Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-023-02542-9 OPEN

Exploring the connection between deep learning and learning assessments: a cross-disciplinary engineering education perspective

Sabrina Fawzia[1] & Azharul Karim[2✉]

It is widely accepted that student learning is significantly affected by assessment methods, but a concrete relationship has not been established in the context of multidisciplinary engineering education. Students make a physiological investment and internalize learning (deep learning) if they see high value in their learning. They persist despite challenges and take delight in accomplishing their work. As student deep learning is affected by the assessment system, it is important to explore the relationship between assessment systems and factors affecting deep learning. This study identifies the factors associated with deep learning and examines the relationships between different assessment systems those factors. A conceptual model is proposed, and a structured questionnaire was designed and directed to 600 Queensland University of Technology (QUT) multidisciplinary engineering students, with 243 responses received. The gathered data were analyzed using both SPSS and SEM. Exploratory factor analysis revealed that deep learning is strongly associated with learning environment and course design and content. Strong influence of both summative and formative assessment on learning was established in this study. Engineering educators can facilitate deep learning by adopting both assessment types simultaneously to make the learning process more effective. The proposed theoretical model related to the deep learning concept can support the key practices and modern learning methodologies currently adopted to enhance the learning and teaching process.

[1] School of Civil and Environmental Engineering, Queensland University of Technology, Brisbane, QLD, Australia. [2] School of Mechanical, Medical and Process Engineering, Queensland University of Technology, Brisbane, QLD, Australia. ✉email: azharul.karim@qut.edu.au

Introduction

Marton and Säljö (1976) coined the term "deep learning" first to distinguish between various ways in which students respond to learning tasks. The deep approach to learning is preferred as it enables students to comprehend the underlying meaning rather than just surface-level details (Marton and Säljö 1976). Lynch et al. (2012) discovered that self- and peer-assessment and feedback practices can promote deep learning. Geitz et al. (2015) studied the relationship between goal orientation, deep learning, and effective feedback in higher education settings. The deep learning approach is associated with intrinsic motivation and interest in the task, an emphasis on understanding the meaning of the material, an attempt to connect different parts of the task, and a focus on linking new concepts with prior knowledge and everyday experiences. In contrast, the surface approach is characterized by extrinsic or instrumental motivation, memorization of discrete facts, and rote learning of procedures, with no connections made between different tasks or to real-life situations. It has been observed that students using the deep approach perform better and retain, integrate, and transfer knowledge at higher rates than those using the surface approach (Ramsden 2003). This learning creates a favorable learning environment (Laird et al. 2008) that leads to better learning outcomes and higher student engagement with the learning process (Laird et al. 2008).

Khilji (2022) examined the continuously growing use of virtual learning environment (VLE) for students' effective engagement as digital technology has transformed teaching and learning and as a result VLE became indispensable part of effectively engaging students through blended learning. His study investigated students' cognitive, emotional, and behavioral engagement through knowledge management and blended learning practices. In the current era of smart classrooms and online learning, creating an appropriate learning environment is crucial for students, and designing effective contents can be instrumental in fostering positive emotional states that facilitate learning (Gupta et al. 2019). Various techniques have been developed to leverage students' effective content analysis for improving their learning outcomes.

According to the National Science Board, employers seek engineers with a passion for their work, a systems thinking approach, the ability to innovate and work in diverse environments, interdisciplinary skills, strong communication and leadership skills, adaptability, and a commitment to lifelong learning (Beering 2007). A deep learning approach can encourage students to engage in comparative and synthetic thinking, making it essential to promote this approach in engineering education. Therefore, undergraduate engineering students should be encouraged to adopt a deep approach to learning (Goodfellow et al. 2016). However, there has been limited research on the factors that contribute to deep learning in engineering education. This study aims to address this research gap by identifying the factors associated with deep learning and establishing their groupings.

Furthermore, there has been a growing focus on students' learning in higher education, with different learning methods used to assess their effectiveness. Summative and formative assessments are two popular systems that each have their advantages, but the relationship between these assessment types and deep learning factors has not been explored. This study aims to fill this gap by identifying the dimensions of deep learning and learning assessments, and testing the relationship between assessment types and the factors associated with deep learning for engineering students.

This paper follows a specific structure, commencing with an Introductory section. Subsequently, the "Literature review" section delves into the literature review and model development, while the "Methodology" section details the methodology employed for conducting research and analysis. The "Results" section presents the study's findings and hypothesis testing. Lastly, the "Discussions" section provides insights into the implications of the study, its limitations, and outlines possible future research directions.

Literature review

In this paper the literature consists of sub-sections to describe the deep learning and learning assessment themes, their associated components and relevant literature.

Deep learning concept. The deep learning approach involves comprehending the learning material, while the surface learning approach involves memorizing for the sake of passing exams (Tynjälä et al. 2005). Deep learning not only enhances students' knowledge, but also facilitates the practical application of that knowledge in real-life situations (Ditcher 2001). Student characteristics play a significant role in determining the learning approach. Previous studies on deep learning found that students who seek a thorough understanding of the material, critically engage with teaching materials, connect prior knowledge and experiences with new ideas, analyze arguments, and link evidence to conclusions are better suited for this approach (Karim et al. 2019; Beattie IV et al. 1997). Therefore, one way to promote deep learning is to personalize or customize the learning process based on students' characteristics (Cross 1998). This approach is characterized by student engagement, integration, synthesis, and reflection, and is supported by a collaborative leadership approach among educators (Laird et al. 2008; Cross 1998).

Factors associated with deep learning. The factors that influence deep learning are complex and multifaceted. According to Warburton (2003), the most significant factors that affect deep learning for engineering students are the learning environment, course content, and individual student factors. However, we argue that it is challenging for educational institutions or policymakers to manage individual student factors such as prior knowledge, personality, motivation, and workload. Educators can only design course content and create a learning environment that aligns with the unique characteristics of their students (Akareem and Hossain 2016). Therefore, this study will focus on the learning environment and course content/design as the primary factors that influence deep learning.

Learning environment. The existing literature suggests that deep learning is closely linked to students' level of engagement (Ramsden 1997), which promotes internalized understanding of subjects instead of just memorizing information for exams (Marton 1997). Therefore, it is crucial for higher education institutions to create an environment that fosters a strong personal interest in learning subjects (Warburton 2003). Educators must make the learning process relevant, utilize diverse content and teaching styles, and understand that imposing a deep learning approach is not possible by merely telling students it is required (Marton 1997). Consequently, understanding the learning environment is essential for implementing deep learning. A favorable learning environment can be ensured by considering multiple dimensions, including student-teacher relationships, student satisfaction and the commitment of support staff (Dunn 1995). These dimensions encompass critical features such as students' interest in learning a subject, access to resources, the effectiveness of learning assessments, and real-world applications of learning.

