA score with participants' age and gender and their media reliability ratings were conducted using correlation analyses (age and media reliability) and a two-sided *t*-test (gender) in Stata 17 (StataCorp, 2021).

Results

As expected, students in higher semesters achieved better CORA scores than those in lower semesters ($r = 0.25$, $p = 0.006$); gender did not play a significant role [$t(116) = -2.00$, $p = 0.05$]. Regarding trust in different types of online media, the analyses revealed significant associations between CORA score and reported trust in video platforms ($r = -0.239$, $p = 0.009$), online encyclopedias ($r = -0.187$, $p = 0.04$), and Google as an information platform ($r = -0.19$, $p = 0.038$) for Internet research. These relationships are also reflected in the actual use of online media, where less frequent use of online encyclopedias ($r = -0.245$, $p = 0.037$) and Google as an information platform ($r = -0.234$, $p = 0.047$) are associated with a better CORA score. Lower trust in video platforms, online encyclopedias, and Google as an information platform as well as less usage of video platforms and online encyclopedias is thus associated with better CORA scores, and greater trust or more use with poorer CORA performance, respectively.

Interpretation

Viewing these results in the context of the external variables we controlled for in our study provides initial evidence that expected correlations exist with respect to both convergent (semester of study, media use) and discriminant (gender) validity. In line with the construct definition, no correlations of CORA performance with gender were found. In contrast, participants who were more advanced in their studies (and thus had already had more learning opportunities in terms of researching information on the Internet as well as writing argumentative texts) showed better CORA performance than students at the beginning of their studies. In addition to institutional learning opportunities, the correlation between general trust in specific types of media for obtaining information and CORA performance, which was expected according to the construct definition, also became evident. These analyses thus confirm that the theoretically formulated basic assumptions regarding the construct, namely the development of COR ability over the course of academic studies and the general importance of media types used, were reflected in the empirically observable correlations with the CORA scores.

Based on the analyses presented here, however, the basic assumptions of the construct cannot yet be considered comprehensively confirmed, since this would first require examining further correlations with other theoretically relevant external variables. With regard to learning opportunities, for example, it would have to be investigated more concretely to what extent COR-developing aspects are actually anchored in the curriculum of the study participants. Furthermore, to gain a better understanding of the (possible) development of COR skills within higher education and beyond, and to what extent these skills can be effectively fostered over the course of academic studies, the actual development of COR skills should also be investigated, for instance in the context of a specific targeted training with comparison groups and pre-post testing, provided that the COR tasks have been proven to be sufficiently valid (Zlatkin-Troitschanskaia et al., 2021a).

Similarly, in terms of media use, the analyses described above cover trust in some media types. While these are highly relevant, especially since, e.g., online encyclopedias are an important source of information for students (Selwyn and Gorard, 2016), the analyses are not exhaustive as they do not (yet) consider other types of higher quality information sources, such as online academic catalogues, professional magazines, or established news sites. Further investigation of the relationship between the CORA score and additional external variables, as well as analyses including participant cohorts other than students, are thus required to provide more comprehensive validation.

Discussion, limitations, and outlook on further research

Today, the Internet has become one of the most important sources of information for learning for university students and young professionals (Brooks, 2016; Newman and Beetham, 2017). However, relying on online resources for information acquisition also presents challenges, as content can be freely distributed on the Internet and vast amounts of unstructured, untrustworthy, inaccurate, or biased information are just as readily available to learners as credible, verified information (Qiu et al., 2017; Ciampaglia, 2018; Maurer et al., 2018). To competently use the information on the Internet, students must be able to critically search, select, review, and evaluate online information and sources based on relevant quality criteria (Molerov et al., 2020; Nagel et al., 2020). To make these skills empirically measurable and to be able to specifically promote them, we developed the new theoretical-conceptual framework of Critical Online Reasoning (Molerov et al., 2020) and a corresponding COR Assessment (CORA) in accordance with the evidence-centered design approach of Mislevy (2017) and the Standards for Educational and Psychological Testing of AERA, APA, and NCME (2014).