Efforts by teaching staff to make units interesting can have a significant positive impact on student learning. When teaching staff create engaging and interesting learning experiences, students are more likely to be motivated to learn and to participate actively in their own learning. The influential paper “Enhancing Student Learning: Seven Principles for Good Practice” by Chickering and Gamson (1987) identified that class activities such as asking questions, solving problems, and discussing ideas with teaching staff as one of the seven principles for good practice in undergraduate education. The paper notes that teaching that facilitates these activities can improve student motivation, attitudes, and learning outcomes. Bridglall (2013) reported higher levels of engagement with their studies (including finding their units interesting) were more likely to achieve high grades and report high levels of satisfaction with their university experience. The author identified one key factor in students’ higher levels of engagement with their studies: the teaching staff’s efforts to make the units interesting and engaging.

The learning environment can have a significant impact on student learning and perceptions of learning evaluation. The learning environment includes both physical environment including lighting, noise, temperature, and physical comfort as well as academic environment including effective communication and interaction between teacher and students, appropriate workload, and good teaching (Lizzio et al. 2002). The above study investigated the relationships between university students’ perceptions of their academic environment, their approaches to study, and academic outcomes. They found that learning outcomes are directly indirectly influenced by the learning environment, availability of necessary resources and their approaches to study.

Zhao et al. (2017) argued that the role of teachers has shifted from class-based teaching to guideline-based teaching to facilitate better learning. For deep learning to occur, teachers must put in efforts to make subjects interesting to students. A crucial aspect of the learning environment is the availability of resources that students can use during the learning process. Both online and offline resources can enhance students’ understanding of a subject (Barnes 1998), indicating that a greater quantity of physical and online resources leads to a better learning environment. Additionally, both learners’ social environment and the way learning is formally tested play a significant role (Baird et al. 2017). This is directly related to learning assessments, which are an essential component of the learning environment. When students have more options, they tend to choose the most effective way to learn (Chambers 1999). Traditional teaching practices do not usually incorporate real-world learning opportunities, leading to lower scores in deep learning among undergraduate students due to low morale and motivation resulting from a perceived lack of employment opportunities (Warburton 2003). To overcome this, educators should include real-world learning opportunities into course content. According to recent research on learning environments in Australasia, there is a positive correlation between innovative learning environments (ILEs) and students’ deep learning (Young et al. 2020).

Course content and design. The second factor that influences deep learning is course content and design, which includes three major dimensions: key concepts and themes, variety of learning opportunities, and relevance to engineering applications (Warburton 2003). These dimensions are reflected in peer learning, group learning, deep understanding and explanation of theories, involvement of industry professionals in lectures, industry visits, and flexibility of teaching and learning, as evidenced by Karim et al. (2015). Group learning has been found to be effective in promoting academic achievement, favorable attitudes toward learning, and increased persistence through science, technology, engineering, and mathematics (STEM) courses (Springer et al. 1999). Therefore, peer or group learning helps students understand the subject matter better than individual learning. Explaining theory in group learning environment is a clear sign of passive learning (Al-Hadad 2013), and educators should use this technique to improve student learning.

Real-world learning getting lot of attention as Universities want their graduates better ready for industry. Therefore, learning activities must be designed taking into consideration the importance, relevance and integration of theory and knowledge with professional practice to develop solutions to real-world issues (Karim et al. 2019). Real-life examples and videos can help students visualize and apply concepts in real-world situations, leading to a better understanding of the topic. Peer learning provides opportunities for students to collaborate and share their knowledge and ideas, enabling them to learn from each other and develop critical thinking and problem-solving skills. Group learning or tutorials allow students to receive personalized feedback and guidance from their instructor and peers, leading to improved comprehension and retention. Deep understanding and explanation of theories are essential for students to grasp complex concepts and apply them in different contexts. These approaches have been shown to improve student engagement, motivation, and academic achievement (Nilson 2016; Sambell and McDowell 1998). Therefore, incorporating a variety of instructional methods can enhance students’ learning experiences and help them achieve better academic outcomes.

Providing students with more choice and more voice can positively impact their learning outcomes. Giving students autonomy in their learning allows them to take ownership of their education, leading to increased engagement, motivation, and satisfaction (Deci and Ryan 2012). When students have a say in what they learn and how they learn it, they are more likely to be invested in the process, leading to deeper and more meaningful learning experiences (Furrer and Skinner 2003). Additionally, when students are given a voice in their education, they feel more valued and supported, leading to increased self-esteem and self-efficacy (Henderson and Mapp 2002). Involving students in the decision-making process can also improve the quality of teaching and learning, as it allows for more diverse perspectives and ideas to be considered (Cairns and Maloney 2017). Therefore, providing more choice and more voice for students can lead to better academic outcomes and contribute to their overall well-being.

In 2021, Zhang et al. investigated the development of a codesign process involving researchers, educators, and students (Zhang et al. 2021). They observed that the educator’s instructional practices improved as a result of this process and that there was a favorable shift in attitudes toward using student perception data to guide instructional improvements. Involving industry professionals in the curriculum enhances students’ engagement and improves learning (Vasiliou et al. 2013). Additionally, when students are given more flexibility and choice, they tend to choose the best approach and course instructor(s) for effective learning (Chambers 1999).

Therefore, students can engage in the learning process when they have the opportunity to work in groups, choose the instructor who can explain the theory, and learn from industry professionals. The following section discusses the aspects of learning evaluation related to deep learning.

Learning evaluation. Learning evaluation can be defined as an assessment of the knowledge acquired during a particular course (Calvert and Carroll 2005). In their study, Masuku et al. (2021)

highlighted the importance of assessment as both a pedagogical and evaluative tool to foster deep learning. The researchers demonstrated how various forms of assessment could facilitate better learning and improve critical thinking and analytical skills. They recommended that assessments should be clearly defined and aligned with learning objectives, taking into account the different levels of deep learning such as knowledge acquisition, comprehension, application, analysis, synthesis, and understanding of fundamental concepts related to the subject matter.

Assessments have direct relationship with the learning outcome. For example, the current authors have reported that designing effective assessments which requires critical thinking, teamwork and communications, as well as technical or discipline-specific skills results in learning outcomes relevant to industry demands and prepares the learners with 'job ready' skills (Karim et al. 2019). The type of assessment used can also affect learning and perceptions of learning evaluation. Assessment methods that are aligned with the learning objectives and that promote meaningful learning tend to be more effective in promoting deep learning (Biggs and Tang 2011).

Assessment methods are an integral part of education and help to measure a student's understanding and application of the course material. This paragraph explains various assessment methods that instructors use in classrooms. A problem-based assignment requires students to apply their knowledge to solve real-world problems. Open-book in class problem solving and open-book final examinations permit students to use their notes or textbooks to solve problems during the test. Multiple-choice question tests provide a quick and efficient way to evaluate students' knowledge of a subject. A seminar/presentation involves students presenting their research or a topic related to the course. Close-book in class problem solving and close-book final examinations test students' knowledge without the use of notes or textbooks. These methods are commonly used in classrooms and have been studied for their effectiveness in assessing students' understanding of the course material (Stevens and Levi 2005; Fredericksen and Collins 1989; Spandel 2012).

Among the commonly used assessment methods, Problem-based assignment is a task-based method that requires students to apply their knowledge and skills to solve real-world problems. Open-book in class problem solving and Open-book final examination allow students to use their textbooks and notes during the exam, but require them to demonstrate their understanding and problem-solving abilities in a limited amount of time. Multiple-choice question tests, on the other hand, are commonly used for assessing students' understanding of a particular subject or topic in a relatively short period of time. Seminar/Presentation is another method where students present their understanding of a topic or subject to the class, which allows them to showcase their presentation skills as well. Close-book in class problem solving and Close-book final examination are assessment methods where students are not allowed to use their textbooks or notes during the exam, and need to rely on their memory and understanding of the subject matter. These methods aim to test students' ability to recall and apply their knowledge in an exam setting. Each of these methods has its strengths and weaknesses and should be used based on the learning objectives and goals of the course. Some of these factors were previously identified by the authors in their previous study (Karim et al. 2015).

There are two different types of learning evaluation namely summative assessment and formative assessment (Bamford et al. 2012). Both types of assessment methods are discussed below.

Summative assessment. The definition states that summative assessment assigns students their course grade at the end of the semester (Al Kadri et al. 2011). Various forms of assessment methods have been identified in the literature for summative assessment, including open-book and close-book final examinations (Krasne et al. 2006), major reports throughout or at the end of the semester, and oral presentations (Nepal and Jenkins 2011). Karim et al. (2015) identified five items reflecting summative assessment, namely open-book final exams, close-book final exams, seminars/presentations, multiple-choice question tests and individual and/or group assignments.

Formative assessment. The other type of learning assessment is formative or problem-based assessment where students are given opportunity to improve their performances and grades (Bamford et al. 2012). Therefore, this type of assessment can enable students to take control of their own learning and become self-regulated learners (Nicol and Macfarlane-Dick 2006). It helps students to get feedback of their strengths and weaknesses in the course (Krasne et al. 2006). Existing literature identifies two different types of formative assessments: open-book problem solving and close-book problem solving (Krasne et al. 2006). These two types of formative/problem-based assessments are considered in this study for further analyses which are supported by Karim et al. (2015).

Hypotheses development. Based on the literature review presented, a conceptual model is proposed to better understand the nature of the interrelationships among learning assessment as considered the exogenous construct and deep learning as the endogenous construct. It is developed based on the previous concepts and the contemporary literature review in relevant fields of learning technologies. The proposed model is depicted in Fig. 1.

Studies by Baird et al. (2017), Tuunila and Pulkkinen (2015), Vos (2000) have shown that the type of assessment used can impact student learning. When students are aware of the assessment criteria and process, they tend to adjust their learning strategies accordingly. Therefore, enhancing the assessment process has the potential to improve students' learning, particularly in the area of deep learning (Vos 2000). The impact of assessment on learning and student behavior may vary depending on whether the assessment is formative or summative (Gielen et al. 2003).

According to Joughin (2010), summative assessment has a significant impact on student learning. Typically, course instructors provide a detailed course outline that includes information on the format and types of questions that will be included in the final exam. Students prepare for the exam based on this information and the instructor's instructions prior to the exam, which is known as the pre-assessment effect (Gielen et al. 2003). As a result, this paper proposes that summative assessments directly affect the factors that influence deep learning. The following hypotheses are suggested based on this proposition (Fig. 2):

Hypothesis (H1): There is a positive and significant relationship between proper summative assessment design and deep learning factors

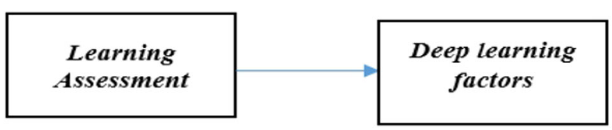

Fig. 1 Conceptual model. The conceptual model depicted in the figure hypothesizes a direct relationship between assessment methods and students' deep learning.

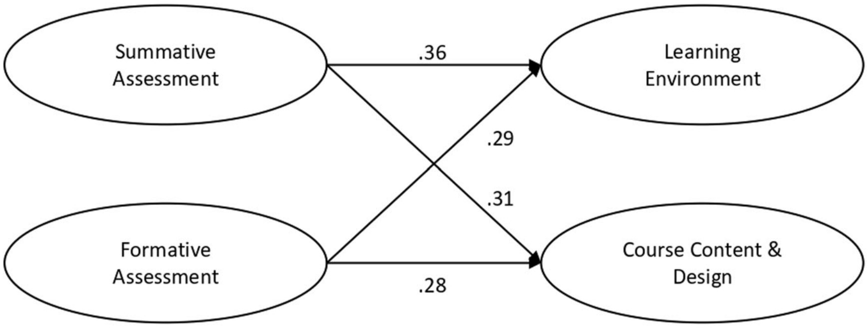

Fig. 2 Conceptual model including final results. The SEM results demonstrate positive relationships of both summative and formative assessments with learning environment and content design.

Table 1 Factors affecting deep learning and perceptions of learning evaluation.

Item no	Statement
Factors affecting deep learning	
i1	Efforts by the teaching staff to make the units interesting
i2	Learning environment in the university
i3	Flexibility in teaching and learning (lecture and assessment)
i4	Availability of necessary resources
i5	Effective assessment strategy
i6	Real-life examples and videos
i7	Peer learning
i8	Group learning/tutorials
i9	Deep understanding/explanation of theories
i10	Involvement of industry people in lectures
i11	More choice and more voice for student
i12	Field trips/industry visit
Learning evaluation	
a1	Problem-based assignment
a2	Open-book in class problem solving
a3	Open-book final examination
a4	Multiple-choice question test
a5	Seminar/Presentation
a6	Close-book in class problem solving
a7	Close-book final examination

(H1a): There is a positive and significant relationship between summative assessment and learning environment.