To ensure that the newly developed assessment actually measures COR abilities as defined by the construct, we followed the argumentative validation process described by AERA, APA, and NCME (2014), according to which the five aspects "test content," "task-and test-response processes," "internal structure of

a test," "interrelationships with other variables," and "consequences of testing" should be analyzed during validation and various sources of information should be used as evidence. In the course of this, evidence for the validity of the CORA scores and their interpretation could already be shown regarding the CORA content (through expert interviews and expert rating of the CORA tasks; Molerov et al., 2020), the task response processes of the test takers (on the basis of log files and eye-tracking data; Schmidt et al., 2020) and, initially, correlations with other variables (regarding number and type of web pages accessed as well as the quality of the content on the web pages; Nagel et al., 2020).

Building on this previous research, the analyses presented here focus on the two criteria "internal structure of the test," carried out via variance component analysis based on Generalizability Theory, and "correlations with other variables." The analyses were conducted with the data of 125 students of economics and economics education. The results of the analysis regarding the internal structure of the CORA confirmed that the largest effect on the CORA score is that of the individual test takers' personal characteristics or their interaction with the different tasks as intended in the assessment, with the effects of tasks and raters being present but much less pronounced. Also, the analyses covering the validity facet "relationship with other variables" confirm that the theoretically formulated basic assumptions regarding the construct exist with respect to both convergent (semester of study, media use) and discriminant (gender) validity. The results of the separate validity analyses are also consistent when viewed as a whole in the sense of a holistic validation argumentation: The presented correlations once again support the expert opinions in Molerov et al. (2020) that the tasks validly measure the COR construct. In addition, they complement the findings of Schmidt et al. (2020) and Nagel et al. (2020) by showing that "good" COR is not only characterized by an appropriate research strategy (i.e., strategic information processing with the use of a larger number of different sources and a larger number of process steps), but that the quality and appropriate evaluation of the sources used also play an important role. At the same time, it can be assumed that individual differences in web search behavior and media use are some of the factors that exert an influence on the CORA score in the context of the direct and interaction effects of the person facet.

For the purpose of further test validation and also a deeper understanding of the COR construct, it is necessary to examine more closely which of the interindividual differences could be identified via the direct effect and thus have an influence independent of the task, and which differences are sensitive to (which) CORA task characteristics as showed in the interaction effects. This concerns both the personal characteristics that were already examined and other characteristics that should be considered additionally, as was also recommended by experts interviewed by Molerov et al. (2020), such as personality traits, prior knowledge, and their relations to the different manifestations of the individual COR facets. While it can be assumed, for example, that personality traits should have a task-overarching effect, the effects of prior knowledge, interests, beliefs, and (political) attitudes may vary depending on the task, and the different characteristics of the individual COR facets could become noticeable via both types of effect. One way to investigate the role of individual influencing factors such as prior knowledge or beliefs during task processing would be by means of cognitive labs with think-aloud commentary (Leighton, 2017), the results of which would also complement the initial eye-tracking studies of Schmidt et al. (2020). These findings would be essential both for the development of specific training tools as well as for ensuring test fairness and in the sense of the validity facet "consequences of testing," which is the only one that has not yet been investigated in detail. In addition, the effects of the raters and tasks found should be examined more closely for their causes, even though they turned out to be rather small, to minimize possible systematic influences by, for example, rater effects, rating scheme or the format or topics of the tasks. This could be done, for instance, by a systematic comparative analysis of the individual ratings and tasks, or, if necessary, expert interviews to also make sure that the assessment does not systematically disadvantage individual groups of people, e.g., due to rater effects or the task topics. These results can also serve as first analyses in terms of validating the CORA regarding its consequences of testing.

In general, analyses including participant cohorts with students from other study subjects and participants other than students are needed to validate the use of the assessment for a broader population. Since the present analyses were conducted with a comparatively small sample, this would also be helpful in confirming the obtained results and expanding their scope. Although the sample size is sufficient for the analyses carried out, the correlations found could become even more significant with a larger sample, particularly with regard to the analyses carried out. In addition, it is necessary for the validation process to take into account the dynamics prevailing on the Internet, which make it difficult to compare results between participants due to the constantly changing information and media landscape and can also lead to a fast outdating of individual CORA tasks. As a result, it may become necessary to continuously develop new task topics, which also have to be examined for their validity.