(H1b): There is a positive and significant relationship between summative assessment and course content and design.

López-Pastor et al. (2013) have demonstrated that, similar to summative assessment, formative assessment also has a significant impact on student learning. Furthermore, a comprehensive meta-analysis of classroom assessments by Black and Wiliam (1998) has revealed that formative assessment plays a crucial role in promoting effective teaching and learning practices. Chin and Brown (2000) have also provided numerous examples of how formative assessment can enhance learning outcomes in higher education, while a study by Gauntlett (2007) has reported positive effects of formative assessment on student learning. Additionally, negotiated assessment with students has been found to improve student engagement in some studies (Boud and Falchikov 2007; López-Pastor 2008). Therefore, this paper proposes that formative assessments directly affect the factors that influence deep learning. The second hypothesis reads as follows:

Hypothesis (H2): There is a positive and significant relationship between formative assessment proper design and deep learning factors

(H2a): There is a positive and significant relationship between formative assessment and learning environment.

(H2b): There is a positive and significant relationship between formative assessment and course content and design.

Methodology

This section is divided into four sub-sections. The first section "Sample size and respondent demography" is about discussing the sample size and respondents profile. The second section "Questionnaire development and administration" is about the survey instrument developed for this study. The items of deep learning and leaning evaluation used in this survey is extracted from literature provided above. The third section "Instrument reliability" provides the reliability of the data while the fourth section "Structural equation modeling" is about the data analysis techniques utilized to empirically investigated the developed hypotheses.

Sample size and respondent demography. Upon developing the proposed questionnaire, it is directed to 600 engineering students of Civil and Mechanical Engineering from Faculty of Engineering (FoE), Queensland University of Technology (QUT). The recipients were randomly selected from 2nd, 3rd and final years students. From the 600 recipients, 243 completed responses and returned to the authors. The sample size is considered very good for further analysis as per recommendation of Oke et al. (2012). Oke et al. (2012) recommended that that the suitable sample size for SEM should be within the range of 200 and 400. Common method biases (CMB) can arise in survey research, which can result in an inflated correlation among the variables, which may not reflect their true relationships. By ensuring participants' anonymity and confidentiality, researchers can reduce the risk of social desirability bias, which occurs when participants provide answers that are socially desirable instead of their actual beliefs or behaviors (Paulhus 1984). In line with Queensland University of Technology (QUT) ethical policy, strict confidentiality and anonymity of the respondents were maintained.

Among the respondents, about 70% were Australian and 30% were international students from 9 countries. Majority of engineering students at Queensland University of Technology are male, which is also reflected in the survey response as 82.6% respondents are male, and the rest are female. The average age of the respondents was 24 years. Although most of the respondents are regular students within the age range 20–23, there are some matured students with age up to 39 years. The students were asked to mention their average GPA and the reported GPA were between 4 and 7 in the scale of 1–7. The demographic data shows that the respondents were similar to the student population of the Faculty of Engineering at QUT.

Questionnaire development and administration. The instrument items for factors affecting deep learning and perceptions of

Table 2 Cronbach's alpha.

Factor	Number of items	Cronbach's alpha
Factors affecting deep learning		
Learning environment	4	0.819
Course content and design	4	0.751
Learning evaluations		
Summative assessment	3	0.764
Formative assessment	2	0.792

assessment methods used in this study was developed from extensive literature review presented in the "Literature review" section above. List of these items used in the survey is depicted below in Table 1.

A traditional paper-based survey questionnaire was distributed among the students during the last week of the academic semester. The first section of the questionnaire consisted the demographic and background information of the students, and the second section consisted of statements related to factors affecting deep learning in engineering courses and perceptions of assessment methods that enhance student attitude and understanding on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Instrument reliability. To test the internal consistency (reliability) of the scales representing two factors affecting deep learning, and the two factors representing perception of assessment methods, we calculated the Cronbach's alpha reliability coefficient. It measures the items' strength with other items of the same factor. Table 2 shows the results, where Cronbach's alpha values ranged from 0.751 to 0.819, which is considered as good reliability for research purpose (Nunnally 1978).

Structural equation modeling. Structural equation modeling (SEM) is used in this research which is the appropriate technique for hypotheses testing. It can be defined as a multivariate technique that permits for the simultaneous analysis of multiple, interrelated relationships between concepts in a model to be considered (Groenland and Stalpers 2012). Further, it has two aspects; the measurement model followed by the path analysis which is the structural model (Dugard et al. 2010).

Results

Results in this paper are depicted to show the analysis of the survey questionnaire and testing the postulated hypotheses consistent with the developed model.

Response rate. As mentioned above, after sending the questionnaire, 243 completed responses from the undergraduate students at QUT were returned to the authors. Thus, based on the number of the submitted surveys and the acceptable returned questionnaire, the response rate which is estimated as 40% is considered high compared to other studies in relevant field. The respondents in this study were from mechanical and civil engineering disciplines. The sample size of in this study is considered appropriate for further analysis in SEM as per recommendation of Oke et al. (2012).

Exploratory factor analysis (EFA) analysis. According to Tabachnick et al. (2013), confirmatory factor analysis (CFA) is appropriate when a researcher wants to confirm or test a pre-existing theory or hypothesis, while EFA is appropriate when a researcher wants to develop a theory or generate hypotheses about the underlying structure of the data. As this study proposes a new hypotheses, EFA is deemed the right approach for the factor analysis.

Table 3 Factor loadings and variance explained by each factor.

	Factors		
Characteristics	F1	F2	
Factors affecting deep learning			
i4	0.808		
i5	0.780		
i6	0.752		
i1	0.728		
i8		0.821	
i7		0.760	
i10		0.660	
i9		0.646	
% Variance	47.939	13.523	
Kaiser-Meyer-Olkin measure of sampling adequacy			0.859
Bartlett's test of sphericity			0.000
Learning evaluation			
a7	0.856		
a3	0.842		
a5	0.765		
a2		0.912	
a6		0.899	
% Variance	44.944	29.354	
Kaiser-Meyer-Olkin measure of sampling adequacy			0.610
Bartlett's test of sphericity			0.000

Extraction method—Principal component analysis (PCA); Based on varimax rotation (Hendrickson and White 1964).

Accordingly, prior to the use of SEM, an initial exploratory analysis was performed using principal component analysis (PCA) and Kaiser Meyer Olkin (KMO) and Bartlett's test for sphericity were used to assess the suitability of PCA. A series of exploratory factor analyses (EFA) were then conducted using IBM SPSS statistics to identify the common aspects and factors.

Following the procedure of Senocak (2009), we conducted the KMO test and the Bartlett's test of sphericity (Bartlett 1954; Kaiser 1974) before determining the EFA to check the suitability of the data for an EFA. KMO value above 0.50 (Kaiser 1974) and significance (below 0.05) of Bartlett's test of sphericity (Bartlett 1954) indicates adequacy for a factor analysis.

Two separate EFAs were conducted, one for factors that are associated with deep learning and the other for perceptions of assessment methods that enhance student attitude and understanding. To ensure the stability of the analysis, a principal component analysis with varimax rotation (Hendrickson and White 1964), an eigenvalue of 1 (Fritz et al. 1984), and a factor loading greater than 0.65 (Jaeger and Adair 2014) were utilized. Additionally, cross-loaded items were deleted to obtain a more accurate grouping (Keery et al. 2004). Four items (i2, i3, i11, and i12) from the factors affecting deep learning and two items (a1 and a4) from the perceptions of assessment methods did not appear in any groups due to poor factor loading with these groups. The refined results of the EFAs are presented in Table 3.

The EFA revealed two underlying factors for items affecting deep learning, explaining 61.5% of the variance, and two items for perceptions of assessment methods, explaining 74.3% of the variance. While factor loadings of 0.30 or higher are considered to reflect sufficiently strong factor loading (Martin-Dunlop and Fraser 2008), we followed a more conservative approach, as suggested by Senocak (2009), by using a factor loading of 0.40 or higher to identify stronger factors. It was found that four characteristics loaded highly on factor 1, four characteristics

Table 4 Mean and standard deviation values of the deep learning factors.

Item	Factors	*N*	Mean	Std. deviation
i1	Efforts by the teaching staff to make the units interesting	243	4.24	0.886
i4	Availability of necessary resources	243	3.94	0.893
i5	Effective assessment strategy	243	3.99	0.935
i6	Real-life examples and videos	243	4.11	0.864
i7	Peer learning	243	3.35	1.178
i8	Group learning/tutorials	243	3.83	1.059
i9	Deep understanding/explanation of theories	243	3.79	0.894
i10	Involvement of industry people in lectures	243	3.71	0.954

Table 5 Mean and standard deviation values of assessment items.

Item	Assessment types	*N*	Mean	Std. deviation
a2	Open-book in class problem solving	220	3.14	1.174
a3	Open-book final examination	221	4.02	1.007
a5	Seminar/Presentation	222	3.67	1.045
a6	Close-book in class problem solving	221	3.13	1.256
a7	Close-book final examination	219	3.93	1.005

loaded highly on factor 2 for factors affecting deep learning, and three characteristics loaded highly on factor 1 and two characteristics loaded highly on factor 2 for perceptions of assessment methods.

The mean and standard deviation values of the items mentioned above (Table 3) for deep learning and assessments are presented in Tables 4 and 5, respectively. We also conducted normality tests for these items using the widely recognized Kolmogorov-Smirnov and Shapiro-Wilk methods and determined that these items met the normality requirements at a statistically significant level.

The next step in the EFA involved identifying commonalities among the characteristics to interpret the general meaning of each factor. Based on the high loadings of the factors affecting deep learning, the four items of factor 1 are associated with environmental features that impact student learning (learning environment), while the later four characteristics of factor 2 are associated with course content and design. The perception of assessment methods includes two factors, with the first factor comprising three characteristics related to the traditional end-of-semester examination-based assessment system (summative assessment), and the second factor composed of characteristics related to assessment based on problem-solving throughout the semester (formative assessment).

Results of hypotheses testing. The model we developed tested the relationship between students' perception of assessment methods and factors affecting deep learning. In this model, summative assessment and formative assessment methods were considered exogenous latent constructs, while learning environment and course content and design were considered endogenous latent constructs. To address the limitations of other statistical methods in analyzing variable relationships, we used structural equation modeling (SEM) with AMOS version 25 (Blunch 2012).

After validating the measurement models for the latent constructs, the next step was to construct the structural model so that the postulated hypotheses could be tested. Goodness of fit indices from the AMOS software were used to assess the appropriateness of the structural model developed in this study. The indices obtained were as follows: the chi-square value of 249.944 with 60 degrees of freedom ($\chi^2/\mathrm{df} = 4.166$, which is less than 5), root mean square error of approximation (RMSEA) of 0.074, comparative fit index (CFI) of 0.921, incremental fit index (IFI) of 0.923, Tucker-Lewis index (TLI) of 0.898, goodness of fit index (GFI) of 0.925, and adjusted goodness of fit index (AGFI) of 0.887. These values indicated good model fitness, as they were all above the recommended threshold of 0.9. This designates the unidimensionality of the measure, as per the guidelines of Anderson and Gerbing (1988) and Zainudin (2015).

The testing of causal hypotheses, as per the developed model in this paper, is illustrated in Table 6. Each hypothesis can be accepted and supported if it is both statistically and practically significant. A p value less than 0.05 is considered statistically significant, and a standardized regression weight (β) greater than 0.2 is considered practically significant.

Table 6 presents the path coefficients for each hypothesized relationship in the structural model, and the results demonstrate that the impact of summative assessment on learning environment ($\beta = 0.36$, $t = 4.208$) and course content and design ($\beta = 0.31$, $t = 3.573$) are statistically significant. Likewise, the effect of formative assessment on learning environment ($\beta = 0.29$, $t = 3.295$) and course content and design ($\beta = 0.28$, $t = 3.128$) is also considered practically significant. Therefore, all hypotheses presented in this study are supported.

Discussions

The aim of this study was to investigate the connections between factors affecting deep learning and learning assessment in the context of Engineering Education, using a structured questionnaire survey. The following discussion is based on the results and findings presented in the previous section, in line with the study's objectives.

The result of EFA confirms in principle the findings of Warburton (2003) who identified the factors affecting deep learning. Out of three factors, students' individual factors were excluded because of two reasons: (1) it is very difficult to control students' prior knowledge, experience, personality and morale by the educators, and (2) most of the universities follow equal opportunity for the students, where excluding a student based on lower score in the individual factors will violate the equal opportunity policy.