The first steps toward implementing the above measures have already been taken in the BRIDGE project (Zlatkin-Troitschanskaia et al., 2021a). There, by using CORA for students in different study phases, the scope of the assessment was once again checked with regard to the suitability of the task difficulty and above that extended to young professionals. In addition, a comprehensive (sociodemographic) accompanying questionnaire was developed, which covers a variety of personal characteristics, e.g., previous knowledge and personal attitudes on a topic, and thus allows more detailed analyses of influencing factors.

In summary, comprehensive validity evidence is available for the CORA for four of the five criteria for valid tests and test score interpretations recommended by AERA, APA, and NCME (2014), with the "consequences of testing" criterion requiring further investigation. Although further analyses are reasonable regarding all validity criteria and necessary in the sense of the argumentative

validity approach according to AERA, APA, and NCME (2014), it can nevertheless be concluded that with the CORA, for the first time, a performance assessment is available for Germany, which can be used in a valid manner to assess the interplay between features of online information acquisition and learning environments and the (cognitive) requirements for critical reasoning from online information in students and young professionals.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

M-TN co-developed and carried out the assessments, conducted the analyses, and co-wrote the article. OZ-T co-developed the assessment, supervised the development and validation process as well as the analyses, and co-wrote the article. JF contributed to the article with a co-workup of the literature relevant to the topic and with co-developing new assessment tasks and a scoring scheme. All authors contributed to the article and approved the submitted version.

Funding

This study is part of the PLATO project, which is funded by the German federal state of Rhineland-Palatinate.

Acknowledgments

We would like to thank the two reviewers and the editor who provided constructive feedback and helpful guidance in the revision of this manuscript. We would also like to thank all students from the Faculty of Law and Economics at Johannes Gutenberg University Mainz who participated in this study as well as the raters who evaluated the written responses. We would also like to thank Katharina Frank, who contributed to the workup of the literature relevant to the topic.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

AERA, APA, and NCME (2014). *Standards for Educational and Psychological Testing*. Washington DC: American Educational Research Association.

Ali, W. (2020). Online and remote learning in higher education institutes: a necessity in light of COVID-19 pandemic. *High. Educ. Stud.* 10, 16–25. doi: 10.5539/hes.v10n3p16

Aspegren, K., and LØnberg-Madsen, P. (2005). Which basic communication skills in medicine are learnt spontaneously and which need to be taught and trained? *Med. Teach.* 27, 539–543. doi: 10.1080/01421590500136501

Banerjee, M., Zlatkin-Troitschanskaia, O., and Roeper, J. (2020). Narratives and their impact on students' information seeking and critical online reasoning in higher education economics and medicine. *Front. Educ.* 5:570625. doi: 10.3389/feduc.2020.570625

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., et al. (2012). "Defining twenty-first century skills," in *Assessment and Teaching of 21st Century Skills*. eds. P. Griffin, B. McGraw and E. Care (New York: Springer), 17–66.

Boh Podgornik, B., Dolničar, D., Šorgo, A., and Bartol, T. (2016). Development, testing, and validation of an information literacy test (ILT) for higher education. *J. Assoc. Inf. Sci. Technol.* 67, 2420–2436. doi: 10.1002/asi.23586

Braasch, J. L. G., Bråten, I., and McCrudden, M. T. (2018). *Handbook of Multiple Source Use*. New York: Routledge.

Brand-Gruwel, S., Wopereis, I., and Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Comput. Educ.* 53, 1207–1217. doi: 10.1016/j.compedu.2009.06.004

Braun, E. (2021). Performance-based assessment of students' communication skills. *Int. J. Chin. Educ.* 10:221258682110062. doi: 10.1177/22125868211006202

Braun, E., and Brachem, J. (2018). Erfassung praxisbezogener Anforderungen und Tätigkeiten von Hochschulabsolvent(inn)en (PAnTHoa). *Zeitschrift für Hochschulentwicklung* 209–232.