The EFA conducted in this study revealed that there are two factors that influence deep learning in engineering courses, namely the learning environment and the course content and design, which is consistent with Warburton's (2003) model. However, four items from the deep learning construct, namely: (1) the university learning environment, (2) student choice and input, (3) teaching and learning flexibility, and (4) field trips/industry visits had poor model fit. It is not surprising as the traditional university learning approach typically does not allow for varying study methods, and course instructors design the course with limited student input. Additionally, arranging field trips or industry visits for courses may not always be feasible due to the nature of different courses, and they may not directly contribute to students' grades. This may be the reason why students did not consider these factors to be the effective factors for deep learning, based on their responses to the questionnaire.

The EFA also revealed that there are two distinct factors in learning evaluation, namely summative assessment and formative

Table 6 Path parameters estimates and hypotheses testing.

Hypotheses	Relationships	β	t	p	Comments	Result
H1(a)	Summative assessment → Learning environment	0.36	4.208	0.000	Significant	Supported
H1(b)	Summative assessment → Course content and design	0.31	3.573	0.000	Significant	Supported
H2(a)	Formative assessment → Learning environment	0.29	3.295	0.000	Significant	Supported
H2(b)	Formative assessment → Course content and design	0.28	3.128	0.002	Significant	Supported

assessment, which supports the findings of Bamford et al. (2012). Summative assessment generally refers to final exams where students cannot improve their performance once the exam is complete, while formative assessment generally refers to in-class evaluations that allow for gradual improvement (Bamford et al. 2012).

The EFA results indicated that summative assessment has the strongest influence on the learning environment. This means that close and open-book examinations, as well as end-of-term presentations, strongly impact whether students utilize the available resources, engage with the teacher's efforts to make the unit interesting, follow the assessment strategy, and pay attention to real-life examples and videos presented during classes. Essentially, when students are aware of the type and format of the final exam and end-of-term presentation, they strive to make the best use of the university's environmental features to perform well in those assessments. The study also found a significantly positive relationship between summative assessment and course content and design, which suggests that students' comprehension of final exams motivates them to work in groups, understand relevant theories, and gain insights from industry professionals. In summary, both factors that affect deep learning are significantly influenced by summative assessment.

The study found that formative assessment has a significant influence on both the learning environment and course content and design. This means that both open and close-book in-class problem-solving based assessments strongly impact various features of the learning environment, including the teacher's efforts to make the unit interesting, the availability of resources, an effective assessment strategy, and the use of real-life examples and videos. Additionally, formative assessment affects course design features, such as peer and group learning, a deep understanding of theories, and the involvement of industry professionals in the class lecture. These results are not surprising, as students tend to engage more with the learning process when they have the opportunity to improve their performance. Although the effect of formative assessment on both factors that influence deep learning is similar, the impact on the learning environment is slightly stronger than that on course content and design.

Accordingly, both summative and formative learning have strong influence on the factors affecting deep learning. The findings of this study indicate that educators should utilize both summative and formatting assessment to engage students with the learning process to make it effective. The combination of formative and summative assessments has been shown to improve student learning outcomes. Formative assessments, such as classroom discussions, provide ongoing feedback to students, allowing them to adjust their learning strategies and focus on areas that need improvement. Summative assessments, such as exams or final projects, provide a measure of student achievement at the end of a unit or course. Together, these assessments can provide a comprehensive picture of student learning and help teachers tailor their instruction to meet the needs of individual students. According to a review of research on assessment practices, combining formative and summative assessments has been shown to improve student achievement and engagement (Black and Wiliam 1998). Additionally, a study conducted by the Educational Testing Service found that incorporating formative assessments into instruction improved student performance on summative assessments (Pellegrino et al. 2016). Therefore, the integration of formative and summative assessments can be a powerful tool for improving student learning outcomes.

Educators can greatly improve the student learning experience by carefully designing course content and assessments. Course content should be organized in a logical and meaningful way, with clear learning objectives that guide the students toward the desired outcomes. Effective course design should also take into consideration the diverse needs and backgrounds of students, and provide opportunities for active engagement, such as discussions, projects, and hands-on activities. Assessments should also be designed with the learning objectives in mind, and should align with the content and format of instruction. Assessment methods should be varied and include both formative and summative assessments to provide students with ongoing feedback and to monitor progress toward the learning objectives. Additionally, educators should ensure that assessments are fair, valid, and reliable, and provide students with opportunities to demonstrate their learning in different ways. By properly designing the course content and assessments, educators can create an engaging and effective learning environment that fosters student achievement and success.

Limitations and future research. Similar to other studies conducted in relevant field, this study is not without limitations which lead to opportunity for further research. Firstly, this study is based on survey where engineering students provided their perception about factors affecting deep learning and their relationships with evaluation system. Sometimes perception-based survey does not reflect the actual interrelationships among variables. An experiment-based study may provide a better insight of such relationships. Future studies can include more variables, such as the outcomes of deep learning as dependent variable.

In this study, students of only one university have been considered. In order to generalize the findings, similar study needs to be done in other universities. Common Method Bias (CMB) is a systematic error that may occur in survey research when the method of data collection influences the response patterns of participants. CMB is a potential source of bias in survey research where respondents may consistently respond to questions in a similar way due to the influence of a common method or instrument. To minimize the effects of CMB in survey research, researchers used several methods such as use different data collection methods, use longitudinal study, use reverse-worded items, use of statistical techniques, and maintain anonymity and confidentiality of respondents. This study only undertaken the approach of 'anonymity and confidentiality' to minimize CMB. Also, the relationships of the factors with demographic variables such as gender, age, nationality and year of study have not been explored.

Conclusion and implications

This study holds significance both theoretically and practically. The theoretical model developed to explore the concept of deep learning can serve as a framework to support modern learning methodologies and enhance the teaching and learning process. Furthermore, it is noteworthy that the study incorporated both types of assessments (formative and summative) into the same multidimensional model. Additionally, the empirical investigation was conducted using real data from a university in Australia (QUT).

The study aimed to achieve two objectives: firstly, to identify the dimensions of factors that influence deep learning and learning evaluation, and secondly, to determine the influence of learning evaluation dimensions on the dimensions of factors affecting deep learning. The study identified two dimensions of factors affecting deep learning—learning environment, and course content and design. Similarly, two dimensions of learning evaluation were identified—summative and formative assessments. The findings suggest that summative assessment has the strongest influence on both dimensions of factors affecting deep learning.