Braun, H. I., Shavelson, R. J., Zlatkin-Troitschanskaia, O., and Borowiec, K. (2020). Performance assessment of critical thinking: conceptualization, design, and implementation. *Front. Educ.* 5:156. doi: 10.3389/feduc.2020.00156

Breakstone, J., Smith, M., Wineburg, S., Rapaport, A., Carle, J., Garland, M., et al. (2021). Students' civic online reasoning: a National Portrait. *Educ. Res.* 50, 505–515. doi: 10.3102/0013189X211017495

Brooks, D. C. (2016). ECAR study of undergraduate students and information technology, 2016. EDUCAUSE Center for Analysis and Research. Available at: https://library.educause.edu/~/media/files/library/2016/10/ers1605.pdf

Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* 56, 81–105. doi: 10.1037/h0046016

Chan, C. K. Y., Fong, E. T. Y., Luk, L. Y. Y., and Ho, R. (2017). A review of literature on challenges in the development and implementation of generic competencies in higher education curriculum. *Int. J. Educ. Dev.* 57, 1–10. doi: 10.1016/j.ijedudev.2017.08.010

Ciampaglia, L. G. (2018). "The digital misinformation pipeline–proposal for a research agenda," in *Positive Learning in the Age of Information. A Blessing or a Curse?* eds. O. Zlatkin-Troitschanskaia, G. Wittum and A. Dengel (Wiesbaden: Springer VS), 413–421.

Cronbach, L. J., Gleser, G. C., Nanda, H., and Rajaratnam, N. (1972). *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles*. New York: John Wiley.

Cronbach, L. J., Nageswari, R., and Gleser, G. C. (1963). Theory of generalizability: a liberation of reliability theory. *Br. J. Statis. Psychol.* 16, 137–163. doi: 10.1111/j.2044-8317.1963.tb00206.x

Davey, T., Ferrara, S., Holland, P. W., Shavelson, R., Webb, N. M., and Wise, L. L. (2015). Psychometric considerations for the next generation of performance assessment: report of the center for K-12 assessment and performance management at ETS [white paper]. Educational Testing Service. Available at: https://www.ets.org/Media/Research/pdf/psychometric_considerations_white_paper.pdf

Desai, S., and Reimers, S. (2019). Comparing the use of open and closed questions for web-based measures of the continued-influence effect. *Behav. Res. Methods* 51, 1426–1440. doi: 10.3758/s13428-018-1066-z

Flanagin, A., and Metzger, M. J. (2017). *Digital Media and Perceptions of Source Credibility in Political Communication. The Oxford Handbook of Political Communication*, vol. 417. doi: 10.1093/oxfordhb/9780199793471.013.65

Goldhammer, F., Naumann, J., and Keßel, Y. (2013). Assessing individual differences in basic computer skills. *Eur. J. Psychol. Assess.* 29, 263–275. doi: 10.1027/1015-5759/a000153

Goldman, S. R., and Brand-Gruwel, S. (2018). "Learning from multiple sources in a digital society," in *International Handbook of the Learning Sciences.* eds. F. Fischer, C. E. Hmelo-Silver, S. R. Goldman and P. Reimann (New York: Routledge), 86–95.

Guath, M., and Nygren, T. (2022). Civic online reasoning among adults: an empirical evaluation of a prescriptive theory and its correlates. *Front. Educ.* 7:721731. doi: 10.3389/feduc.2022.721731

Hahnel, C., Schoor, C., Kroehne, U., Goldhammer, F., Mahlow, N., and Artelt, C. (2019). The role of cognitive load in university students' comprehension of multiple documents. *Zeitschrift für pädagogische Psychologie* 33, 105–118. doi: 10.1024/1010-0652/a000238

Harrison, N., and Luckett, K. (2019). Experts, knowledge and criticality in the age of 'alternative facts': reexamining the contribution of higher education. *Teach. High. Educ.* 24, 259–271. doi: 10.1080/13562517.2019.1578577

Harrison McKnight, D., and Kacmar, C. J. (2007). "Factors and effects of information credibility," in *ICEC'07: Proceedings of the ninth international conference on Electronic commerce*. eds. D. Sarppo, M. Gini, R. J. Kauffman, C. Dellarocas and F. Dignum; Association for Computing Machinery, 423–432.

Herrero-Diz, P., Conde-Jiménez, J., Tapia-Frade, A., and Varona-Aramburu, D. (2019). The credibility of online news: an evaluation of the information by university students. *Cult. Educ.* 31, 407–435. doi: 10.1080/11356405.2019.1601937

Huang, K., Law, V., Ge, X., Hu, L., and Chen, Y. (2019). Exploring patterns in undergraduate students' information problem solving: a cross-case comparison study. *Knowledge Manag. E-Learn.* 11, 428–448. doi: 10.34105/j.kmel.2019.11.023

Jiang, Z. (2018). Using the linear mixed-effect model framework to estimate generalizability variance components in R. *Methodology* 14, 133–142. doi: 10.1027/1614-2241/a000149

Johnson, F., Sbaffi, L., and Rowley, J. (2016). Students' approaches to the evaluation of digital information: insights from their trust judgments. *Br. J. Educ. Technol.* 47, 1243–1258. doi: 10.1111/bjet.12306

Koltay, T. (2011). The media and the literacies: media literacy, information literacy, digital literacy. *Media Cult. Soc.* 33, 211–221. doi: 10.1177/0163443710393382

Korn, J. (2004). Teaching talking: Oral communication skills in a Law course. *J. Leg. Educ.* 54, 588–596.

Ku, K. Y. (2009). Assessing students' critical thinking performance: urging for measurements using multi-response format. *Think. Skills Creat.* 4, 70–76. doi: 10.1016/j.tsc.2009.02.001

Leeder, C. (2019). How college students evaluate and share "fake news" stories. *Libr. Inf. Sci. Res.* 41:100967. doi: 10.1016/j.lisr.2019.100967

Leighton, J. P. (2017). *Using Think-Aloud Interviews and Cognitive Labs in Educational Research*. New York: Oxford University Press

Limberg, L., Sundin, O., and Talja, S. (2012). Three theoretical perspectives on information literacy. *Hum. IT* 11, 93–130.

Liu, O. L., Frankel, L., and Crotts Roohs, K. (2014). Assessing critical thinking in higher education: current state and directions for next-generation assessment. *ETS Res. Rep. Ser.* 2014, 1–23. doi: 10.1002/ets2.12009

Makhmudov, K., Shorakhmetov, S., and Murodkosimov, A. (2020). Computer literacy is a tool to the system of innovative cluster of pedagogical education. *Eur. J. Res. Reflect. Educ. Sci.* 8, 71–74. doi: 10.6084/m9.figshare.12310661

Maurer, M., Quiring, O., and Schemer, C. (2018). "Media effects on positive and negative learning," in *Positive Learning in the Age of Information. A Blessing or a Curse?* eds. O. Zlatkin-Troitschanskaia, G. Wittum and A. Dengel (Wiesbaden: Springer VS), 197–208.

Maurer, M., Schemer, C., Zlatkin-Troitschanskaia, O., and Jitomirski, J. (2020). "Positive and negative media effects on university students' learning: Preliminary findings and a research program," in *Frontiers and Advances in Positive Learning in the Age of Information (PLATO).* ed. O. Zlatkin-Troitschanskaia (Wiesbaden: Springer), 109–119.

Mayer, R. E. (2009). *Multimedia Learning. 2nd* Edn Cambridge University Press.

McGrew, S., Breakstone, J., Ortega, T., Smith, M., and Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theor. Res. Soc. Educ.* 46, 165–193. doi: 10.1080/00933104.2017.1416320

McGrew, S., Ortega, T., Breakstone, J., and Wineburg, S. (2017). The challenge that's bigger than fake news. Civic reasoning in a social media environment. *Am. Educ.* 41, 4–9.