The findings suggest that engineering educators, should carefully design the assessment system of a course which lead the students to deep learning. The learning assessments should be a combination of both summative assessment and formative assessment because both type of assessments have significant influence on deep learning.

Data availability

Data can be provided on request.

Received: 14 July 2022; Accepted: 18 December 2023;
Published online: 02 January 2024

References

Akareem HS, Hossain SS (2016) Determinants of education quality: what makes students' perception different? Open Rev Educ Res 3(1):52–67

Al Kadri HM, Al-Moamary MS, Magzoub ME, Roberts C, van der Vleuten C (2011) Students' perceptions of the impact of assessment on approaches to learning: a comparison between two medical schools with similar curricula. Int J Med Educ 2:44

Al-Hadad O (2013) Student–centered learning, analyzing approaches and techniques in University Malaysia Pahang

Anderson JC, Gerbing DW (1988) Structural equation modeling in practice: a review and recommended two-step approach. Psychol Bull 103(3):411

Baird J-A, Andrich D, Hopfenbeck TN, Stobart G (2017) Assessment and learning: fields apart? Assess Educ Princ Policy Pract 24(3):317–350

Bamford D, Karjalainen K, Jenavs E (2012) An evaluation of problem-based assessment in teaching operations management. Int J Oper Prod Manag 32(12):1493–1514

Barnes BE (1998) Creating the practice-learning environment: using information technology to support a new model of continuing medical education. Acad Med 73(3):278–281

Bartlett MS (1954) A further note on the multiplying factors for various X2 approximations in factor analysis. J R Stat Soc 16(2):296–298

Beattie IV V, Collins B, McInnes B (1997) Deep and surface learning: a simple or simplistic dichotomy? Account Educ 6(1):1–12

Beering SC (2007) Moving forward to improve engineering education. National Science Foundation, Washington DC

Biggs J, Tang C (2011) Teaching for quality learning at university. McGraw-Hill Education, UK

Black P, Wiliam D (1998) Assessment and classroom learning. Assess Educ Princ policy Pract 5(1):7–74

Blunch N (2012) Introduction to structural equation modeling using IBM SPSS statistics and AMOS. Sage

Boud D, Falchikov N (2007) Rethinking assessment in higher education: learning for the longer term. Routledge

Bridglall BL (2013) Teaching and learning in higher education: studies of three student development programs. Lexington Books

Calvert C, Carroll J (2005) Training for ERP: does the IT training literature have value? In Proceedings of the sixteenth Australiasian conference on information systems. University of Technology, Sydney, pp 1–12

Cairns E, Maloney C (2017) Student voice in higher education: the importance of inclusive partnerships. High Educ Res Dev 36(2):321–334

Chambers DW (1999) Faculty ratings as part of a competency-based evaluation clinic grading system. Eval Health Prof 22(1):86–106

Chickering AW, Gamson ZF (1987) Seven principles for good practice in undergraduate education. AAHE Bulletin, p 3–7

Chin C, Brown DE (2000) Learning in science: a comparison of deep and surface approaches. J Res Sci Teach 37(2):109–138

Cross KP (1998) Classroom research: implementing the scholarship of teaching. N Dir Teach Learn 75:5–12

Deci EL, Ryan RM (2012) Motivation, personality, and development within embedded social contexts: an overview of self-determination theory. The Oxford handbook of human motivation 18(6):85–107

Ditcher AK (2001) Effective teaching and learning in higher education, with particular reference to the undergraduate education of professional engineers. Int J Eng Educ 17(1):24–29

Dunn SV (1995) The development of a clinical learning environment scale. J Adv Nurs 22(6):1166–1173

Dugard P, Todman J, Staines H (2010) Approaching multivariate analysis a practical introduction, 2nd edn. Routledge, New York, NY

Fredericksen EE, Collins A (1989) A systems approach to educational *testing*. Educ Res 18(9):27–32

Fritz W, Hilger H, Raffée H, Silberer G, Förster F (1984) Testnutzung und Testwirkungen im Bereich der Konsumgüterindustrie. Warentest und Unternehmen–Nutzung, Wirkungen und Beurteilung des vergleichenden Warentests in Industrie und Handel, p 27–114

Furrer C, Skinner E (2003) Sense of relatedness as a factor in children's academic engagement and performance. J Educ Psychol 95(1):148

Gauntlett N (2007). Literature review on formative assessment in higher education. Middlesex University. http://www.mdx.ac.uk/HSSc/cetl/docs/formrevnov07.pdf

Geitz G, Brinke DJT, Kirschner PA (2015) Goal orientation, deep learning, and sustainable feedback in higher business education. J Teach Int Bus 26(4):273–292

Gielen S, Dochy F, Dierick S (2003) Evaluating the consequential validity of new modes of assessment: the influence of assessment on learning, including pre-, post-, and true assessment effects. In Optimising new modes of assessment: in search of qualities and standards, Dordrecht: Springer Netherlands, p 37–54

Goodfellow I, Bengio Y, Courville A, Bengio Y (2016) Deep learning, vol. 1. MIT Press, Cambridge

Groenland E, Stalpers J (2012) Structural equation modeling. A verbal approach. Nyenrode Research Paper Series (12-02)

Gupta SK, Ashwin TS, Guddeti RMR (2019) Students' affective content analysis in smart classroom environment using deep learning techniques. Multimed Tools Appl 78:25321–25348

Henderson AT, Mapp KL (2002) A new wave of evidence: the impact of school, family, and community connections on student achievement. National Center for Family & Community Connections with Schools

Hendrickson AE, White PO (1964) Promax: a quick method for rotation to oblique simple structure. Br J Math Stat Psychol 17(1):65–70

Jaeger M, Adair D (2014) The influence of students' interest, ability and personal situation on students' perception of a problem-based learning environment. Eur J Eng Educ 39(1):84–96

Joughin G (2010) The hidden curriculum revisited: a critical review of research into the influence of summative assessment on learning. Assess Eval High Educ 35(3):335–345

Kaiser HF (1974) An index of factorial simplicity. Psychometrika 39(1):31–36

Karim A, Campbell M, Hasan M (2019) A new method of integrating project-based and work-integrated learning in postgraduate engineering study. Curric J 31:1–17

Karim A, Fawzia S, Islam MM (2015) Factors affecting deep learning of engineering students. In: Proceedings of the 26th annual conference of the Australasian Association for Engineering Education (AAEE2015), School of Engineering, Deakin University, p 1–8

Khilji N (2022) The use of virtual learning environment for students effective engagement in the Higher Education Institutions through knowledge management and blended learning. Eur Conf Knowl Manag 23(1):601–610

Keery H, Van den Berg P, Thompson JK (2004) An evaluation of the Tripartite Influence Model of body dissatisfaction and eating disturbance with adolescent girls. Body Image 1(3):237–251

Krasne S, Wimmers PF, Relan A, Drake TA (2006) Differential effects of two types of formative assessment in predicting performance of first-year medical students. Adv Health Sci Educ 11:155–171

Laird TF, Shoup R, Kuh GD, Schwarz MJ (2008) The effects of discipline on deep approaches to student learning and college outcomes. Res High Educ 49:469–494

Lizzio A, Wilson K, Simons R (2002) University students' perceptions of the learning environment and academic outcomes: implications for theory and practice. Stud High Educ 27(1):27–52

López-Pastor V (2008) Implementing a formative and shared assessment system in higher education teaching. Eur J Teach Educ 31(3):293–311
López-Pastor VM, Pintor P, Muros B, Webb G (2013) Formative assessment strategies and their effect on student performance and on student and tutor workload: the results of research projects undertaken in preparation for greater convergence of universities in Spain within the European Higher Education Area (EHEA). J Furth High Educ 37(2):163–180
Lynch R, McNamara PM, Seery N (2012) Promoting deep learning in a teacher education programme through self-and peer-assessment and feedback. Eur J Teach Educ 35(2):179–197
Martin-Dunlop C, Fraser BJ (2008) Learning environment and attitudes associated with an innovative science course designed for prospective elementary teachers. Int J Sci Math Educ 6:163–190
Marton F (1997) Approaches to learning. In The experience of learning, 2nd edn. Edinburgh: Scottish Academic Press, p 39–58
Marton F, Säljö R (1976) On qualitative differences in learning: I—outcome and process. Br J Educ Psychol 46(1):4–11
Masuku MM, Jili NN, Sabela PT (2021) Assessment as a pedagogy and measuring tool in promoting deep learning in institutions of higher learning. Int J High Educ 10(2):274–283
Nepal KP, Jenkins GA (2011) Blending project-based learning and traditional lecture-tutorial-based teaching approaches in engineering design courses. In: Australasian Association for Engineering Education (AAEE) conference proceedings, p 338–343
Nicol DJ, Macfarlane-Dick D (2006) Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. Stud High Educ 31(2):199–218
Nilson LB (2016) Teaching at its best: a research-based resource for college instructors. John Wiley & Sons
Nunnally JC (1978) Psychometric theory, 2nd edn. Mcgraw Hill Book Company
Oke AE, Ogunsami DR, Ogunlana S (2012) Establishing a common ground for the use of structural equation modelling for construction related research studies. Australas J Constr Econ Build 12(3):89–94
Paulhus DL (1984) Two-component models of socially desirable responding. J Personal Soc Psychol 46(3):598–609. https://doi.org/10.1037/0022-3514.46.3.598
Pellegrino JW, DiBello LV, Goldman SR (2016) A framework for conceptualizing and evaluating the validity of instructionally relevant assessments. Educ Psychol 51(1):59–81
Ramsden P (1997) The context of learning in academic departments. Exp Learn 2:198–216
Ramsden P (2003) Learning to teach in higher education. Routledge
Sambell K, McDowell L (1998) The construction of the hidden curriculum: messages and meanings in the assessment of student learning. Assess Eval High Educ 23(4):391–402
Senocak E (2009) Development of an instrument for assessing undergraduate science students' perceptions: the problem-based learning environment inventory. J Sci Educ Technol 18:560–569
Spandel V (2012) Creating writers: 6 traits, process, workshop, and literature. Pearson Higher Ed
Springer L, Stanne ME, Donovan SS (1999) Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis. Rev Educ Res 69(1):21–51
Stevens J, Levi A (2005) Introduction to rubrics: an assessment tool to save grading time, convey effective feedback, and promote student learning. Stylus Publishing, LLC
Tabachnick BG, Fidell LS, Ullman JB (2013) Using multivariate statistics, vol 6, Pearson; Boston, MA, p 497–516
Tuunila R, Pulkkinen M (2015) Effect of continuous assessment on learning outcomes on two chemical engineering courses: case study. Eur J Eng Educ 40(6):671–682
Tynjälä P, Salminen RT, Sutela T, Nuutinen A, Pitkänen S (2005) Factors related to study success in engineering education. Eur J Eng Educ 30(2):221–231
Vasiliou C, Ioannou A, Zaphiris P (2013) Technology enhanced PBL in HCI education: a case study. In: Human-computer interaction–INTERACT 2013: 14th IFIP TC 13 international conference, Cape Town, South Africa, September 2–6, proceedings, part IV 14, Springer, Berlin Heidelberg, p 643–650
Vos H (2000) How to assess for improvement of learning. Eur J Eng Educ 25(3):227–233
Warburton K (2003) Deep learning and education for sustainability. Int J Sustain High Educ 4(1):44–56
Young F, Cleveland B, Imms W (2020) The affordances of innovative learning environments for deep learning: Educators' and architects' perceptions. Aust Educ Res 47(4):693–720
Zainudin A (2015) SEM made simple. Agentle approach to learning structural equation modeling. MPWS Publisher, Selangor
Zhang L, Jackson HA, Yang S, Basham JD, Williams CH, Carter RA (2021) Codesigning learning environments guided by the framework of Universal Design for Learning: a case study. Learn Environ Res 25:1–19
Zhao H, Piao M, Li L, Liu D (2017) The application of network teaching in applied optics teaching. In Education and Training in Optics and Photonics. Optica Publishing Group, Washington, DC 20036 USA

Acknowledgements

The authors acknowledge the contributors who helped in analyzing the survey data.

Author contributions

SF: concept development, instrument design, questionnaire survey, data analysis, manuscript development, review and modification, preparation of responses to reviewer comments, final review. AK: concept development, instrument design, questionnaire survey, data analysis, manuscript development, review and modification, preparation of responses to reviewer comments.

Competing interests

The authors declare no competing interests.

Ethical approval

All procedures performed in this study were in accordance with the ethical standards of the university as ethical clearance and approval were granted by the university. Approval number 1700001004, approving body: Office of Research Ethics and Integrity, QUT.

Informed consent

Informed consent was obtained from all participants.

Additional information

Correspondence and requests for materials should be addressed to Azharul Karim.