Mislevy, R. J. (2017). *Sociocognitive Foundations of Educational Measurement*. New York: Routledge.

Mislevy, R. J., Behrens, J. T., Dicerbo, K. E., and Levy, R. (2012). Design and discovery in educational assessment: Evidence-centered design, psychometrics, and educational data mining. *Journal of Educational Data Mining* 4, 11–48.

Molerov, D., Zlatkin-Troitschanskaia, O., Nagel, M.-T., Brückner, S., Schmidt, S., and Shavelson, R. J. (2020). Assessing university students' critical online reasoning ability: a conceptual and assessment framework with preliminary evidence. *Front. Educ.* 5:577843. doi: 10.3389/feduc.2020.577843

Molerov, D., Zlatkin-Troitschanskaia, O., and Schmidt, S. (2019). "Adapting the civic online reasoning assessment cross-nationally using an explicit functional equivalence approach [paper presentation]." in *Annual meeting of the American educational research association*, Toronto, Canada.

Nagel, M.-T., Schäfer, S., Zlatkin-Troitschanskaia, O., Schemer, C., Maurer, M., Molerov, D., et al. (2020). How do University students' web search behavior, website characteristics, and the interaction of Both influence students' critical online reasoning? *Front. Educ.* 5:565062. doi: 10.3389/feduc.2020.565062

National Research Council (2012). *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. Washington, DC: National Academies Press

Naumann, J., Richter, T., and Groeben, N. (2001). Validierung des INCOBI anhand eines Vergleichs von Anwendungsexperten und Anwendungsnovizen. *Zeitschrift für Pädagogische Psychologie* 15, 219–232. doi: 10.1024//1010-0652.15.34.219

Newman, T., and Beetham, H. (2017). Student digital experience tracker 2017: the voice of 22,000 UK learners. Jisc. Available at: https://repository.jisc.ac.uk/6662/1/Jiscdigitalstudenttracker2017.pdf

Nygren, T., and Guath, M. (2020). Students evaluating and corroborating digital news. *Scand. J. Educ. Res.* 66, 549–565. doi: 10.1080/00313831.2021.1897876

Osborne, J., Pimentel, D., Alberts, B., Allchin, D., Barzilai, S., Bergstrom, C., et al. (2022). *Science Education in an Age of Misinformation*. Report. Stanford. Available at: https://sciedandmisinfo.stanford.edu/

Oser, F. K., and Biedermann, H. (2020). "A three-level model for critical thinking: Critical alertness, critical reflection, and critical analysis," in *Frontiers and Advances in Positive Learning in the Age of InformaTiOn (PLATO)*. ed. O. Zlatkin-Troitschanskaia. (Wiesbaden: Springer), 89–106.

Park, H., Kim, H. S., and Park, H. W. (2021). A Scientometric study of digital literacy, ICT literacy, information literacy, and media literacy. *J. Data Info. Sci.* 6, 116–138. doi: 10.2478/jdis-2021-0001

Qiu, X., Oliveira, D. F. M., Shirazi, A. S., Flammini, A., and Menczer, F. (2017). Limited individual attention and online virality of low-quality information. *Nat. Hum. Behav.* 1, 1–22. doi: 10.1038/s41562-017-0132

Rammstedt, B. (Ed.) (2013). *Grundlegende Kompetenzen Erwachsener im internationalen Vergleich: Ergebnisse von PIAAC 2012*. Münster: Waxmann.

Reddy, P., Sharma, B., and Chaudhary, K. (2020). Digital literacy: a review of literature. *Int. J. Technoethics* 11, 65–94. doi: 10.4018/IJT.20200701.oa1

Rowley, J., Johnson, F., and Sbaffi, L. (2015). Students' trust judgements in online health information seeking. *Health Informatics J.* 21, 316–327. doi: 10.1177/1460458214546772

Sanders, L., Kurbanoğlu, S., Boustany, J., Dogan, G., and Becker, P. (2015). Information behaviors and information literacy skills of LIS students: an international perspective. *J. Educ. Libr. Inf. Sci.* 56, 80–99. doi: 10.12783/issn.2328-2967/56/S1/9

Schlebusch, C. L. (2018). Computer anxiety, computer self-efficacy and attitudes toward the internet of first year students at a south African University of Technology. *Africa Educ. Rev.* 15, 72–90. doi: 10.1080/18146627.2017.1341291

Schmidt, S., Zlatkin-Troitschanskaia, O., Roeper, J., Klose, V., Weber, M., Bültmann, A.-K., et al. (2020). Undergraduate students' critical online reasoning: process mining analysis. *Front. Psychol.* 11:576273. doi: 10.3389/fpsyg.2020.576273

Selwyn, N., and Gorard, S. (2016). Students' use of Wikipedia as an academic resource—patterns of use and perceptions of usefulness. *Internet High. Educ.* 28, 28–34. doi: 10.1016/j.iheduc.2015.08.004

Sendurur, E. (2018). Students as information consumers: a focus on online decision making process. *Educ. Inf. Technol.* 23, 3007–3027. doi: 10.1007/s10639-018-9756-9

Shavelson, R. J., and Webb, N. M. (1981). Generalizability theory: 1973–1980. *Br. J. Math. Stat. Psychol.* 34, 133–166. doi: 10.1111/j.2044-8317.1981.tb00625.x

Shavelson, R. J., Webb, N. M., and Rowley, G. L. (1989). Generalizability theory. *American Psychologist* 44, 922–932. doi: 10.1037/0003-066X.44.6.922

Shavelson, R. J., Zlatkin-Troitschanskaia, O., Beck, K., Schmidt, S., and Mariño, J. P. (2019). Assessment of university students' critical thinking: next generation performance assessment. *Int. J. Test.* 19, 337–362. doi: 10.1080/15305058.2018.1543309

Shavelson, R. J., Zlatkin-Troitschanskaia, O., and Marino, J. P. (2018). International Performance Assessment of Learning in Higher Education (iPAL): Research and Development. *Assessment of Learning Outcomes in Higher Education: Cross-National Comparisons and Perspectives*, 193–214. doi: 10.1007/978-3-319-74338-7_10

Siddiq, F., Hatlevik, O. E., Olsen, R. V., Throndsen, I., and Scherer, R. (2016). Taking a future perspective by learning from the past–a systematic review of assessment instruments that aim to measure primary and secondary school students' ICT literacy. *Educ. Res. Rev.* 19, 58–84. doi: 10.1016/j.edurev.2016.05.002

Solano-Flores, G. (2020). Boolean analysis of Interobserver agreement: formal and functional evidence sampling in complex coding endeavors. *Educ. Meas. Issues Pract.* 40, 26–36. doi: 10.1111/emip.12409

Speicher, M., Both, A., and Gaedke, M. (2015). "SOS: does your search engine results page (SERP) need help?" in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. (eds.) B. Begole, J. Kim, K. Inkpen and W. Woo; Association for Computer Machinery, 1005–1014.

StataCorp. (2021). *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC.

Tribukait, M., Baier, K., Grzempa, H., Loukovitou, A., Sijakovic, R., Tettschlag, N., et al. (2017). Digital learning in European education policies and history curricula. Eckert. Dossiers, 13. Available at: https://repository.gei.de/handle/11428/232

Virtanen, A., and Tynjälä, P. (2018). Factors explaining the learning of generic skills: a study of university students' experiences. *Teaching in Higher Education* 24, 880–894. doi: 10.1080/13562517.2018.1515195

Walton, G., Barker, J., Pointon, M., Turner, M., and Wilkinson, A. (2020). "Information literacy and the societal imperative of information discernment," in *Informed Societies. Why Information Literacy Matters for Citizenship, Participation and Democracy*. ed. S. Goldstein (London: Facet Publishing), 149–164.

Wathen, C. N., and Burkell, J. (2001). Believe it or not: factors influencing credibility on the web. *J. Am. Soc. Inf. Sci. Technol.* 53, 134–144. doi: 10.1002/asi.10016

Weber, H., Becker, D., and Hillmert, S. (2019). Information-seeking behaviour and academic success in higher education: which search strategies matter for grade differences among university students and how does this relevance differ by field of study? *High. Educ.* 77, 657–678. doi: 10.1007/s10734-018-0296-4

Whitelock-Wainwright, A., Laan, N., Wen, D., and Gašević, D. (2020). Exploring student information problem solving behaviour using fine-grained concept map and search tool data. *Comput. Educ.* 145:103731. doi: 10.1016/j.compedu.2019.103731

Wineburg, S., Breakstone, J., McGrew, S., and Ortega, T. (2018). "Why Google can't save us. The challenges of our post-Gutenberg moment," in *Positive Learning in the Age of Information. A Blessing or a Curse?* eds. O. Zlatkin-Troitschanskaia, G. Wittum and A. Dengel (Wiesbaden: Springer VS), 221–228.

Wineburg, S., and McGrew, S. (2016). Why students can't google their way to the truth: fact-checkers and students approach websites differently. *Educ. Week* 36, 22–28.

Wineburg, S., and McGrew, S. (2019). Lateral reading and the nature of expertise: reading less and learning more when evaluating digital information. *Teach. Coll. Rec.* 121, 1–40. doi: 10.1177/016146811912101102

Wineburg, S., McGrew, S., Breakstone, J., and Ortega, T. (2016). Evaluating information: The cornerstone of civic online reasoning. Stanford Digital Repository.

Zieky, M. J. (2014). An introduction to the use of evidence-centered design in test development. *Psicol. Educ.* 20, 79–87. doi: 10.1016/j.pse.2014.11.003

Zlatkin-Troitschanskaia, O., Brückner, S., Nagel, M.-T., Bültmann, A.-K., Fischer, J., Schmidt, S., et al. (2021a). Performance assessment and digital training framework for young professionals' generic and domain-specific online reasoning in law, medicine, and teacher practice. *J. Supranat. Pol. Educ.* 13, 9–36. doi: 10.15366/jospoe2021.13.001

Zlatkin-Troitschanskaia, O., Hartig, J., Goldhammer, F., and Krstev, J. (2021b). Students' online information use and learning progress in higher education – a critical literature review. *Stud. High. Educ.* 46, 1996–2021. doi: 10.1080/03075079.2021.1953336

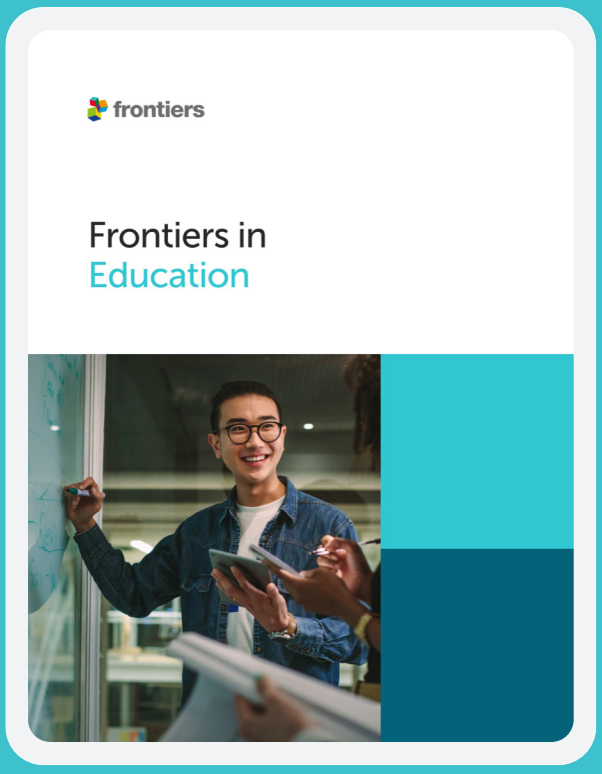
Frontiers in Education

Explores education and its importance for individuals and society

A multidisciplinary journal that explores research-based approaches to education for human development. It focuses on the global challenges and opportunities education faces, ultimately aiming to improve educational outcomes.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact