



## OPEN ACCESS

## EDITED BY

Reza Fauzan,  
Banjarmasin State Polytechnic, Indonesia

## REVIEWED BY

Rakha Asyofi,  
Universitas Hang Tuah, Indonesia  
Anis Amna,  
Brawijaya University, Indonesia  
Noor Cholis Basjaruddin,  
Politeknik Negeri Bandung, Indonesia

## \*CORRESPONDENCE

Safiah Mohammed Alamr  
✉ sma2v21@soton.ac.uk

RECEIVED 02 April 2025

ACCEPTED 11 August 2025

PUBLISHED 05 September 2025

## CITATION

Alamr SM, Leon Urrutia M and Carr LA (2025)  
success factors of E-assessment in computer  
science for higher education: perspectives  
from instructors.  
*Front. Educ.* 10:1595171.  
doi: 10.3389/feduc.2025.1595171

## COPYRIGHT

© 2025 Alamr, Leon Urrutia and Carr. This is  
an open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Success factors of E-assessment in computer science for higher education: perspectives from instructors

Safiah Mohammed Alamr\*, Manuel Leon Urrutia and  
Leslie Alan Carr

Web and Internet Science Group, School of Electronics and Computer Science, Faculty of  
Engineering and Physical Sciences, University of Southampton, Southampton, United Kingdom

Assessment is an essential element in higher education processes, and its use is becoming increasingly widespread with the digitalisation of universities. E-assessment plays a major role in the success of the educational process due to its benefits, such as improving students' experience and performance and optimising instructors' effort. This study aims to identify success factors for the use of e-assessment methods from instructors' perspectives, focusing on computer science (CS) instructors who assess coding assignments in higher education in the UK and Saudi universities. A mixed-methods approach was employed, combining 16 interviews with CS instructors in the UK and Saudi universities and a survey as the quantitative element. Survey data revealed clear disparities in the perception of personalised feedback between the UK and Saudi Arabia. Findings indicate a lack of training courses for instructors and a shortage of specific automated tools for assessing coding. Few tools or systems are available for providing automated and personalised feedback, and existing e-assessment tools do not meet instructors' expectations regarding flexibility, accessibility, scalability, practicality, validity, reliability, and authenticity. Instructors currently spend significant time and effort assessing coding assignments, which may hinder learning outcomes and the enhancement of educational activities. This thesis proposes a comprehensive framework outlining success factors for effective assessment, which can benefit e-assessment tool suppliers, instructors, and universities, ultimately enhancing the quality of instructors' assessments.

## KEYWORDS

E-assessment, higher education, computer science education, E-assessment framework, coding assessment

## Introduction

E-assessment (EA) tools at universities, is arguably the linchpin of modern higher education. Its origin arguably dates back to the early 20th century, where a number of attempts were directed at automation related to grading processes with enhancement in scalability and efficiency (Pressey, 1926). Over time, especially with the introduction of the World Wide Web during the 1990s, these EA systems began to capture greater momentum, which offered heightened access and functionality (Combéfis, 2022).

There is a quick improvements in many fields, particularly in computer science higher education, transpired with the involved of the World Wide Web in the 1990s. Subsequently, other companies implemented their proprietary EA systems. It was not until the UK had come up with standardizing the idea of EA through the Joint Information Systems Committee that formulated principles and guidelines for implementation in its education system (Gilbert et al., 2009).

Some organizations like Cisco, Intel, and Microsoft launched several frameworks, one of which was “Transforming Education: Assessing and Teaching 21st Century Skills,” that included the role of EA in confidence-building of new competencies of learners (Care et al., 2012). However, notwithstanding these advances, there remain some issues concerning certain gaps in question automation, feedback mechanisms, and instructor training. Which detract from its effectiveness particularly in coding exams, both from pedagogical and practical viewpoints.

In the continuous development of higher education, EA is seen as a characteristic that differentiates, integrating technology into all components of the assessment process. E-assessment, abbreviated as E-assessment, is the integration of technological innovations into the assessment process (Almuhanna, 2023). The development of this concept has caused a high variation in the presentations, leading to computer-based tests, web-based exams, technology-supported assignments, and automated grading systems (Palmer, 2023). EA is versatile, and it may appear in diverse contexts, such as hybrid learning environments, fully online settings, and conventional educational institutions using digital tools (Loureiro and Gomes, 2022). EA is a significant integration of education and technology, with the emphasis placed on expanding the scope, effectiveness, and scalability of the process. Notably, its scope goes over and above delivery mechanisms to include complex algorithms and analytics of data that enhance understanding and enhance learning outcomes. The infusion of technology is particularly prominent in subjects like computer science.

EA application within computer science at the tertiary level supports the assessment of coding assignments. It is also comprised of the efficacious attributes which involve the automated analysis of coding proficiency (Barra et al., 2020). This topic is a mix of the foundations of computer science with methods for assessing student learning in computer programming and algorithm design (Kanika and Chakraborty, 2020). This kind of EA goes past the uncomplicated conversion of customary assessment techniques into electronic format. It utilizes the unique capabilities of computing devices to analyze code attributes, including structure, logic, and output (Paiva et al., 2022). Within this learning space, EA is not just a grading system, but also a vehicle for providing timely and comprehensive feedback on students' problem-solving abilities and coding abilities.

The aim of this research is to examine the success factors of EA, more particularly its impact on computer science courses in universities from the perspective of teachers. It talks about the pedagogical and practical aspects of EA for programming assignments, and how it can reduce the time and effort required by teachers in creating these assignments while maximizing learning outcomes. This study provides a unique perspective on the intersection of technology and education, highlighting the challenges and opportunities associated with EA in higher education computer science.

## Literature review

### The term “E-assessment” in higher education (HE)

Though e-learning has been a component of higher education for a while, e-assessment is a new concept (Buzzetto-More and Alade,

2006). E-assessment is the use of electronic technology to assess learners throughout the instructional strategies (Hettiarachchi et al., 2015). In addition, Singh & De Villiers state that EA is a way of evaluating learners' skills and knowledge while utilizing information and communication technology (ICT) in all the assessment steps, including design, implementation, recording responses, and feedback. EA encompasses a vast multitude of assessment methodologies. Nevertheless, the analysis of complex coding assignments might present certain limitations (Conrad and Openo, 2018).

There are many definitions of the EA but for the purpose of this research, we will focus on the following definition. EA is an aspect of higher education that is constantly evolving and developing; it is a major development in the ways learning is assessed through the inclusion of electronic technologies in the assessment process. In this regard, a holistic approach is involved whereby ICT is applied from the very beginning in designing an assessment to giving feedback rather than using ICT only in designing and delivering tests. Moreover, while EA is a very broad area that contains many different tools, the scope of this research tries to represent the concept of EA of coding assignments.

## Qualities of assessment

Assessment in higher education can really be effective only when done in tandem with clearly defined learning objectives. That way, the methods adopted could directly measure the skills and knowledge that they are intended to measure (Fernandes et al., 2023). Assessment will actually be effective only when it is valid assessment must measure what it purports to measure and adequately covers the subject matter so as to elicit from the students the required knowledge or skills (Adamu, 2023). These studies confirm that the complexity of validity assurance is one of the major factors contributing to a lack of confidence in results and feedback produced through EA tools.

Apart from this, there has to be reliability, that is, assessments should produce consistent results over time and across diverse groups of students, therefore offering dependability in the measure of student learning (Rao and Banerjee, 2023). This actually would mean implications of EA for instructors' teaching practices and following a holistic approach in the development of expertise in assessment and promoting instructor experience, which is seen as a big challenge arising in higher education regarding the importance of assessment.

On the other hand, effective assessment plays a crucial role in supporting student learning and development by providing timely, specific, and actionable feedback that helps students recognize their strengths and areas for improvement (Subheesh, 2023; Rothwell, 2023). By integrating these varied assessment approaches, instructors can better assist all students in reaching their learning objectives, making the assessment process an essential component of the educational journey.

## Feedback in E-assessment

Feedback in EA is the most important element in higher education; it allows for timely diagnoses and changing of learning strategies aimed at achieving better outcomes. The traditional methods are not offering any immediate feedback, but EA provides this and

enables students to point out strengths and weaknesses on their own. This immediacy reinforces the academic experience and encourages the learning process. There again, research underlined how the ease of use within assessment-enhanced technology is an important adoption determinant of these tools, hence proving worthwhile to assure immediate action in order to inform the students and instructors (Apostolellis et al., 2023; Heil and Ifenthaler, 2023; Dimitrijević and Devedžić, 2021).

## Impact on learning outcomes feedback

It is one of the strongest drivers for improving learning (Torrisi-Steele et al., 2023). E-assessments provide better, timely, and constructive feedback to help students make focused improvements (Rajaram, 2023). Gaps in the use of feedback persist because of a lack of frameworks and instructor training that limits its effectiveness. While automated feedback functions well within summative assessments, these generally lack the depth and richness found within human-led formative assessments (Evans, 2013; Goldin et al., 2017).

## Challenges and integration

The integration of EA feedback into teaching strategies is in itself a challenge. Instructors need training and frameworks for using feedback in a balance with pedagogical methods. Full realization of EA feedback for student learning and outcomes development requires professional development and an evolution of teaching practices. These gaps, at the institutional levels, have to be addressed to maximize the potential of e-assessments in higher education (Kaya-Capocci et al., 2022).

Based on the literature reviewed, as well as the results of the interviews and surveys conducted in this study, several core principles emerged as significant in shaping the effective adoption of e-assessment (EA) tools in coding education: authenticity, reliability, validity, flexibility, accessibility, practicality, and scalability. While these concepts are grounded in existing research, their specific relevance and interconnection became more evident through the empirical data. A preliminary overview of these factors in the provided framework. These findings will be discussed in greater detail in the Discussion section.

## Theoretical framework

### Theoretical underpinning

The theoretical foundation of this study is different from that of the current research, although they also have some similarities. The foundation of this study is the Technology Acceptance Model (TAM), whereas current research has heavily employed the Unified Theory of Acceptance and Use of Technology (UTAUT). Both theories emphasize the determinants of technology acceptance and utilization. Nevertheless, the primary distinction lies in their emphasis. UTAUT encompasses a more extensive range of determinants such as performance expectancy, effort expectancy, social influence, and facilitating conditions, all of which lead to technology adoption in an organizational setting (Terzis and Economides, 2011). TAM, on the

other hand, concentrates on perceived usefulness and perceived ease of use as the primary determinants of technology acceptance (Ramdhani, 2009). This focused approach enables a more compact examination of how these determinants have a direct influence on the acceptance of EA tools in universities. By applying TAM, this study seeks to provide clear and usable information on the particular features of technology that influence user acceptance, enabling a focused approach to complement the more overall view of UTAUT.

## Technology acceptance model (TAM)

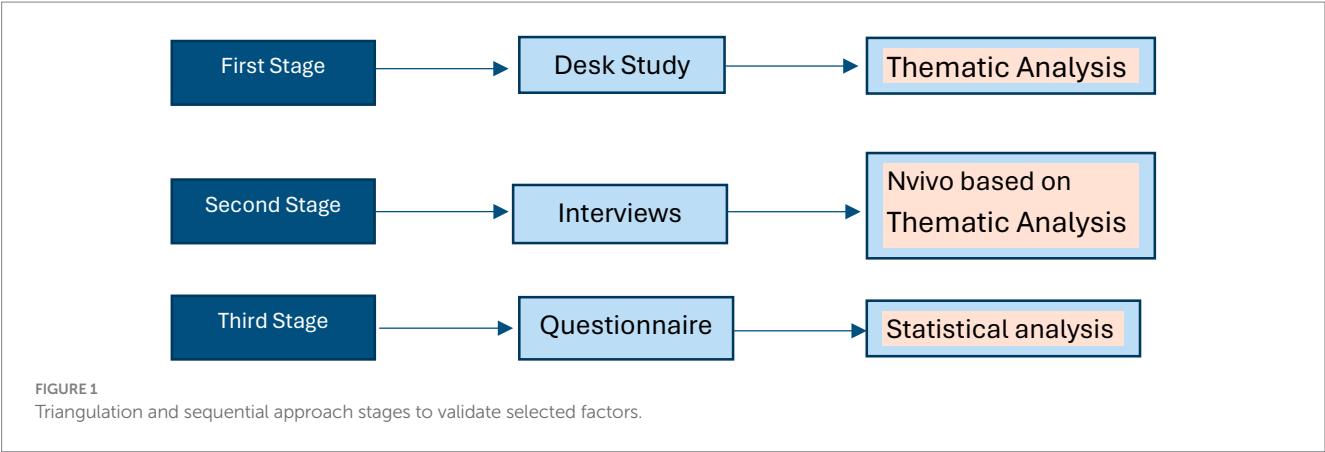
### Definition and history

Technology Acceptance Model (TAM) is one of the most widely applied theories measuring users' intention to use information and communication technology (ICT) by being influenced by the Theory of Reasoned Action (TRA) developed by Fishbein and Ajzen (Ramdhani, 2009)(Mekhiel, 2016). TAM proposes that Perceived Ease of Use (PEOU) and Perceived Usefulness (PU) affect the attitude of users, which, in turn, affects Behavioral Intention (BI) along with actual usage behavior (Malatji et al., 2020). TAM has been modified over time by adding or deleting variables and introducing moderators or mediators to improve its applicability to emerging technologies (Stubenrauch and Lithgow, 2019). The model has been applied to assess the adoption and acceptance of various innovations, with identified limitations being issues with measuring behavior accurately and criticism for not accounting for such factors as costs and structural constraints in shaping adoption decisions (Fatmawati, 2015). Despite being put under criticisms for its outdatedness, a bibliometric review of 2,399 articles published between 2010 and 2020 attests to the ongoing usefulness and adaptability of TAM in various domains such as e-commerce, banking, education, and healthcare (Al-Emran and Granić, 2021). Therefore, a grasp of the genesis, utilization, extensions, limitations, and critiques of TAM is indispensable in understanding technology adoption and successful integration of new technologies such as EA that is the subject of this research.

### TAM in this research

As discussed in the section above, this research has been formulated within the theoretical framework of the Technology Acceptance Model (TAM). The different components of TAM were redefined for the context of this study. While the majority of the components are analogous to those of other adaptations of this model, there are others that have never been investigated in the same manner previously in related research. For instance, Venkatesh and Davis's (2000) perceived usefulness and perceived ease of use constructs are identical to those of this study, as indicated in Figure 1. Nonetheless, the external variables and their influence on behavioral intention and actual system usage differ from those in the model proposed by (Venkatesh and Davis, 2000).

The Technology Acceptable Model (TAM) is an important role in this research since it accurately predicts and explains technology adoption and usage by the users, specifically in educational settings. The TAM focuses on the perceived usefulness as well as ease of use, whereas it also demonstrates a concise but explanatory foundation designed for instructional use measurement of the EA tools. As information technologies become ever more pervasive in higher education, it's essential to ascertain the factors enabling or hindering the adoption of EA systems. This research aims to establish the success



factors of EA tools that influence their acceptance through TAM. In this way, it provides practical implications for improving the design and implementation of the tools. This focus not only responds to the increasing need for effective digital assessment methods, but also helps toward the incorporation of technology into education, causing EA tools to become beneficial and functional for all stakeholders.

Research method

In this research, a multiphase mixed sequential techniques approach was used. The first stage entailed conducting desk study, the second stage entailed conducting interviews with computer science instructors at university of Southampton and Saudi Arabia universities. This stage comprised both open- and closed-ended questions. The open-ended questions were helpful in identifying current use of EA, reasons behind the response of computer science instructors to closed-ended questions, challenges and opportunities in using EA when assessing programming courses, and providing additional factors other than those in selected ones. The third stage involved sending an online questionnaire with closed-ended questions that were distributed to all computer science instructors at the university of Southampton and Saudi universities. This was aimed at confirming the factors resulted from the literature and participants’ interviews (Figure 2).

Data collection method

Participants

As stated in the introduction, the main objective of this research was to identify success factors that impact on using EA at higher education computer science from the perspectives on CS instructors. Participants with varying levels of expertise were examined across different stages. These participants were categorized into experienced instructors (lecturers, associate professor), and experts. Table 1 provides the participants number in each of the stages include in this research.

Data collection tools

Interview

In order to accomplish substantial outcomes, it is necessary to interview an appropriate number of experts. It is crucial to determine

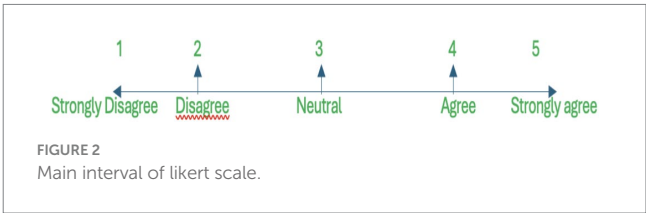


TABLE 1 Stages of this research.

Stage	Description	Number of participants
Desk study	Analysis of the literature review	0
Experienced instructors	Interview one to one with lecturers, associate professor from UK universities	8
Experts	Interview one to one with lecturers, associate professor from Saudi Arabia universities	6
Online Survey	Sending an online survey to CS instructors from UK and Saudi Arabia universities.	44

the minimum sample size that will yield reliable and valid results at this juncture (Banerjee et al., 2009). The number of experts that should be interviewed for the study of content validity is not a fixed answer; however, the majority of experts agree on a panel of three to twenty (Grant and Davis, 1997)(Lynn, 1986). Marshall et al. (2013) suggest that the saturating approach be implemented (Marshall et al., 2013). In qualitative research, data analysis is the process of collecting data until a point of redundancy is attained, at which point no new data is acquired (Bowen, 2008). The researcher in this study achieved data collection duplication. Consequently, 16 lecturers, associate professors, and experts from Southampton University and Saudi Arabia universities were interviewed.

The participants were invited to participate in this investigation via email. The email contained a brief summary of the research, its objectives, and a request for the instructors to arrange a time for the interview. The Team app was employed to conduct the interviews. Upon commencing each interview, participants were requested to examine the participant information page. Following this, they were posed with approximately 12 closed-ended queries. The framework was refined and any additional factors that influence the use of EA were identified in the concluding open-ended segment. All sessions were recorded with the participants’ consent, and each interview lasted approximately 30 to 45 min.



The interviews were conducted with the aim of establishing the potential benefits and constraints of employing EA within programming modules at the University of Southampton and Saudi Arabian universities, as well as uncovering factors that were not previously considered. The research used an interview framework that was semi-structured, incorporating open-ended and closed-ended questions. These questions were reviewed by three specialists who provided constructive feedback with a view to facilitating their clarification and simplification. Although no significant changes were proposed, they advised that certain ambiguous points be clarified.

All interviews in this study were conducted as one-time sessions with each of the 16 participants. Follow-up interviews were not conducted, as the initial sessions offered sufficient depth and adequately addressed the key areas of interest outlined in the study. The data obtained from these individual interviews contributed significantly to the identification of recurring themes and informed the development of the success factors framework.

The data saturation had been reached, as recurring themes began to emerge and no substantially new insights were introduced. However, arranging and conducting these interviews presented challenges, particularly in coordinating schedules and obtaining participants' agreement to take part. Despite these logistical difficulties, a diverse range of perspectives was gathered, and the consistency of responses suggested that the key themes were sufficiently explored to inform the development of the success factors framework.

### Experienced instructors and experts

In this stage, a total of 16 participants took part, all of them instructors in programming, and comprising of lecturers, associate professors and experts. To obtain significant results, it was imperative to conduct interviews with a suitable number of experts. The choice of these participants was grounded on their proficiency and pertinence to the primary emphasis of the research, guaranteeing that the gathered data would be comprehensive (Table 2).

### Questionnaire

Foddy and Foddy (1993) asserted that shorter scales, i.e., five-point scales, are more appropriate when a clear decision is needed (Foddy and Foddy, 1993). Hence, in order to determine the selected influential factors of EA adoption, a five-point Likert scale (strongly disagree to strongly agree) was utilized in this study for close-ended questions (Likert, 1932). The reasoning underlying such a practice was to verify the validity of comments made by experts and to provide them with response diversity options. The factors that influence the use and acceptance of EA were identified by the mean intervals of the Likert Scale. The intervals of the scale were 1 to 5. (Refer to Figure 2). It was deemed significant if the mean factor was 3 and above; otherwise, it was not included.

## Data analysis (methods)

### Thematic analysis (interview data)

#### Thematic analysis

Thematic analysis is a popular method for analysing qualitative data that assists researchers in identifying, analysing, and

TABLE 2 Participants' career.

Participants	Career
P1	Lecturer in Computer Science, Electronics and Computer Science (staff) at a British university.
P2	Senior Teaching Fellow, Electronics and Computer Science (staff) at a British university.
P3	Principal Teaching Fellow, Higher Education Practice (staff) at a British university.
P4	Associate Professor, Electronics and Computer Science (staff) at a British university.
P5	Professor, Electronics and Computer Science (staff) at a British university.
P6	Senior Teaching Fellow, Electronics and Computer Science (staff) at a British university.
P7	Lecturer, Electronics and Computer Science (staff) at a British university.
P8	Associate Professor, Electronics and Computer Science (staff) at a British university.
P9	Assistant Professor, Computer Science at a Saudi university.
P10	Lecturer, Computer Science at a Saudi university.
P11	Assistant Professor, Computer Science at a Saudi university.
P12	Lecturer, Computer Science at a Saudi university.
P13	Lecturer, Computer Science at a Saudi university.
P15	Senior Teaching Fellow, Computer Science at a Saudi university.
P16	Associate Professor, Computer Science at a Saudi university.

interpreting patterns or themes within a dataset. It permits a thorough investigation of the data's underlying meaning and patterns (Castleberry and Nolen, 2018). There are several ways to analyze qualitative data generated by interviews, but thematic analysis was considered to be the most appropriate for the current. Thematic analysis, content analysis, narrative analysis, discourse analysis, grounded theory, framework analysis and phenomenological analysis, all of which involve identifying patterns and themes within the transcribed interview responses by systematically reviewing the data to understand the meaning behind the participants' words and experience (Abell and Myers, 2008).

However, thematic analysis was chosen to identify key patterns in teachers' experiences with EA. This method, as articulated by Braun and Clarke (2006), offers adaptability in coding while preserving a methodical approach to data processing (Braun and Clarke, 2006). The practical focus on real-world applications directed the coding process, ensuring that subjects were relevant to the development of a framework of success factors of e-assessment.

The purpose of utilising thematic analysis in this study on EA is to produce comprehensive, context-specific insights that enable a thorough examination of digital assessment instruments in HE. Due to the intricate and dynamic characteristics of EA, a descriptive and adaptable methodology is crucial to encompass the varied experiences of instructors in different institutional settings. Thematic analysis is well-suited for this objective, since it facilitates the identification of

patterns and themes within qualitative data, providing a comprehensive insight of the perception, adoption and utilization of digital assessment tools.

NVivo is specifically designed for qualitative data analysis, including thematic analysis. It offers a variety of tools and features to assist researchers with the organization, coding and analysis of qualitative data (Sotiriadou et al., 2014). Typically, researchers follow several stages to conduct a thematic analysis with NVivo (Dhakal, 2022):

1. Familiarization: The researcher familiarises themselves with the data by perusing or listening to the qualitative material. This phase requires repetitive exposure to the data in order to comprehend its content.
2. Coding: NVivo permits researchers to designate their data by labelling or tagging particular sections or segments. Codes can be used to identify key concepts, ideas or patterns within the data. Either a coding structure or an extant framework can serve as a guide for the process.
3. Theme development: The researcher then identifies patterns, relationships or themes that emerge from the coded data segments. Themes are comprehensive concepts that encapsulate the essence of the data and yield insightful conclusions. NVivo facilitates the analysis process by enabling researchers to organise and combine codes into themes.
4. Reviewing and refining themes: Examining their coherence, relevance, and relationship to the data, researchers examine and refine the themes that have been identified. NVivo facilitates the manipulation and reorganization of themes and codes, allowing for the iterative refinement of the analysis.
5. Interpretation and reporting: After developing and refining the themes, the researcher interprets and analyses the data in light of these themes. From the qualitative data, this phase involves drawing connections, identifying patterns, and generating insights. NVivo facilitates the creation of reports, summaries, visualisations, and other outputs to effectively communicate findings (Dhakal, 2022).

I therefore used NVivo software for the qualitative data analysis, and conducted a comprehensive analysis of the interview data I had collected. Initially, I meticulously transcribed the interviews, ensuring the accuracy of every detail. After preparing the transcripts, I imported them into NVivo and commenced a methodical process of coding. Using an iterative and inductive methodology, I analyzed the data, identified patterns, and allocated appropriate identifiers to capture the essence of each theme. The intuitive interface and tools of NVivo facilitated the organization and interconnection of these themes, resulting in a coherent narrative encompassing the prevalent challenges, systems, tools and categories of EA in this project.

## Statistical analysis (survey data)

In this research, there are four types of statistical tests have been used which are Likert scale, T-test, ANOVA, and Regression analysis. The analysis of Likert scale data is linked with an ordinal response code in numerical terms typically of the 1 (strongly disagree) to 5 (strongly agree) kind of response, with descriptive statistics like means and standard deviations in order to determine central tendency and

dispersion (Hutchinson and Chyung, 2023). Besides, one of the most significant applications of t-tests for survey data analysis is a comparison of the means of two groups with a view to determining statistical differences, especially when analyzing the impact of binary independent variables on continuous dependent variables (O'Brien and Muller, 1993). Statistical tests such as the *t*-test for one mean or the difference between two group means provide a powerful method for a researcher to determine if differences are random or significant (Ewens and Brumberg, 2023).

Furthermore, ANOVA is a very powerful statistical technique that can be used in the comparison of three or more group means with the aim of determining whether there exists any significant group difference (ANOVA, 2008; Mage and Marini, 2023). The test is simply an extension of the *t*-test since it allows comparison of several groups simultaneously. Thus, ANOVA is extremely suitable for independent variables with multiple categories at more than a single level in survey research, such as types of educational levels or age groups (Das et al., 2022).

Finally, regression analysis in survey data is utilized in understanding how dependent and independent variables are associated with one another by predicting and explaining data patterns (Mostoufi and Constantinides, 2023). Multiple regression analysis enables a researcher to understand the combined effects of several predictors and the relative importance of these variables by considering several predictors simultaneously (Cleophas and Zwinderman, 2023).

## Methodological limitations

It was difficult to plan and conduct interviews in both Saudi Arabia and the UK in this study. Coordinating interviews or consent across these areas consumed significant time and effort, as institutions had to give their blessing and people had to be found to participate. In addition, collecting participants' emails from both contexts took much longer than anticipated. This hampered the onset of data collection.

In addition, a low response rate in the survey was a further constraint. Despite trying to reach an audience more widely, it was expected that the number of respondents who filled out the questionnaire would be more than it actually turned out to be. This was despite having sent out 332 emails to CS instructors which were obtained from university websites, and sending reminders several times. I also used my social media network to try and find more participants.

Adequate sample sizes are essential to ensure the reliability and validity of statistical tests. A minimum sample size of 50 plus eight times the number of predictors is recommended by Green (1991) for regression analysis, and 104 plus the number of predictors is recommended for evaluating individual predictors (Green, 1991). The significance of taking the particular analytic approach into account when calculating sample size is highlighted by Brooks and Johanson's (2011) demonstration that a suitable sample size for the omnibus test does not always give sufficient statistical power for *post hoc* multiple comparisons in the context of ANOVA (Brooks and Johanson, 2011).

One of the main limitations of this study is the relatively small sample size in both the survey (44 responses), which may restrict the generalizability of the findings beyond the specific institutions

involved. Although efforts were made to include participants from different universities, the sample may not fully represent the broader diversity of experiences with e-assessment across higher education.

Additionally, the modest number of responses is below the recommended threshold for certain statistical techniques like ANOVA and regression, which could have limited the detection of statistically significant results. However, despite these constraints, the study offers important insights by highlighting emerging patterns and perceptions among academic staff regarding the use of e-assessment. It provides a valuable starting point for understanding the factors influencing adoption and implementation, especially within under-researched contexts. The mixed-methods approach especially the qualitative data allowed for a deeper exploration of participants' experiences, adding meaningful depth to the findings. Future research should build on this foundation by involving larger, more diverse samples across institutions and adopting longitudinal designs to capture how practices and attitudes evolve over time, thereby enhancing the generalizability and impact of the results.

Although  $p$ -values were reported accurately throughout the analysis, it is important to interpret them within the context of this study. Some factors did not reach statistical significance; however, this should not be taken as evidence of their irrelevance. Due to the relatively small sample size, the statistical power of the analysis was limited, which may have hindered the detection of meaningful effects. As such, non-significant results should be interpreted with caution. Nevertheless, the findings still offer valuable insights particularly when considered alongside the qualitative data and may highlight emerging trends worth exploring. It is recommended that future research with larger sample sizes further investigate these factors to confirm or challenge the patterns identified in this study.

I acknowledge that the proposed framework has not yet been empirically tested in real educational settings. However, all prior steps leading to the development of the framework including the research design, data collection tools, and analysis were carefully validated to ensure their reliability. The framework itself was grounded in both the existing literature and the findings from surveys and interviews. While its practical implementation has not yet been formally evaluated, this represents an important direction for future research to assess its effectiveness and applicability in diverse educational contexts.

In this study, the ANOVA test was used to explore differences in e-assessment usage across varying levels of teaching experience. The assumptions underlying ANOVA such as normality and homogeneity of variances were not formally tested, given the exploratory nature of the analysis. The primary objective was to identify general trends and patterns rather than to draw definitive statistical conclusions. Nevertheless, the absence of assumption testing is acknowledged as a limitation, as these assumptions are essential for ensuring the validity and reliability of ANOVA results. Future research should address this by incorporating assumption checks, such as the Shapiro Wilk test for normality and Levene's test for homogeneity of variances, to strengthen the robustness and interpretability of the findings.

This represents a limitation of the current study. Future research should aim to validate the framework through pilot implementations or case studies within higher education settings. Doing so would help assess its effectiveness, identify potential improvements, and ensure its relevance and adaptability across different institutional contexts.

## Ethical considerations

Ethical approval for this research was obtained via ERGO (Ethics and Research Governance Online) at the University of Southampton prior to commencing the study. I submitted an application detailing the research objectives, methods and ethical issues. The framework guaranteed that my research conformed to ethical standards, safeguarded participants' rights, and met institutional and legal obligations.

The University Ethics Committee evaluated and approved my application, granting authorisation for data collection (research ethics number 81152). ERGO enabled me to make any modifications, and guarantee complete adherence to ethical norms. This systematic approach guaranteed transparency and responsibility in my research, bolstering ethical integrity throughout the project. All participants were apprised of the study's objectives and provided written consent. Confidentiality was preserved through the anonymization of replies and secure arrangements for storing data.

To prevent ethical concerns, I instituted rigorous protocols to guarantee participant anonymity and data confidentiality. All replies were completely anonymized, with no identifying information gathered or associated with individuals. Data was securely held on password-protected computers, available solely to the researcher, in accordance with the University of Southampton's ethical norms. Participants were clearly informed of their right to withdraw at any point without the necessity of giving a reason. Furthermore, they were notified that the data would be obliterated at the study's completion.

## Results

This research provides the results of a mixed-method research design triangulating both quantitative and qualitative approaches for gaining an understanding of the impacts and adoption of EA tools in computer science instruction among instructors in the UK and Saudi Arabia. Qualitative inputs for the present study are from these interviews conducted with instructors, while quantitative data are from the survey, to explore different facets of using EA tools for coding assessment in higher education.

In addition, this research provides the findings around the main aim identified earlier. Data have been classified into themes representing benefits and challenges of using EA tools, whereby this classification allows an in-depth analysis of the experiences of the instructors. Direct quotes from participants contributed to the nuanced understanding of the instructors' perceptions. Through combined analysis, it aims at bringing out the play of practical and pedagogical factors that might go into shaping the adoption of EA tools by instructors of computer science within various education contexts.

## Success factors and challenges for E-assessment

### Pedagogical factors: authenticity, reliability, validity

#### Authenticity

Participants had concerns about the authenticity of EA tools and the ability of these tools to measure learning outcomes. Participant 5 said.

*“Assessment methods may not effectively support learning objectives, leading to superficial knowledge” [P5].*

Despite these concerns, the survey data reveals that most respondents have a positive view of the authenticity of their EA tools. More precisely, 44.19% agreed while 41.86% strongly agreed that their tools are mapped to the nature of the curriculum, with 13.95% being neutral toward this. These findings suggest that there is an overall belief that EA tools can genuinely measure learning, though the qualitative insights of interviews bring out a very strong dependence on context and conditions. The view taken over the tools was still a good one, but there are notable exceptions of superficial assessment where the inadequacy of set-up can undermine their perceived authenticity (Figure 3).

### Reliability

Another significant concern is the reliability, especially about consistency and accuracy in feedback from the EA tools. Participant 3 also pointed out potential challenges regarding maintaining assessments that are reliable.

*“Setting up the exam for competency assessment can be challenging due to the need for technical proficiency” [P3].*

Such concerns are complemented by the quantitative survey data, which indicates that while a substantial portion of respondents 48.84% agreed that the tools give consistent scores when the same code is assessed multiple times, a significant minority expressed reservations: 25.58% neither agreed nor disagreed and 20.93% strongly agreed, indicating some level of confidence in the reliability of the tool. For instance, 4.65% of the respondents ticked off “Other,” suggesting in some way that specific nuances or conditions are affecting the reliability for which the overall question did not account in its standardized answer options. The results proposed here indicate that though there is a general agreement over the reliability of tools, basic concerns remain on consistency and accuracy, coupled with the dependence on proper setup and manual oversight to achieve reliable outcomes (Figure 4).

For instance, 4.65% of the respondents selected the “Other” option when asked about factors influencing reliability. This suggests that there may be specific nuances or contextual factors affecting perceptions of reliability that were not fully captured by the predefined response options. Although the overall results show general agreement regarding the reliability of e-assessment tools, these “Other” responses indicate that underlying concerns remain particularly related to consistency, accuracy, and the dependence on correct setup and manual oversight. These elements may not have been explicitly addressed in the survey but appear to play a role in shaping user confidence in the system’s reliability.

### Validity

Concerns about validity were also raised regarding EA tools, particularly concerning their ability to measure what they are designed to measure. Participant 1 commented.

*“There are doubts about the effectiveness of computer-aided assessments in measuring deep learning outcomes” [P1].*

Data from the survey subscribes to the apprehensions, where a significant percentage of respondents doubted over the validity of their EA tools. In fact, 46.51 percent of all respondents do not agree that their EA tool provides quality feedback and 30.23% neither agreed nor disagreed. Of all the participants, it could be gathered that only a few had a positive view of the tool at 16.28 and 6.98% that agreed and strongly agreed to have valid feedback produced by it. These findings thus show that while a substantial number of users have faith in the appropriateness of their EA tools, there is a large amount of doubt, particularly in the respect that these tools can really measure deep-learning outcomes or give feedback which reasonably reflects student understanding. In the following section, there will be an overview of the comparison between feedback quality and time/effort-saving ratings among frequent and less frequent users EA tools (Figure 5).

T-test of the results showed no significant difference between frequent and less frequent users in their perceived feedback quality and time/effort saved. This is represented by the very high  $p$ -values:

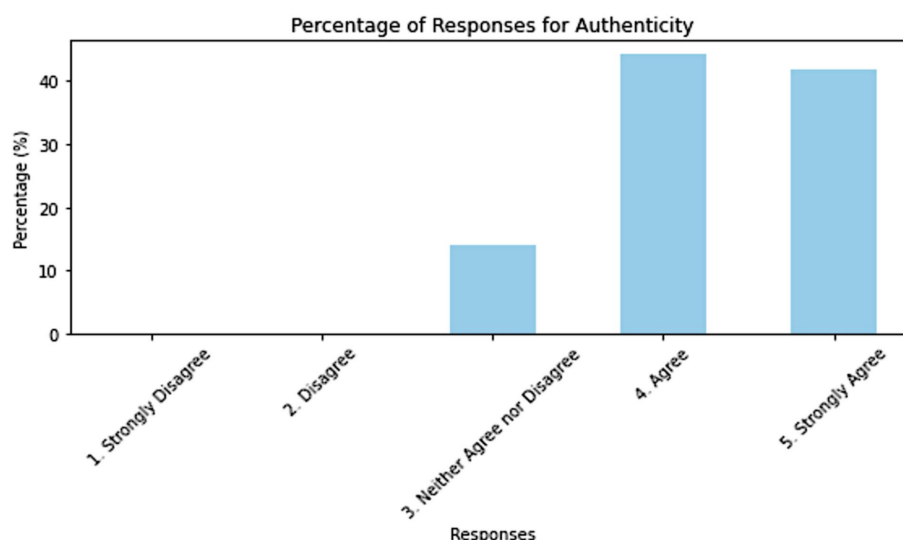


FIGURE 3  
Percentage of responses for authenticity.



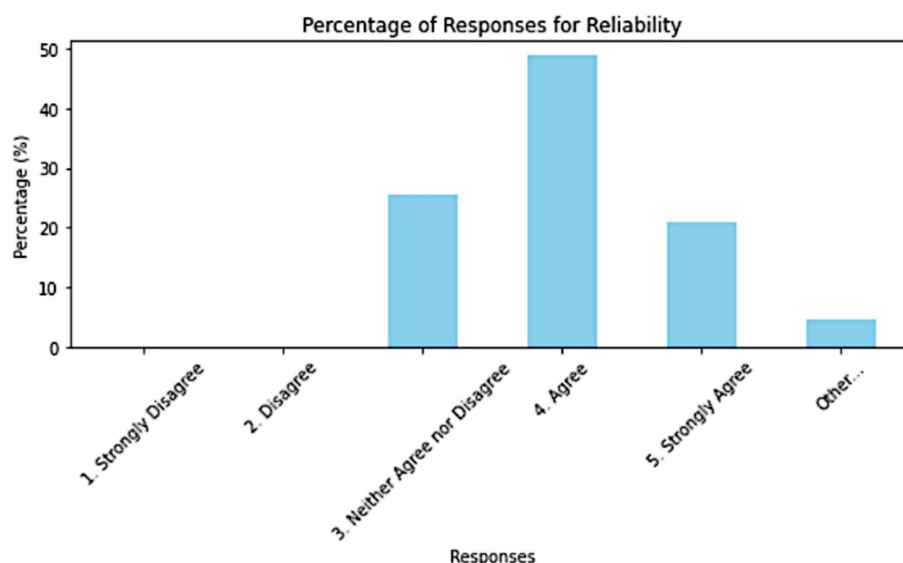


FIGURE 4  
Percentage of responses for reliability.

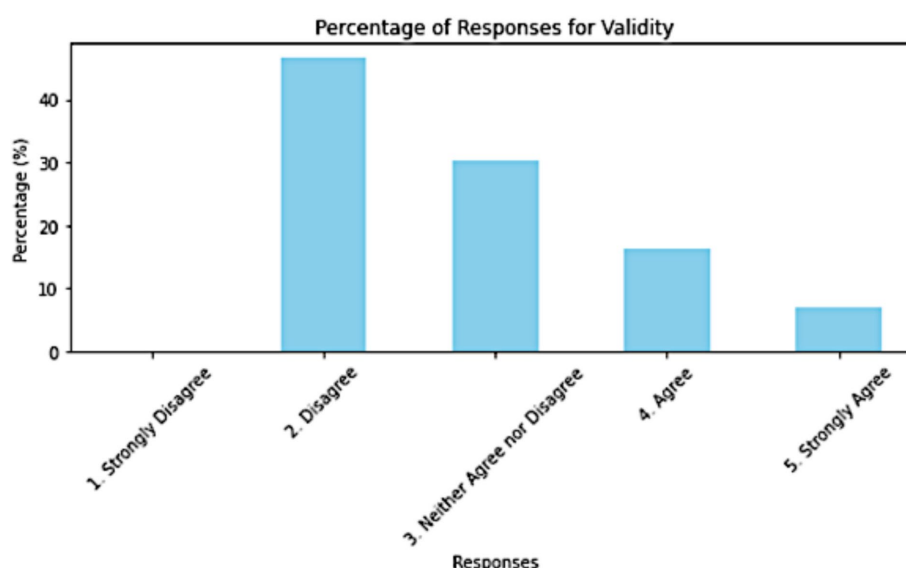


FIGURE 5  
Percentage of responses of validity.

0.9 with respect to feedback quality, where for time/effort saving it stands at 0.9 and are well above the common threshold level of 0.05. A high p-value means that the difference between the groups is probably due to random chance and not a true difference of opinion between the two groups (Figure 6).

## Practical factors: flexibility, accessibility, practicality, and scalability

### Flexibility

Flexibility was one of the important features for a successful EA tool. Participant 9 from the UK stated.

*“I’ve created my own e-assessment tool. Therefore, it’s extremely flexible” [P9].*

This is also demonstrated by the same findings in the quantitative graph, where most of the respondents (44.19%) strongly agreed and (27.91%) agreed that their EA tools were flexible and could be taken anywhere and anytime. However, it is also clear from the graph that a substantial number of respondents (20.93%) had concerns about flexibility, sharing exact limitations regarding web setup and institutional policies. This means that, although flexibility is positively rated overall, there must be some critical aspects that EA tools need to adapt to fully satisfy the

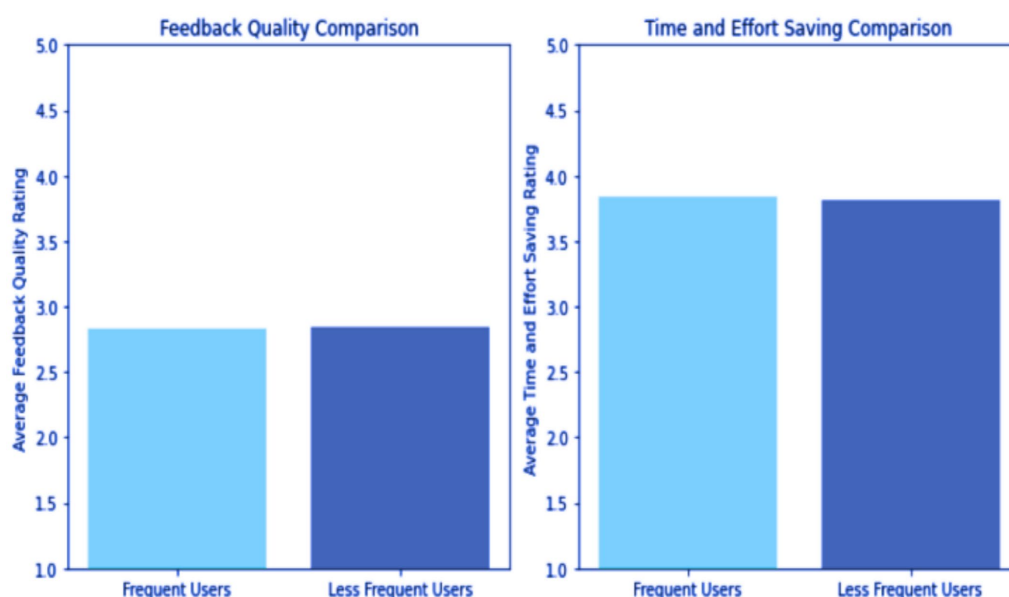


FIGURE 6  
Comparison between feedback quality and time/effort-saving ratings.

diversities of instructors. This decreasing trend is evident from “Strongly Agree” to “Security policy of the uni” in both frequency and percentage, indicating general satisfaction along with important exceptions where flexibility could stand to be improved (Figure 7).

### Accessibility

Accessibility problems significantly impacted the experience of instructors, particularly in the proper utilization of EA tools. Participant 4 from the UK stated,

*“The university’s assessment system is accessible to all students, regardless of technological or language background, and how the system is designed to support students with disabilities” [P4].*

The qualitative findings illustrate the illustrative perception of the EA tools in the survey data. More than half of the respondents, 55.81%, agreed that their EA tool has an easy user interface, where 9.30% strongly agreed; this paints a general positive perception of the ease of use of the tool. However, a substantial number of them, 27.91%, neither agreed nor disagreed, which implies uncertainty or mixed experiences. Besides, 6.98% of the users disagreed; this is really strong argument because for every user for whom the tools are easy to use, there is a significant group of users faced with difficulty in navigating or using the tools. In the next section, there will be ANOVA test to compare between the mean tool accessibility rating by teaching experience (Figure 8).

ANOVA test results show an F-statistic of 0.685 and a *p*-value of 0.606. The F-statistic indicates the ratio of the variance between different groups (in this case, groups defined by teaching experience) to the variance within the groups. A *p*-value of 0.606 is much higher than the common significance threshold of 0.05, suggesting that the differences in mean tool accessibility ratings across the different levels of teaching experience are not statistically significant. Consequently, we are unable to refute the null hypothesis, which

implies that there is no evidence to suggest that teaching experience has a substantial effect on perceptions of tool accessibility among the respondents.

### Practicality

A number of participants continually mentioned the ease-of-use practicality of the EA tools. Participant 4 pointed out that the tools could have potential time-saving advantages.

*“Marking on papers is error-prone and time-consuming, but the new assessment platform streamlines the process” [P4].*

These concerns are further illuminated by the quantitative data from the survey. Respondents were generally positive about whether EA tools were practical in terms of saving time and effort 41.86% agreed and 27.91% strongly agreed. On the other hand, 16.28% disagreed and 13.95% neither agreed nor disagreed. Still, a significant minority has problems in this respect. When asked if EA tools are adaptable with their work, 44.19% agreed and 30.23% strongly agreed, which shows that many of the respondents find these tools adaptable to their practice. However, 25.58% neither agreed nor disagreed, whereby some of the respondents are not sure or have mixed experiences regarding the practicality of these tools in their specific contexts (Figure 9).

### Scalability

Scalability was recognized as an essential factor for effectively using EA tools, especially in managing large classes and maintaining consistency in grading and feedback. Participant 4 discussed the difficulties of scaling assessments for larger classes, mentioning the,

*“Challenges of scaling up the assessment for a large class, including the need for consistency in marking and the number of demonstrators required” [P4].*

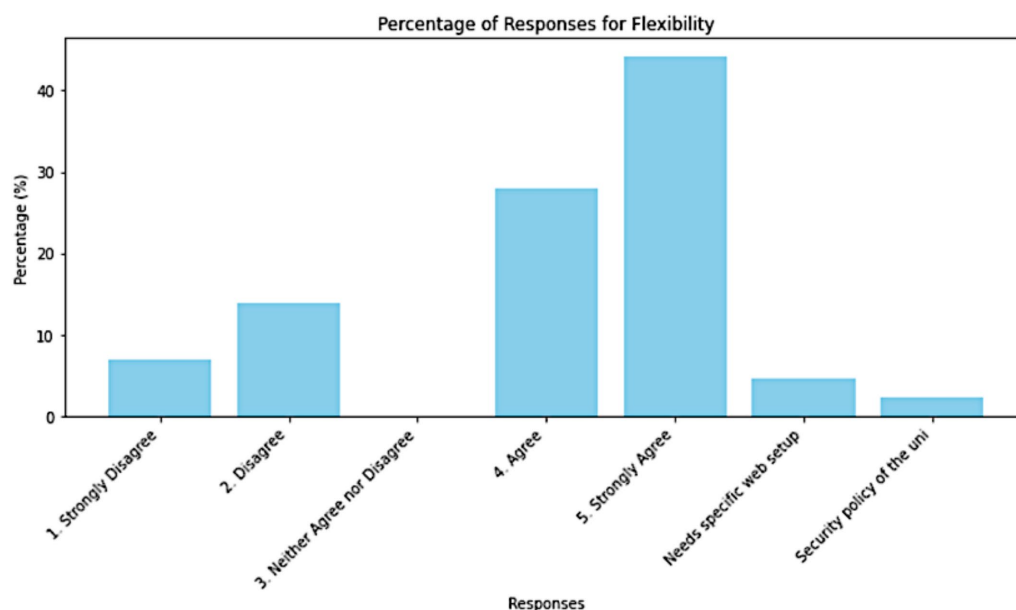


FIGURE 7  
Percentage of responses for flexibility.

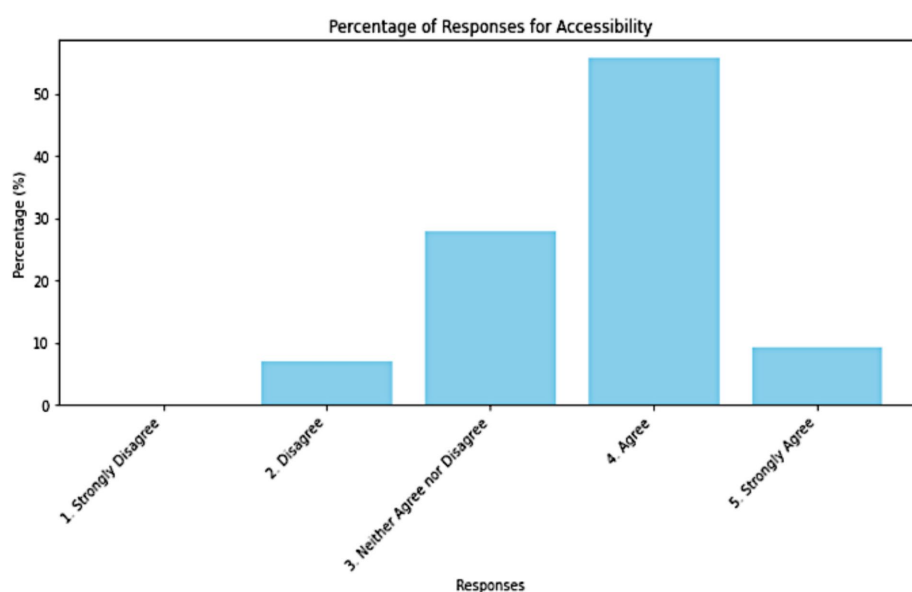


FIGURE 8  
Percentage of responses for accessibility.

The quantitative survey data support the mixed perceptions of scalability by the users. This is clear in the statistics, with 41.86% agreeing and 39.53% strongly agreeing that their EA tools could handle many students simultaneously; this seems to reflect a generally good perception of scalability. However, 13.95% strongly disagreed, and 4.65% disagreed, highlighting that there are still significant concerns about the ability of these tools to scale effectively, especially in contexts where logistical or technical challenges are prominent (Figure 10).

## Factors affecting adoption of E-assessment

### Pedagogical and practical influences on adoption

Personalized feedback was another key feature that several participants alluded to within their reflections on the adoption of EA tools. For instance, Participant 8 said,

*“Assessment alone is not adequate for student feedback, must be part of a hybrid strategy” [P8].*

Such concerns on the adequacy of personalized feedback from EA tools are indeed reflected within the quantitative survey data. Data indicated that 30.23% strongly disagreed that automated feedback provided by EA tools is personalized for each student, while 25.58% just disagreed. This serves as an indication that over half of the respondents have negative perceptions of personalization capabilities in these tools. Meanwhile, 25.58% agreed and 18.60% neither agreed nor disagreed, which therefore reveals that a very small minority finds the tools somewhat adequate or has mixed experiences about personalized feedback. In the next section, the results of logistic regression, which investigate the influence of feedback quality and curriculum alignment on the adoption of EA tools will be presented (Figure 11).

The logistic regression results examine whether feedback quality and curriculum alignment, influence the adoption of EA tools. The coefficients for both “Feedback Quality” and “Curriculum Alignment” are very close to zero (0.0231 and 0.03, respectively), and their associated  $p$ -values (0.9 and 0.9) are significantly higher than the common significance level of 0.05. This indicates that neither feedback quality nor curriculum alignment has a statistically significant effect on the likelihood of adopting EA tools. The model’s pseudo-R-squared value is close to zero, suggesting that the model does not explain much of the variability in the adoption of e-assessment tools.

### Perceived benefits and challenges in adoption

While the perceived benefits of e-assessment tools, such as efficiency and scalability, were strong motivators for their adoption, these benefits were often weighed against the potential loss of personalized feedback. Participant 12 exemplified this tension by saying,

*“The tool reduced my workload, but I found myself supplementing it with manual feedback to ensure students received the guidance they needed” [P12].*

This quotation really sums up what makes maintaining the depth and quality of feedback in automatic ways a challenge (Figure 12).

*“Sometimes, students receive incorrect feedback with no explanation, hindering learning outcomes” [P5].*

These findings are confirmed by the survey data, which indicates mixed responses to the perception of personalized feedback through EA tools. The result is that 30.23% of the respondents strongly disagreed that their EA tool provided personalized feedback for each student, with 25.58% simply disagreeing with the statement. Such a reading would therefore point to the fact that more than half of the respondents feel there exist significant challenges with the level of personalized feedback offered by these tools. On the other hand, 25.58% agreed that the feedback is personalized and 18.60% did not agree nor disagree, thus suggesting that although some users find the feedback adequate or have mixed experiences about it, quite a big proportion still sees it as a major limitation (Figure 13).

## Differences between UK and Saudi Arabian contexts

### Cultural and educational differences impacting pedagogical and practical factors

It was observed that cultural differences generated a significant difference between the two educational contexts about the usage and valuation of personalized feedback. Participant 15 from KSA reflected that.

*“In our context, personalized feedback is not just a pedagogical tool but also a cultural expectation. Students expect detailed, individualized feedback, which automated tools sometimes fail to deliver” [P15].*

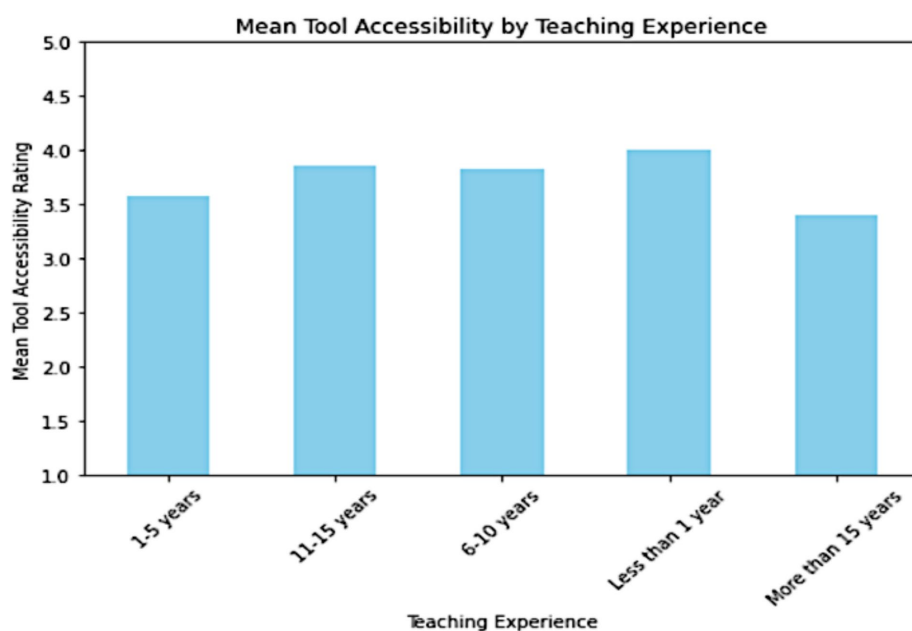


FIGURE 9  
Mean tool accessibility rating by teaching experience.



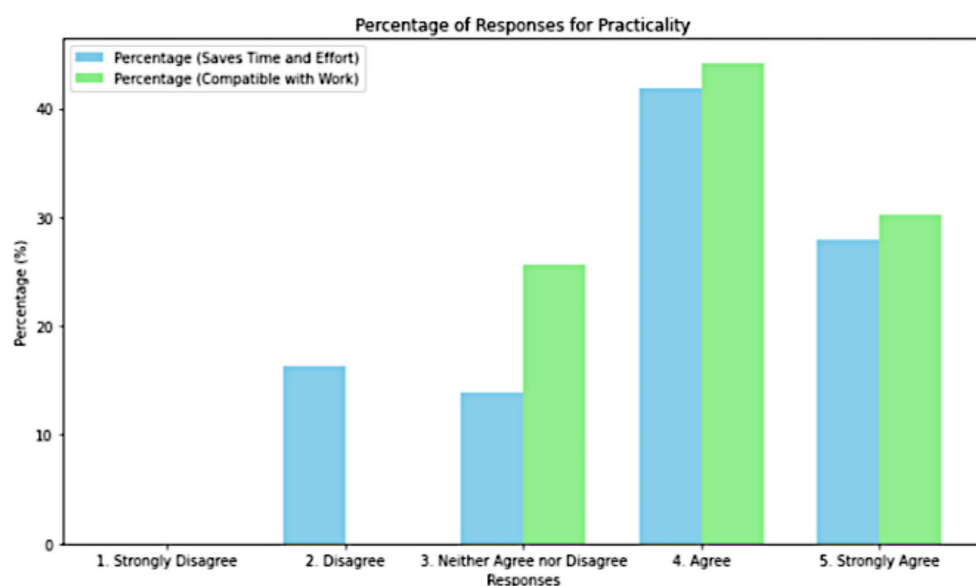


FIGURE 10  
Percentage of responses for practicality.

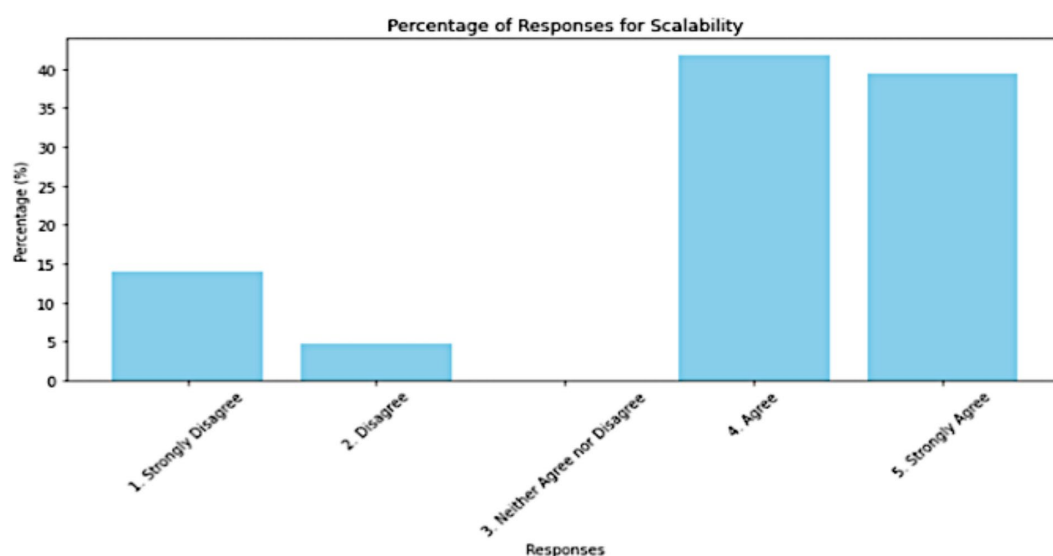


FIGURE 11  
Percentage of responses for scalability.

The practical challenges in implementing the EA tools were different in the UK and Saudi Arabia, mainly related to the infrastructures and the adaptation of the tool. Participant 7 from the UK also said that (Figure 14).

*“The tools that I have learned are easy, but sometimes it takes time to learn. You have to practice; you have to find the pathways to the correct setting. And that makes it very time consuming” [P7].*

In synthesis, the analysis draws differences in how practical challenges are considered in the UK and Saudi Arabian contexts with the cultural and infrastructural factors presented in the interviews.

Mainly, the problems that the UK instructors raise are technical barriers and the experience of the use of EA factors. On the other hand, Saudi Arabian instructors also feel that these tools are compatible and offer more sufficiency in terms of training time. Such findings emphasize the role of perceived usefulness, as instructors in different regions are more likely to adopt EA tools when they believe these tools meet their unique cultural and infrastructural needs, thereby enhancing their effectiveness in specific educational contexts (Davis, 1989). Such results indicate that strategies for increasing the effective and successful use of EA tools need to be very context-specific, attending to unique challenges and requirements within each educational environment. This will ensure that tools are better attuned

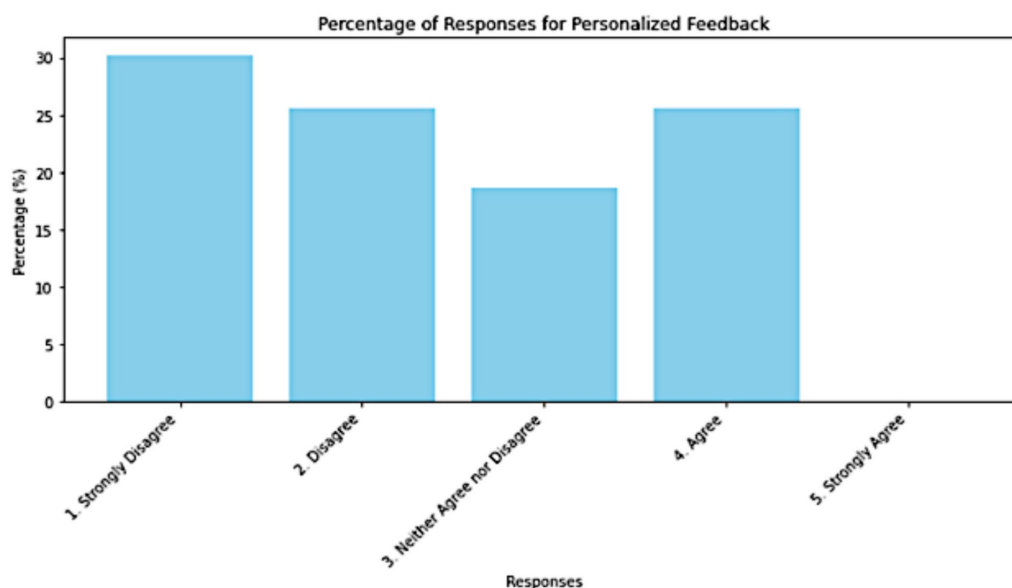


FIGURE 12  
Percentage of responses for personalized feedback.

to the specific infrastructural capabilities and cultural expectations of different regions (Figure 15).

### Specific barriers and incentives

The effectiveness of EA tools was restricted, despite the fact that their adoption was predominantly motivated by their potential to enhance feedback mechanisms. Participant 12 from Saudi Arabia said,

*“The tool reduced my workload, but I found myself supplementing it with manual feedback to ensure students received the guidance they needed” [P12].*

In synthesis, there are rather ambivalent views regarding the perceived benefits of EA tools in relation to feedback mechanisms. It is clearly indicated in this regard that while there are obvious incentives related to efficiency and more of workload reduction, at the same time, serious concerns are expressed about how these tools could be useful in providing quality-based personalized feedback helping the diverse requirements of instructors. This reflects the role of perceived usefulness, as instructors are more likely to support and use EA tools if they believe these tools not only improve efficiency but also effectively meet their pedagogical needs by delivering meaningful and individualized feedback (Davis, 1989). In the UK and Saudi Arabia alike, there is a noted tension between the perceived benefits of EA tools in terms of streamlining assessment processes and the perceived limitations in delivering feedback that is both meaningful and individualised (Figure 16).

### Impact of E-assessment tools on workload and feedback

#### Influence on workload and personalized feedback

As e-assessment reduced the time spent on grading, the adoption of it had a profound effect on the amount of time that instructors spend on an assignment. Participant 14 from Saudi Arabia commented,

*“The process of manually assessing student assignments, including downloading, assessing, and providing written feedback, all that let us spend lots of time and effort unlike e-assessment” [P14].*

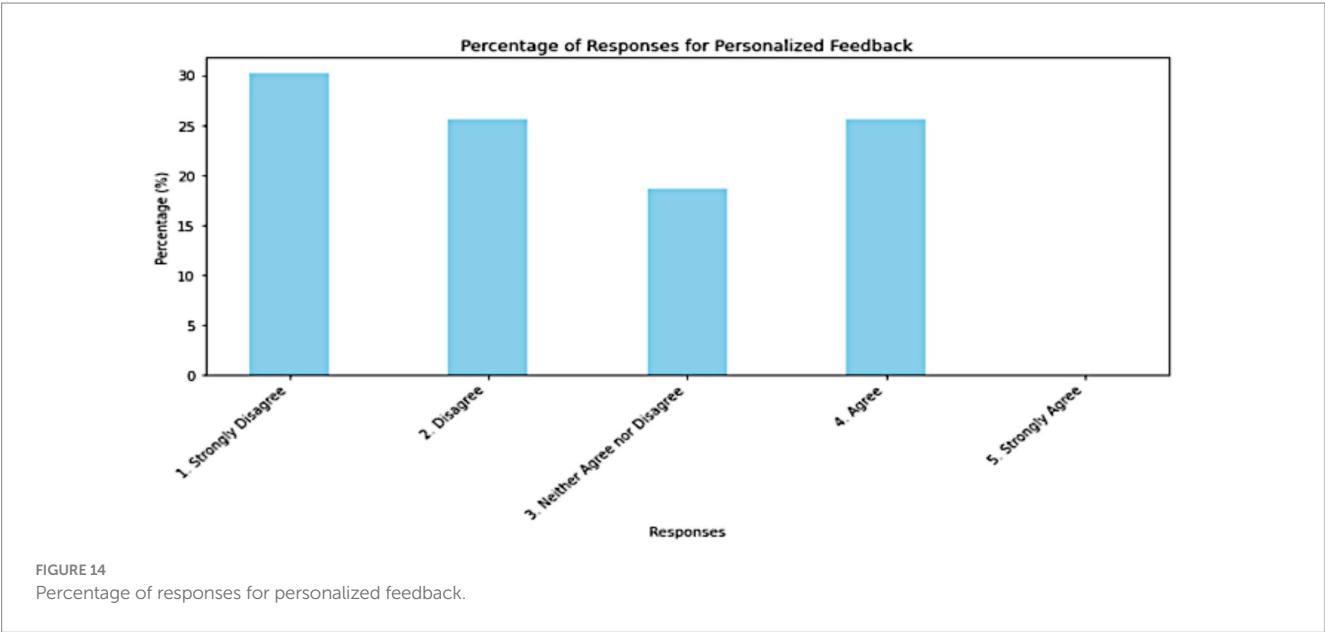
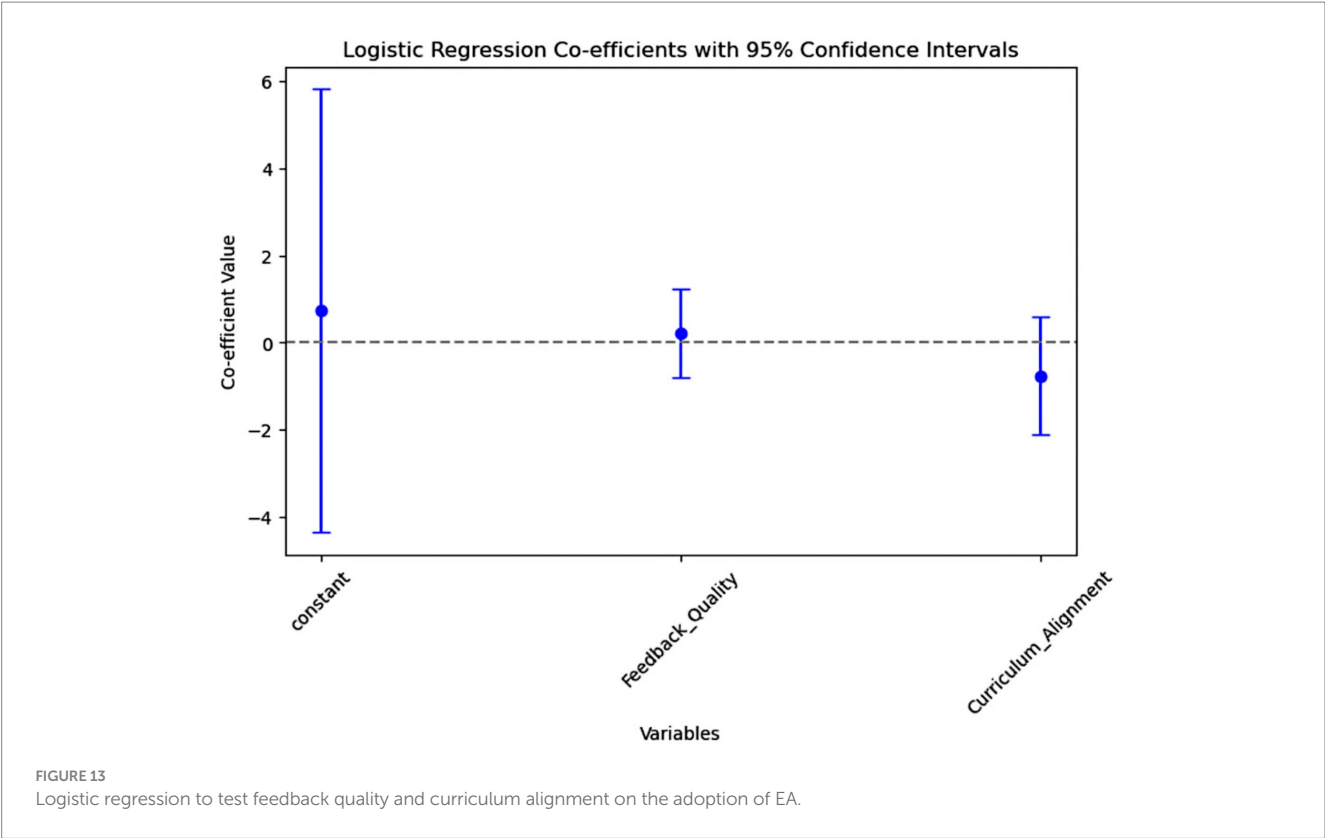
In synthesis, the results suggest that reducing workload and providing personalized feedback are perceived as benefits of EA tools, while at the same time there are associated perceived shortcomings. These perceived benefits, especially in grading tasks, are accruing evidence on such time-saving benefits, while the delivery of meaningful individualized feedback is also a great source of concern. This aligns with the perceived usefulness, as instructors are more likely to adopt EA tools if they see them as enhancing efficiency without compromising the depth and quality of feedback, indicating the need to balance between time-saving benefits and the tools' capability to deliver personalized, meaningful feedback (Davis, 1989). Both groups of instructors, UK and Saudi, are somewhat dissatisfied with the ability of the tools to match their expectations for personalized feedback suggesting the tools need balancing between efficiency and feedback's depth and quality (Figure 17).

#### Effects on feedback quality and instructors' teaching

The quality of feedback was a major concern, mainly on the ground that it should be not only timely but also personalized. Participant 6 stated,

*“Automated tools can provide feedback quickly, but they often lack the nuance needed to address individual student needs. Personalized feedback is critical for helping students improve their coding skills” [P6].*

To synthesize, the data clearly suggests a tension between efficiency and quality of feedback as delivered by EA tools. The tools are hailed for their timely delivery of feedback, but there is quite a bit of concern with respect to their ability to deliver nuanced, personalized guidance that hits the right spot in addressing individual student needs. The results showed that educators from the UK and



Saudi Arabia were very dissatisfied with the quality and personalization of feedback (Figure 18).

On the other hand, the use of EA tools significantly impacted the practices of instructors in giving feedback and engaging the learners. While these took away the amount of time needed to grade, at times they also added in terms of set-up and tweaking times. Participant 11, Saudi Arabia, reported.

“Main challenge in assessing code is length, takes too much time” [P11].

indicating how EA tools improved efficiency while at the same time constituting a source of workload investment necessary to assure quality feedback. Relatedly, Participant 13 noted,

“Teacher struggles with time constraints due to grading coding assignments, impacting ability to provide students with additional activities” [P13].

To synthesize, the data brings out a detailed image of how EA tools affect teaching practices in the UK and Saudi Arabia. Though

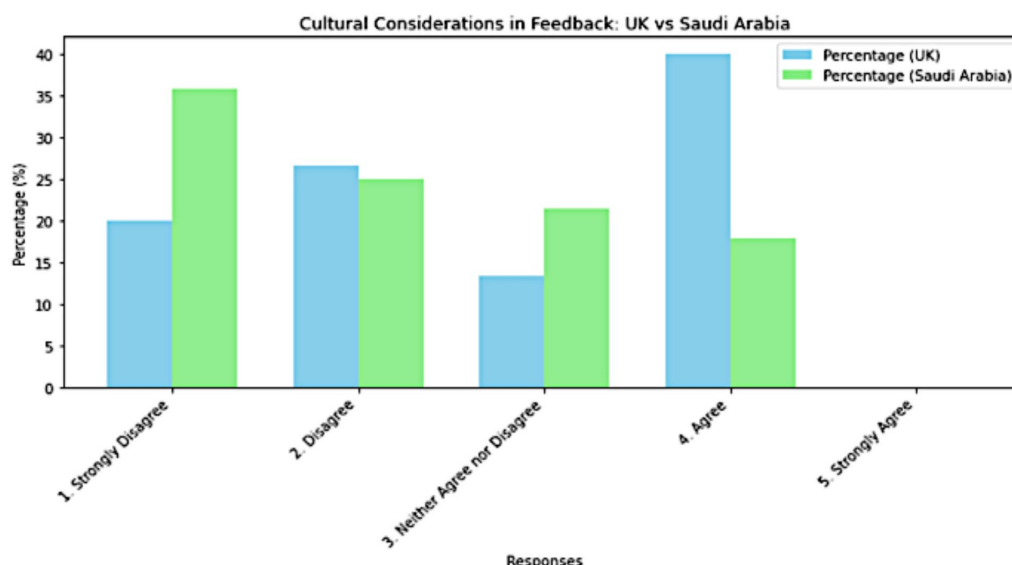


FIGURE 15  
Cultural consideration in Feedback UK vs. Saudi Arabia.

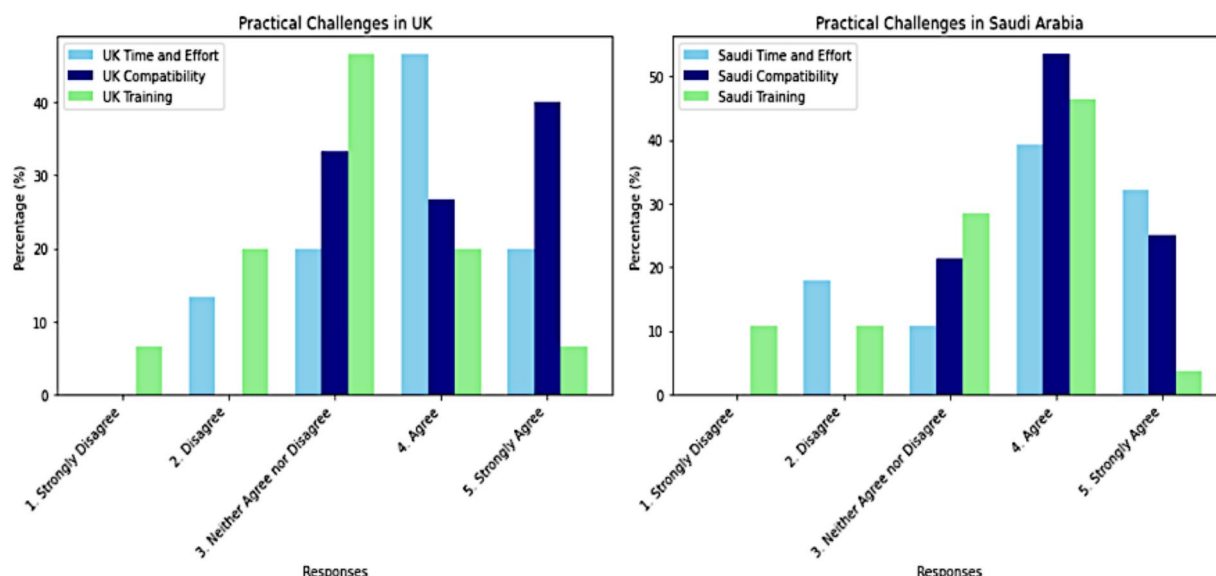


FIGURE 16  
Practical challenges in the UK and Saudi Arabia.

these tools are appreciated for efficiency and handling of work in bulk, there is still a lot of worry on their capacity to provide personal feedback. Further, the responses have also demonstrated that in general there is a favorable opinion on the compatibility of those tools with teaching practices; the share of those finding these tools not sufficient or only partially meeting their needs is also quite big. This indicates the role of perceived usefulness, as instructors are more likely to adopt EA tools if they see them as not only efficient but also as valuable in delivering detailed, personalized feedback that aligns with their teaching practices and supports meaningful learning outcomes (Davis, 1989).

These findings highlight the fact that the use of EA tools in higher education is fundamentally difficult, with practical balance enabling efficiency in one hand and provision of meaningful, personal feedback in the other. While these tools have undoubtedly brought major practical gains, for example, in scalability of the assessment process, flexibility of usage, and expanded access for instructors, such instrumental efficiency only defines their effectiveness in a very limited sense. Value in EA tools is derived from their capacity to provide authentic, reliable, valid, flexible, accessible, sailable, and practical (Figure 19).



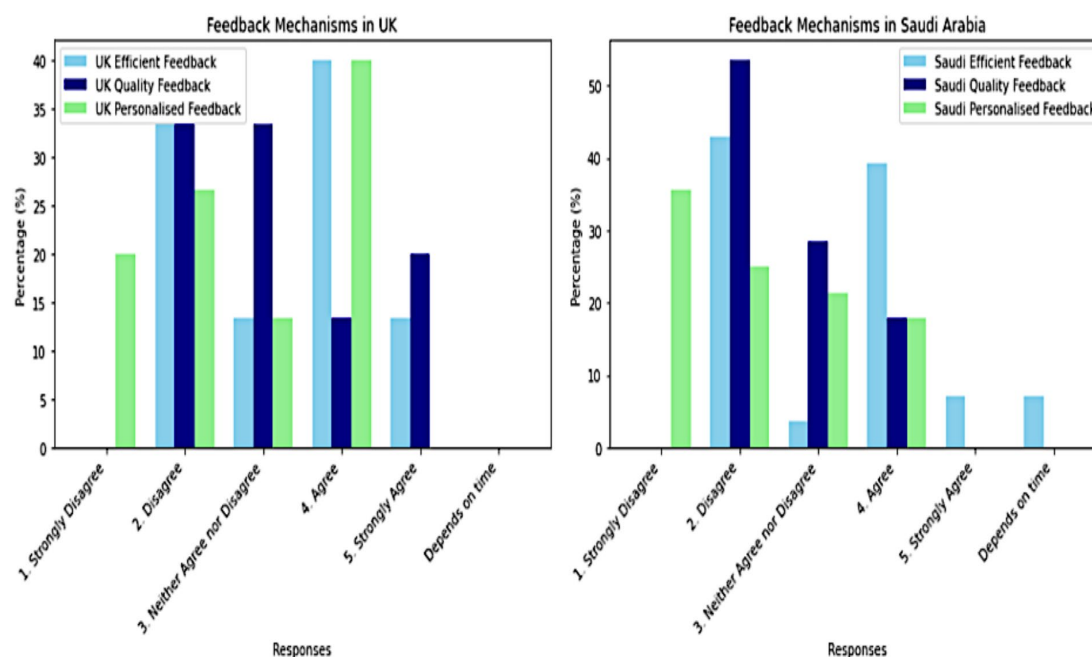


FIGURE 17  
Feedback mechanism in the UK and Saudi Arabia.

## Discussion

This section is a discussion based on the insights gained from the findings presented in the previous chapter, chapter 4 “Findings,” and informed by the relevant literature reviewed in chapter 2 “Background and Literature Review.” The four sections are organized to correspond with the research questions (RQ), facilitating a thorough and targeted examination of each inquiry. The first section shall discuss the success factors as well as challenges of using e-assessment in education from both pedagogical and practical perspectives. The second section outlines the main influences affecting adoption decisions about EA, while developing a fine-grained understanding of the perceived benefits and challenges constituting the nature of such decisions. The third section offers an analytical comparison of differences in UK and Saudi Arabian settings, showing ways that cultural and educational contexts influence the implementation and acceptance of EA tools. Section four specifically discusses how the use of EA tools by instructors impacts workload and feedback, therefore carving a niche in how teaching innovation in assessment can affect the quality of teaching and feedback. The chapter summarizes these discussions as it reflects on the broader implications for educational practice, focusing on practical and pedagogical to enhance the adoption and effectiveness of EA (Figure 20).

## Discussion of seven success factors and challenges for E-assessment

### Pedagogical factors

#### Authenticity

Findings indicated that the perceived usefulness of EA tools is effective in aligning with curricular goals and supporting practical

utility, but there was often concern that these tools lacked authenticity, and the depth needed to measure meaningful learning outcomes. As was discussed in Chapter 2, Dhawan (2020) clarified that technology enhanced learning platforms, while theoretically capable of facilitating continuous assessment to support curricular goals, depend significantly on how well they are integrated into the learning environment and the quality of their implementation (Dhawan, 2020). This suggests that EA tools need to be designed to capture higher-order cognitive skills and provide authentic learning experiences in terms of focusing on variety of coding tasks, real-world application and relevant feedback.

#### Reliability

The findings hint that while EA tools are reputedly capable of providing reliable and consistent feedback, this will only take place with proper setting up and adequate instructor training. This finding agrees with those of some previous studies. Appiah and Van Tonder, (2018) found that well-constructed EA tasks require a great deal of effort and skill on the part of instructors to ensure they elicit evidence of higher-order thinking skills rather than less desirable superficial learning. In support, Baartman et al. (2007) drew on the fact that the introduction of different assessment tools to simulations and scenario-based tasks can establish the validity and reliability of the e-assessments if incorporated into competency-based programs (Baartman et al., 2007). In synthesis, simulations and scenario-based tasks can be integrated with EA tools and comprehensive instructor training, this may significantly enhance the reliability of the assessment by allowing for deeper learning and critical thinking skills.

Additionally, the findings have established that there is a need of high reliability in EA tools through the reduction of technical complications and improvement of user interfaces. Similarly, studies like (Tinoca et al., 2014) have reiterated that intuitive designs and

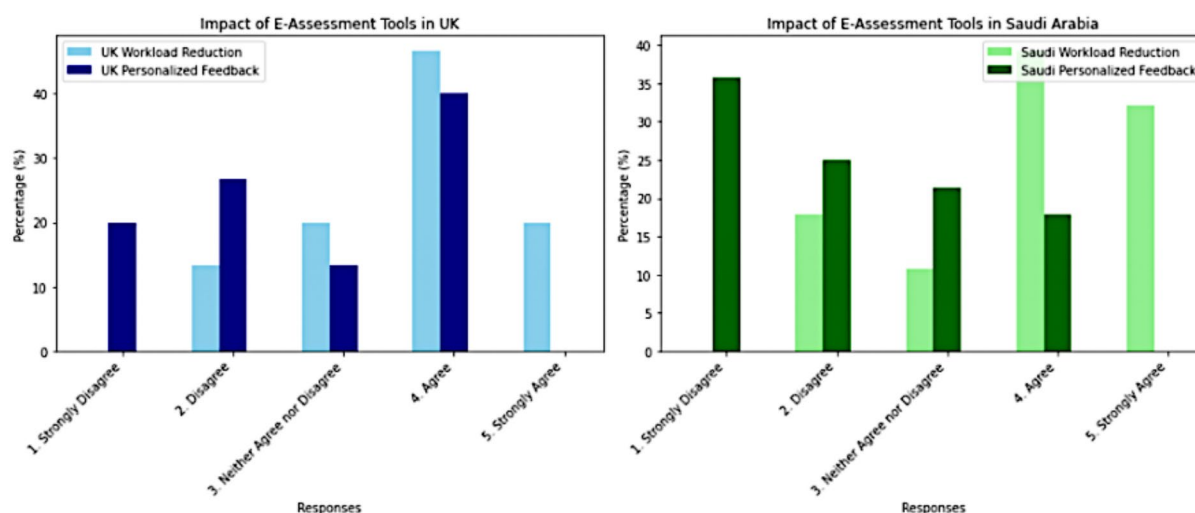


FIGURE 18  
Impact of EA tools in the UK and Saudi Arabia.

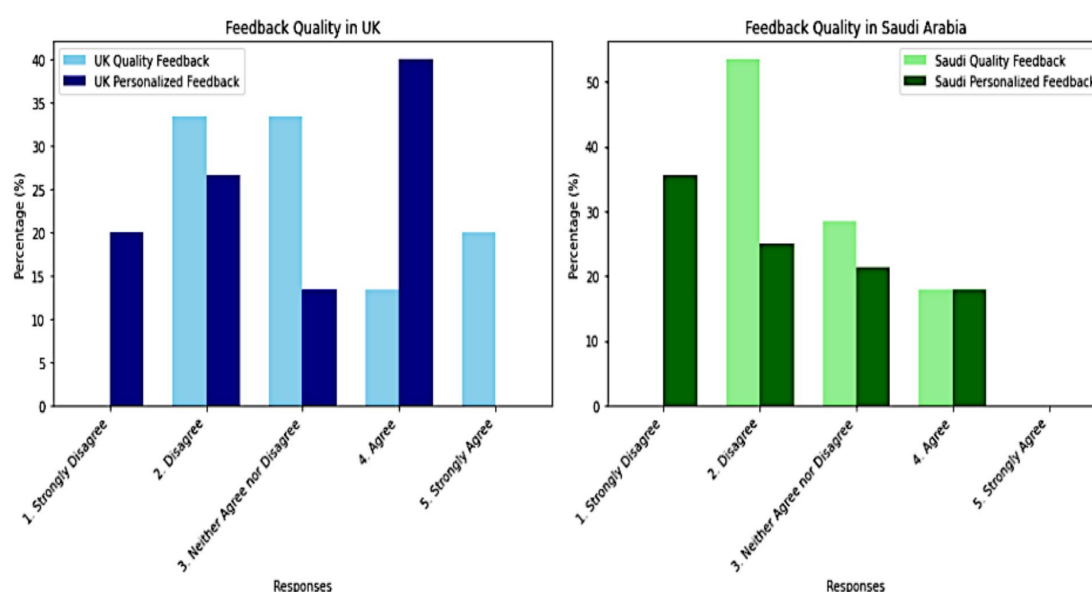


FIGURE 19  
Feedback quality in the UK and Saudi Arabia.

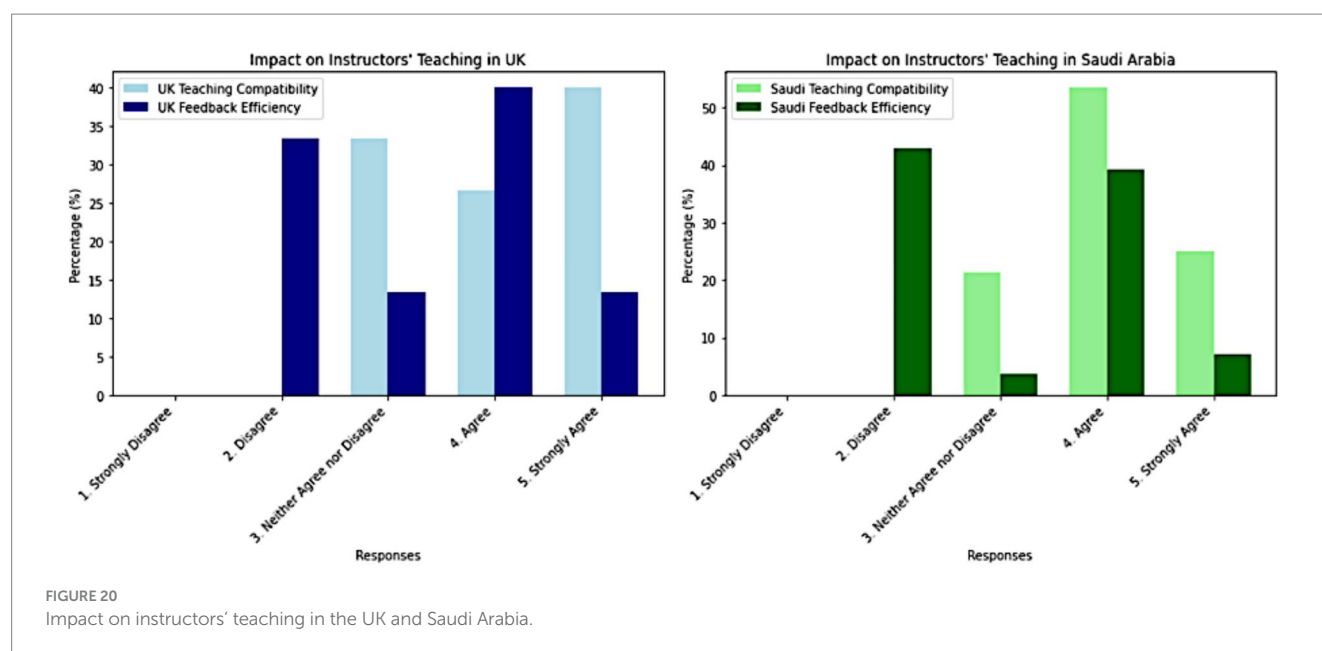
supportive training is essential to help instructors develop the assessment that will be able to validly measure the learning outcomes. This means that, increasing the support for the instructors and having more accessible tools leads to fostering confidence in EA tools through focusing on important features such as consistency over time, inter-rater reliability, test-retest reliability and supportive training.

### Validity

The findings suggest that while EA tools are valued for their flexibility and adaptability, the validity of these tools would thus be one of the largest impediments to broader acceptance. Recent study points out that instructors doubt that these tools can truly represent higher-order cognitive skills and provide valid assessment (Flanigan and Babchuk,

2022). This study supports the fact that complexity regarding validity assurance is one of the major factors leading to a lack of confidence in results and feedback produced through an EA tool, especially in the measurement of complex learning outcomes such as assessing coding.

Additionally, results indicate that there is a necessity of sufficient training on the part of the instructor and sufficient time to develop valid and reliable assessment tools. Similarly, the challenges about the validity of technology enhanced assessment discussed by (Wahas and Syed, 2024) draw attention to the need for extended training and support. As this is important, instructors should make provisions for appropriate design and implementation of e-assessment tools that go beyond surface-level evaluation. Training for e-assessment tools needs to consist of developing higher-order learning outcomes in terms of



content validity, consistency, and accuracy of the feedback, which would improve perceived usefulness and encourage wider diffusion.

## Practical factors

### Flexibility

The results indicate that most instructors value flexibility in the use of EA tools, yet technical and institutional constraints remain ticklish issues in their effective implementation. This corroborates the work done by [Al-Rahmi et al. \(2018\)](#), where they established that flexibility and adaptability were among the most important factors in the adoption of educational technologies ([Al-Rahmi et al., 2018](#)). Yet, these also identified similar issues; for example, some of the technical preconditions include web setup and institutional needs ([Al-Rahmi et al., 2018](#)). Consequently, it is imperative to overcome these technical and institutional obstacles to fully realise the potential of EA tools, thereby guaranteeing their effective and adaptable implementation in a variety of educational environments.

On the other hand, the perceived ease of use, as indicated in the findings, has remained one of the major determinants of ensuring that EA tools are effectively used. Similarly, several studies place ease of use, customizability, and cross-platform compatibility as key features that are instrumental in facilitating EA tools usage ([Khan and Khan, 2019](#)). In this study, while positive feedback from the questionnaire suggests appreciation for these adaptable features, however, interviews by comparison suggest that not all the tools assess coding successfully. This means that addressing these issues based on consideration of such factors with support for multiple programming languages, customization of assessment criteria, feedback mechanism and accessibility inclusivity would enhance perceived ease of use and increase adoption rates among instructors.

### Accessibility

The results indicate that most instructors found EA tools relatively accessible. However, the notable barriers were in terms of needing more practice and ongoing support from the technical

staff which means usability challenge. This key finding is also reflected in recent studies. For instance, ([Vayre and Vonthron, 2019](#)) state that continuous support and frequent training lead to the integration of technology. This means without proper support structures and training, instructors are likely to face challenges in using the EA tools to assess coding, which in turn somewhat discourages wider adoption and integration into teaching practices.

In support, [Aljawarneh \(2020\)](#), illustrates that a reduction in user interface complexity and encouragement of regular usage can lead to a reduction of time spent on training and an increase in overall acceptance of the tools ([Aljawarneh, 2020](#)). This study has further shown that ease of use design combined with effective training programmes can substantially improve perceived ease of use and ensure full realization of such technology's potential among instructors ([Aljawarneh, 2020](#)). Therefore, to solve these challenges through comprehensive training and support based on enhance platform compatibility and user interface design would be right at the heart of any strategy for improving perceived ease of use and adoption of EA tools.

### Practicality

Results from this study show that while highly valued by those instructors trying EA tools for their potential contribution to improved teaching, significant barriers to effective use remain. The most important of these are setup, maintenance, and implementation. These are all very time- and labour-intensive, hence complicating their integration into current workflows. This is evidenced by previous studies on technology adoption pointed out that perceived ease of use impacts on technology adoption ([Dimitrijević and Devedžić, 2021](#); [Scherer et al., 2021](#)). The study's findings also show the same indications of complexity and difficulties in setting up as recent studies on the adoption of educational technology ([Francom, 2020](#)). Therefore, making such tools simpler and reducing any complexities in setting them up becomes necessary to enhance perceived ease of use.

This study further identifies that limited support and lack of necessary training are other restraining factors in effectively implementing the different EA tools. In support of this, recent studies by Hamutoglu (2021) and Hébert et al. (2021) have also identified that if instructors are not properly trained, they cannot use new technologies even with recognition of their potential benefits (Hamutoglu, 2021; Hébert et al., 2021). Ensuring instructors are sufficiently supported and trained is a key to making EA tools a success. This means that, if the challenges are resolved based on enhancements such as these factors: ease of use, comprehensive training and feedback quality, then the use of EA tools within assessing coding could be considerably increased, resulting in ease of use and lessening the burden on instructors.

### Scalability

The findings imply that while EA tools are taken to be capable of scaling to meet larger class demands, there are challenges related to proper support, resources provided, and consistency and quality which need to be solved. Recent studies such as Cheng et al. (2021) and García-Morales et al. (2021) mentioned that if EA systems are to be 'truly scalable' they must be properly supported, resourced and configured otherwise for increases in class size consistency and quality of assessment will not be maintained (Cheng et al., 2021; García-Morales et al., 2021). This emphasis on the need for proper support and resources fits within the broader literature, suggesting that scalability is not solely a function of technological capability but rather how effectively these tools are integrated into educational practice.

Moreover, the critical role played by perceived usefulness, as explained by the findings, aligns with recent research studies on technology acceptance in education. Accordingly, Chiu et al. (2023) and Erguvan (2021) found that educators are more prone to adopt EA tools if they recognize these tools as having the capacity to handle higher volumes without compromising assessment quality. Therefore, consistency and reliability are the most important factors to address to increase adoption by instructors. Additionally, with adequate support, proper resources, volume capacity, quality of assessment scaling up EA tools in more effectively managed, thereby enhancing their effectiveness in larger educational settings.

### Differences between the UK and Saudi contexts: cultural and educational influences on E-assessment adoption

Cultural and educational contexts play a significant role in shaping how e-assessment (EA) tools are perceived, adopted, and integrated into teaching practices. The findings of this study highlight a marked contrast between the UK and Saudi Arabia in terms of instructors' attitudes toward EA particularly regarding feedback provision and overall system acceptance.

In Saudi universities, instructors tend to view EA tools more cautiously, especially when feedback is automated or lacks the depth of personal interaction. This is largely due to cultural and pedagogical expectations that prioritize close, personalized relationships between instructors and students. The limited availability of resources for generating meaningful feedback further reinforces negative perceptions of EA in this context. These observations align with previous research, such as Algahtani (2011), which emphasizes the importance of human interaction and personalization in Saudi educational environments [206].

By contrast, instructors in UK universities generally demonstrate a more pragmatic and open attitude toward EA tools. They tend to value the efficiency and productivity that technology can offer, particularly when the system is seen as capable of maintaining feedback quality and aligning with educational objectives. As noted by Wongvorachan et al. (2022), the perceived usefulness of EA especially its ability to deliver timely, goal-oriented feedback is a key determinant of its acceptance in UK settings [119].

These cross-cultural differences suggest that the adoption of EA tools is not merely a matter of functionality but one of cultural compatibility. Instructors are significantly more inclined to adopt EA platforms when these tools are designed to reflect and respect the cultural and educational values of their context.

For instance, in the Saudi context:

- Flexibility through personalization is essential, as tools must adapt to individual learning styles and cultural expectations.
- Validity must reflect hierarchy and respect, ensuring that assessment criteria resonate meaningfully with both learners and instructors.
- Accessibility must support collectivism, aligning with the group-oriented nature of teaching and learning in many Saudi institutions.
- Authenticity must incorporate contextual learning, presenting scenarios that are not only realistic but culturally relevant and engaging.

These culturally embedded dimensions are not just desirable features—they are essential enablers of EA adoption. When tools are designed with such cultural sensitivities in mind, they are more likely to be embraced by instructors and more effective in supporting meaningful, equitable learning experiences.

Ultimately, these findings reinforce the argument that successful implementation of EA tools in diverse higher education environments depends on more than just technical capabilities. It requires thoughtful integration of cultural values, pedagogical traditions, and institutional expectations. Future EA development should therefore adopt a culturally responsive approach, ensuring that educational technologies are not only functional but also contextually appropriate and inclusive.

### Challenges in preserving personalized feedback in E-assessment and workload reduction for instructors

According to the results, EA tools are thought to be helpful in decreasing the workload of instructors, particularly when it comes to assessing assignments, yet challenges remain regarding their ability to deliver meaningful, personalized feedback. This is consistent with findings from Emira et al. (2020), who emphasize that while EA tools can enhance efficiency, they often lack the depth of feedback needed to meet diverse educational goals [103]. Instructors in both the UK and Saudi Arabia expressed concerns over the tools' inability to provide rich, individualized feedback, pointing to a broader issue of balancing the need for personalized student support with the goal of reducing instructor workload a tension that lies at the core of this study.

Ryan et al. (2021) reinforce this perspective, noting that in digital learning environments, feedback quality remains a persistent challenge.



Their research shows that although automation can streamline processes, it often sacrifices the nuance and personalisation of feedback that many educators and learners consider essential [104]. These findings strengthen the argument that for EA tools to gain wider acceptance, especially in contexts like Saudi Arabia and the UK, developers must enhance their capacity to offer personalized, student-centred feedback without increasing the burden on instructors. Thus, while EA tools contribute significantly to efficiency, there is still considerable room for development to ensure they meet the pedagogical expectations and feedback standards required in higher education.

## The proposed framework

The framework identifies seven success factors for implementing EA in coding education (Figure 21).

### Pedagogical factors

#### Authenticity's influence on adoption

Authenticity is one of the determining factors for encouraging effective EA tools in coding at HE through a proposed framework. Authenticity of EA involves the creation of assessment tasks that have been created to represent real-world applications and relevant practical contexts. This approach aids in making the evaluation more relevant and appealing for students. The authentic assessment approach should, therefore, be directed toward various types of coding activities that can help students achieve deep understanding by exposing them to different kinds of challenges they could face in a real-life situation. In putting their knowledge of coding to practical use for real-life applications, students can build deeper levels of understanding by connecting theoretical knowledge.

Timely and personalized feedback are critical aspects of this approach. When this is achieved, students' learning process can be more emphatic and motivating. This in turn can lead to improvement in their coding practices. Another aspect is contextual learning assurance. Attention to culturally subtle nuances provides a means to secure relevant assessments, representative of the diversity of students.

All these aspects combine into one coherent and effective EA framework, fostering skill acquisition and authentic engagement in learning to code.

#### Reliability's influence on adoption

Reliability in EA concerns the consistency of assessment results over time, ensuring dependability and reproducibility across instances and different assessors. Reliability is a very important factor in the adoption of EA for coding in HE. It ensures that the assessments will come out consistent and dependable. Most importantly, there is a time consistency unstable results would yield the same data across different instances meaning instructors would have little faith in the process.

Another aspect is inter-rater reliability which focuses on different evaluators giving the same score for the same work. This discrimination would reduce subjective biases and develop fairness. In addition, test-retest reliability furthers trust by showing the repeated assessments would yield similar results under similar conditions. In turn, this demonstrates the robustness of the tool.

Moreover, simulations and scenario based tasks would enable students to reveal coding skills in dynamic, real life like contexts. This

leads to adding depth to the reliability of assessment. Lastly, supportive training for the instructors would mean that there is a harmonized application of the evaluation criteria applied and that they are inculcated with the EA tools.

#### Validity's influence on adoption

Validity refers to the precision with which EA measures what is supposed to be measured. So that the results reflect the students' real capabilities in coding and their learning outcomes. In other words, validity guarantees preciseness in the measurement of the coding competencies by the e-assessment.

In addition, content validity, indicates that the assessment should correspond as closely as possible to the skills and knowledge it purports to measure. Another aspect is consistency in the design of assessment whereby standards are set clear and uniformly across tasks.

Equally relevant is the accuracy of the feedback provided to the students with constructive, actionable insights that substantially improve the skills of students in coding. Moreover, supportive training equips instructors with necessary skills for effective assessment, just like in the case of reliability, while an emphasis on hierarchy and respect corresponding to cultural features develops respect and inclusiveness.

In other words, reliability and validity together provide a sound framework that furthers the appropriateness, inclusivity, and efficiency of EA for coding within HE.

### Practical factors

#### Flexibility's influence on adoption

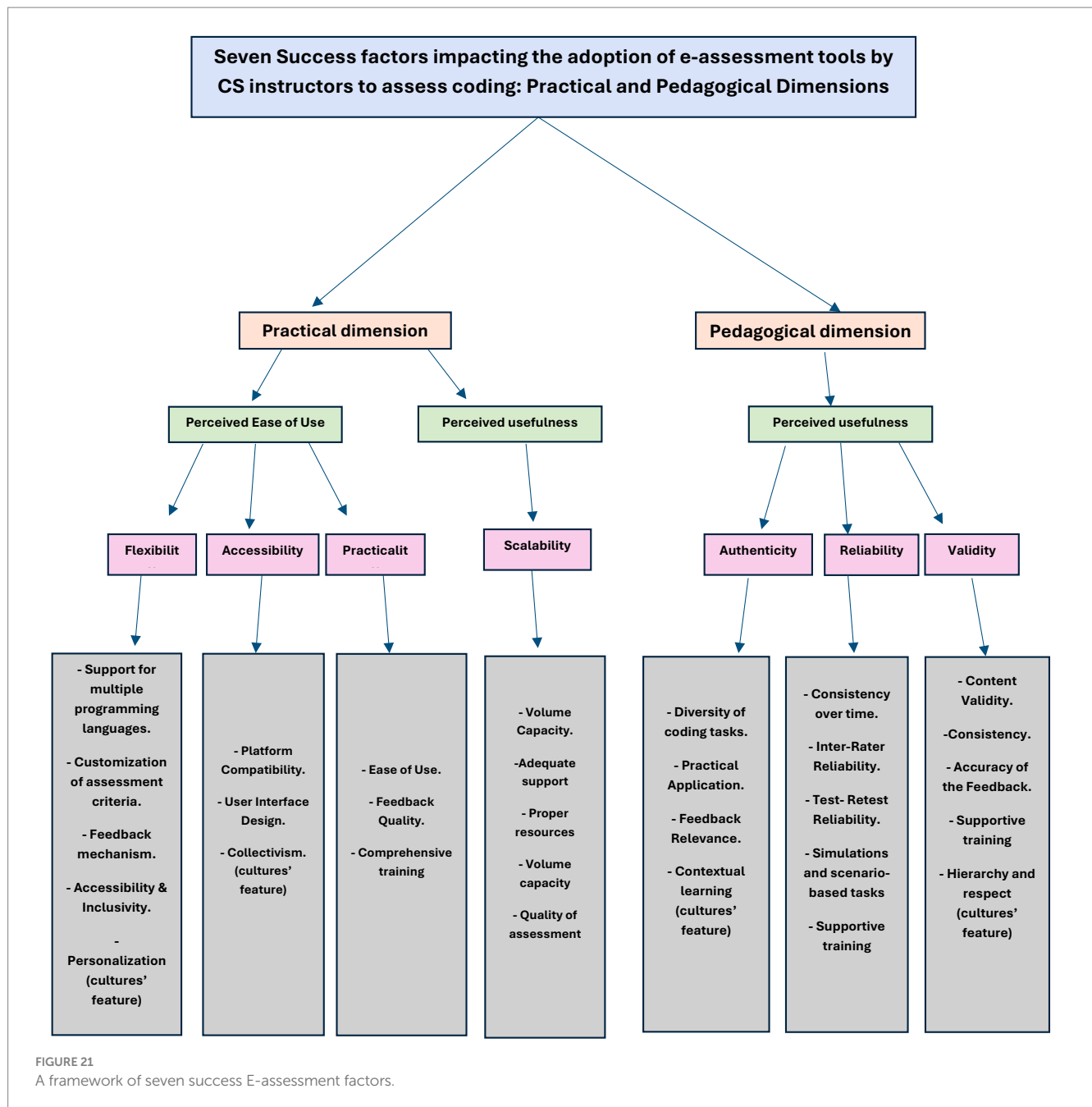
Flexibility in EA means that the platform can address a variety of instructional needs. So, instructors are allowed to adapt assessments, assessing multiple programming languages and providing responsive feedback. Therefore, they can implement features enabling inclusion of various teaching preferences. Flexibility provides the ability to accommodate what is necessary for ensuring the adoption rate of EA tools for coding education within HE institutions. This because instructors are afforded the opportunity to adapt this technology toward a multitude of goals based on diverse learning needs and educational contexts.

Another aspect is supporting many different programming languages. This makes EA more versatile for a great many coding classes and skill levels. In addition, having customized assessment criteria further provides flexibility in the instructor's work. This leads to designing ways of assessing work that more simply fit into course goals and competencies.

A thoughtfully designed feedback mechanism assists the instructors in providing timely, personalized feedback. This directly feeds into student learning. In addition, the accessible and inclusive platform of e-assessment can lead to using assistive technology in assessing coding. This can help in making changes in the user interface for better accommodation of diverse students' needs. Another important aspect of flexibility is cultural nuances and the variety of teaching contexts. This best empowers instructors to create assessments that are both relevant and inclusive. Lastly, with all-around support, flexible EA will form a base that will be able to answer the.

#### Accessibility's influence on adoption

Accessibility means that tools must be designed and developed in a manner that makes them accessible to every instructor through easy usage. Irrespective of their different technical skills or



resources. Accessibility has become one of the prime concerns in the adoption of e-assessment tools in coding education, which can help instructors provide appropriate inclusive assessments. It is further manifested in platform compatibility that lets instructors create content seamlessly across devices, operating systems and browsers. Hence managing flexible assessments without technical barriers.

In addition, intuitive usability at the user interface reinforces such accessibility by simplifying instructors' workflows. This can make it easy for them to set up, administer, and grade coding assessments with ease. Moreover, collectivism fostered within the platform encourages shared community and purpose among these instructors. This leads to fostering collaboration and resource-sharing across departments. Enabling instructors to have all these facilities, can help to focus on

delivery of high-quality, culturally responsive EA. Hence, features enable enhancements in the overall effectiveness of coding education in diverse HE environments.

### Practicality's influence on adoption

Practicality concerns the extent to which tools support instructors in integrating, managing, and conducting assessments efficiently, which means that there is no extreme complexity and time demands associated with the process of assessing coding. Practicality forms one of the important factors in the diffusion of EA platforms in coding education within HE institutions. Most important of all is ease of use. Instructors being able to navigate and use the platform with as little effort as possible; freeing them to focus on the evaluation of student competencies without technical obstacles.

Feedback quality also contributes to practicality. Since clear and personalized feedback gives the instructor a means to provide meaningful insights to students efficiently, thereby augmenting learning without increased workload. Additionally, comprehensive training further supports practicality by equipping the instructors with the necessary competencies and knowledge in the effective use of an EA system. Hence reducing time wasted over troubleshooting and building confidence in the use of the tool.

All these aspects of practicality taken together ensure efficient e-assessment experience; which can serve the needs of the instructors. Therefore, it makes reaching out to a wider assessment possible and seamless integration into coding education successful.

### Scalability's influence on adoption

Scalability is a significant factor affecting the adoption of EA tools in coding education. It refers to the ability of the platform to scale up to big volumes. Therefore, it can deal with big numbers in terms of students and assessments with no performance or quality issues. Key aspects include the quality of automated feedback. This means timely and accurate feedback are given to a significant number of students, which is critical in maintaining instructional quality when usage scales up.

In addition, support resources and good customer service help support instructors a great deal when issues come up and serve to provide a seamless experience irrespective of class size. This will further support scalability by ensuring that the platform responds

well. Moreover, it is reliable at times of peak usage with proper resources such as processing power and storage. There is finally the quality of the assessment that would remain constant, even while the platform scales to provide non-discriminatory, precise assessments that maintain the standards of education in coding programs.

All these scalability elements combined mean that EA tools can grow with the demands of education and enable their wider adaptation in higher education coding courses (Table 3).

### Future work

This thesis tackled the challenge and opportunity of an underexplored area, which is using EA in the assessing coding within HE from pedagogical and practical dimensions. Comprehensive insight into the practices of EA in this thesis is underpinned by a mixed method research approach. Therefore, more research is needed based on these findings to enhance further their robustness and to investigate new dimensions of this important field.

To further extend the findings from the present research, three main directions have been identified which are, advanced data analytics to further exploration into EA systems; ongoing engagement of various stakeholders, including students and industrialists, to ensure that assessments provide real-world ends, and the use of artificial intelligence to make electronic assessments scalable, adaptive and ethical.

TABLE 3 The novelty of the proposed framework.

Principle	Description	Key advantages	Examples in EA	Linked TAM component
Authenticity	Designing assessment tasks that reflect real-world coding applications and practical contexts. Encourages deep learning through real-life relevance.	<ul style="list-style-type: none"> <li>- Enhances engagement and motivation</li> <li>- Encourages deeper understanding</li> <li>- Promotes contextual and culturally responsive learning</li> </ul>	<ul style="list-style-type: none"> <li>- Real-world coding projects</li> <li>- Personalized feedback</li> <li>- Culturally relevant tasks</li> </ul>	<b>Perceived usefulness</b> Enhances relevance and learning outcomes, increasing users' belief in the system's value
Reliability	Ensures consistent, stable, and reproducible assessment outcomes across time and assessors.	<ul style="list-style-type: none"> <li>- Builds confidence in system accuracy</li> <li>- Reduces bias- Ensures fair evaluation</li> </ul>	<ul style="list-style-type: none"> <li>- Inter-rater rubrics- Test-retest reliability</li> <li>- Assessor training</li> </ul>	<b>Perceived usefulness</b> Builds trust in results, encouraging instructors to rely on the system
Validity	Accurately measures real coding skills and aligns with learning outcomes.	<ul style="list-style-type: none"> <li>- Measures true abilities</li> <li>- Improves fairness and learning</li> <li>- Informs better feedback</li> </ul>	<ul style="list-style-type: none"> <li>- Content-aligned tasks</li> <li>- Clear scoring standards</li> <li>- Constructive feedback</li> </ul>	<b>Perceived usefulness</b> Ensures assessment outcomes are meaningful and reflective of actual student performance
Flexibility	Adapts to different teaching styles, coding languages, and educational contexts.	<ul style="list-style-type: none"> <li>- Customizable</li> <li>- Inclusive for diverse learners</li> <li>- Adapts to course goals</li> </ul>	<ul style="list-style-type: none"> <li>- Multi-language support</li> <li>- Personalized rubrics</li> <li>- Assistive tech compatibility</li> </ul>	<b>Perceived ease of use</b> Flexibility reduces barriers and increases relevance to various user needs
Accessibility	Easy to use for all instructors regardless of technical background.	<ul style="list-style-type: none"> <li>- Lowers technical entry</li> <li>- Simplifies processes</li> <li>- Encourages collaboration</li> </ul>	<ul style="list-style-type: none"> <li>- Cross-platform compatibility</li> <li>- Clean UI</li> <li>- Community sharing features</li> </ul>	<b>Perceived ease of use</b> Reduces complexity, making the system more approachable and manageable
Practicality	Simple to integrate, set up, and use effectively without major effort.	<ul style="list-style-type: none"> <li>- Saves time</li> <li>- Reduces learning curve</li> <li>- Streamlines grading</li> </ul>	<ul style="list-style-type: none"> <li>- Instant feedback tools</li> <li>- Minimal setup</li> <li>- Instructor onboarding</li> </ul>	<b>Perceived ease of use</b> Supports quick adoption by minimizing time and effort needed to operate the system
Scalability	Maintains performance and quality as user demand increases.	<ul style="list-style-type: none"> <li>- Handles large classes</li> <li>- Maintains consistent outcomes</li> <li>- Supports institutional growth</li> </ul>	<ul style="list-style-type: none"> <li>- Automated feedback at scale</li> <li>- Reliable servers</li> <li>- High-volume performance</li> </ul>	<b>Perceived usefulness</b> Demonstrates capacity to support large-scale education efficiently

## Advanced data analytics

Although this thesis has followed a mixed-method approach to determining the pedagogical and practical effectiveness of EA, use of advanced data analytics would further make the system adaptable and accurate. Such analytics would complement the qualitative and quantitative insights, as already discussed, by the following means:

Firstly, continuous improvement of nuanced performance, identifying patterns of coding errors, time-on-task metrics, and response rates to feedback with machine learning and statistical tools. Secondly, enhancing system adaptability through real-time analysis of assessment data with predictions of student learning trajectories to inform modifications in subsequent assessments.

## Stakeholder engagement

These findings confirm that an e-assessment system needs stakeholder collaboration in the making and refining process. Further research in the area should focus on fostering stakeholders' collaboration in the proposed frameworks' validation and scaling in the following areas:

**Students:** Expanding the study to include students' experiences, perceptions of usability on e-assessment tools. Online questionnaires and participatory research methods could serve to highlight gaps in the current practice or investigate their needs and difficulties related to AI-enhanced e-assessment tools. Workshops and focus groups may reveal several practical tips for integrating these instruments into the existing curriculum.

**Industry Professionals:** Examining the alignment of e-assessment frameworks with professional-based coding competencies. Working with industry stakeholders will help ensure students are being assessed on skills relevant and transferable to the workplace.

## Leveraging AI for adaptive and ethical E-assessment

AI has the potential to reach completely new levels of personalization, scalability and ethics in EA in the education of coding. According to the Russell Group's principles on AI in education, any uses of AI should be designed explicitly to enhance learning outcomes with fairness and accountability.

Further studies can be conducted to understand how AI can be integrated into the EA systems as follows:

**Adaptive assessments:** using AI algorithms to automatically adapt the assessment tasks in real-time, immediately taking into consideration students' performance and learning profiles.

**Academic integrity:** through AI, design and develop detection tools showing practices amounting to unethical behavior, including plagiarism or abuse of code generation facilities. Academic integrity needs to be considered alongside issues of privacy and fairness in this aspect.

**Ethics of deployment:** the treatments of the AI system should be on an equal basis, and any deployment should be considered as a diverse data problem supported by guiding principles in terms of fairness and transparency.

**Scalability and efficiency improvement:** the automate complex assessment coding regarding code quality, functionality and creativity, in order to enable better management of large student cohorts.

Furthermore, the embedding of AI within ongoing ethical and pedagogical frameworks will point to ways in which the deployment of such technologies in HE can be done responsibly.

## Conclusion

This paper discusses success factors regarding e-assessment in coding education, how practical it would be for instructors to provide better learning outcomes with personalized feedback, and how seamlessly it would fit into the curriculum. Some of the key advantages are reduced effort by instructors and efficient evaluation, while some of the challenges pertain to training needs and technical limitations, which again indicates a way to improve and enhance the acceptance.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://doi.org/10.5258/SOTON/D2607>.

## Ethics statement

The studies involving humans were approved by the research ethics number 81152/Safiah Alamr/University of Southampton. Thee studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin.

## Author contributions

SA: Writing – original draft, Resources, Data curation, Visualization, Conceptualization, Validation, Writing – review & editing, Methodology, Investigation. ML: Validation, Writing – review & editing, Supervision, Conceptualization. LC: Validation, Conceptualization, Supervision, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Southampton University.

## 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.

## Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy,



including review by the authors wherever possible. If you identify any issues, please contact us.

## 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.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2025.1595171/full#supplementary-material>.

## References

- Abell, J., and Myers, G. (2008). Analyzing research interviews. *Qual. Discourse Anal. Soci. Sci.* 145:161. doi: 10.1007/978-1-137-04798-4\_7
- Adamu, N. (2023). "What is the point of assessments?" in Higher education. eds. L. Waller and S. K. Waller (Rijeka, Croatia: IntechOpen).
- Al-Emran, M., and Granić, A. (2021). "Is it still valid or outdated? A bibliometric analysis of the technology acceptance model and its applications from 2010 to 2020," in Recent Advances in Technology Acceptance Models and Theories, (Cham: Springer), 1–12.
- Algahtani, A. (2011). "Evaluating the effectiveness of the e-learning experience in some universities in Saudi Arabia from male students' perceptions." Durham University.
- Aljawarneh, S. A. (2020). Reviewing and exploring innovative ubiquitous learning tools in higher education. *J. Comput. High. Educ.* 32, 57–73. doi: 10.1007/s12528-019-09207-0
- Almuhanna, M. (2023). Improving E-Assessment based on university students' experiences. *Turkish Online Journal of Educational Technology* 22, 130–143.
- Al-Rahmi, W. M., Alias, N., Othman, M. S., Alzahrani, A. I., Alfarraj, O., Saged, A. A., et al. (2018). Use of e-learning by university students in Malaysian higher educational institutions: a case in Universiti Teknologi Malaysia. *IEEE Access* 6, 14268–14276. doi: 10.1109/ACCESS.2018.2802325
- ANOVA (2008). Analysis of variance (ANOVA). *Group* 1:3.
- Apostolellis, P., Wheeler, L., and Mandeltort, L. (2023). Assessing the impact of feedback on student learning using e2Logos: a novel grading tool for online student reports. Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V (Finland: Turku), 649
- Appiah, M., and Van Tonder, F. (2018). "E-Assessment in higher education: A review," *International Journal of Business and Management Economic Research*, 9, 1454–1460.
- Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A., and Van der Vleuten, C. P. M. (2007). Evaluating assessment quality in competence-based education: a qualitative comparison of two frameworks. *Educ. Res. Rev.* 2, 114–129. doi: 10.1016/j.edurev.2007.06.001
- Banerjee, A., Chitnis, U. B., Jadhav, S. L., Bhawalkar, J. S., and Chaudhury, S. (2009). Hypothesis testing, type I and type II errors. *Ind. Psychiatry J.* 18, 127–131. doi: 10.4103/0972-6748.62274
- Barra, E., López-Pernas, S., Alonso, Á., Sánchez-Rada, J. F., Gordillo, A., and Quemada, J. (2020). Automated assessment in programming courses: a case study during the COVID-19 era. *Sustainability* 12:7451. doi: 10.3390/su12187451
- Bowen, G. A. (2008). Naturalistic inquiry and the saturation concept: a research note. *Qual. Res.* 8, 137–152. doi: 10.1177/1468794107085301
- Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp0630a
- Brooks, G. P., and Johanson, G. A. (2011). Sample size considerations for multiple comparison procedures in ANOVA. *J. Mod. Appl. Stat. Methods* 10, 97–109. doi: 10.22237/jmasm/1304222940
- Buzzetto-More, N. A., and Alade, A. J. (2006). Best practices in e-assessment. *J. Inf. Technol. Educ. Res.* 5, 251–269. doi: 10.28945/246
- Care, E., Griffin, P., and McGaw, B. (2012). Assessment and teaching of 21st century skills. eds. Esther, C., Patrick, G., and Mark, W. Dordrecht, Netherlands: Springer.
- Castleberry, A., and Nolen, A. (2018). Thematic analysis of qualitative research data: is it as easy as it sounds? *Curr. Pharmacy Teaching Learning* 10, 807–815. doi: 10.1016/j.cptl.2018.03.019
- Cheng, M.-M., Lacaste, A. V., Saranza, C., and Chuang, H.-H. (2021). Culturally responsive teaching in technology-supported learning environments in marine education for sustainable development. *Sustainability* 13:13922. doi: 10.3390/su132413922
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., and Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Comp. Educ.* 4:100118. doi: 10.1016/j.caeai.2022.100118
- Cleophas, T. J., and Zwinderman, A. H. (2023). "Regression analysis" in Modern survival analysis in clinical research: Cox regressions versus accelerated failure time models. eds. T. J. Cleophas and A. H. Zwinderman (Cham, Switzerland: Springer), 1–8.
- Combéfis, S. (2022). Automated code assessment for education: review, classification and perspectives on techniques and tools. *Software* 1, 3–30. doi: 10.3390/software1010002
- Conrad, D., and Openo, J. (2018). Assessment strategies for online learning: Engagement and authenticity. Athabasca, AB, Canada: Athabasca University Press.
- Das, B. K., Jha, D. N., Sahu, S. K., Yadav, A. K., Raman, R. K., and Kartikeyan, M. (2022). "Analysis of variance (ANOVA) and design of experiments" in Concept building in fisheries data analysis. eds. B. K. Das, D. N. Jha, S. K. Sahu, A. K. Yadav, R. K. Raman and M. Kartikeyan (Singapore: Springer), 119–136.
- Davis, F. D. (1989). "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly*, 13, 319–340. doi: 10.2307/249008.
- Dhakal, K. (2022). NVivo. *J. Med. Libr. Assoc.* 110, 270–272. doi: 10.5195/jmla.2022.1271
- Dhawan, S. (2020). Online learning: a panacea in the time of COVID-19 crisis. *J. Educ. Technol. Syst.* 49, 5–22. doi: 10.1177/0047239520934018
- Dimitrijević, S., and Devedžić, V. (2021). Utilitarian and experiential aspects in acceptance models for learning technology. *Educ. Technol. Res. Dev.* 69, 627–654. doi: 10.1007/s11423-021-09970-x
- Erguvan, I. D. (2021). The rise of contract cheating during the COVID-19 pandemic: a qualitative study through the eyes of academics in Kuwait. *Lang. Test. Asia* 11:34. doi: 10.1186/s40468-021-00149-y
- Emira, M., Craven, P., Frazer, S., and Rahman, Z. (2020). "Assessment in the modern age: challenges and solutions," in *Learning and Performance Assessment: Concepts, Methodologies, Tools, and Applications*, IGI Global, 1248–1279.
- Evans, C. (2013). Making sense of assessment feedback in higher education. *Rev. Educ. Res.* 83, 70–120. doi: 10.3102/0034654312474350
- Ewens, W. J., and Brumberg, K. (2023). "Tests on means" in Introductory statistics for data analysis (Cham, Switzerland: Springer), 173–203.
- Fatmawati, E. (2015). Technology acceptance model (TAM) untuk menganalisis penerimaan terhadap sistem informasi perpustakaan. *J. Iqra* 9, 1–13.
- Fernandes, S., Abelha, M., and Oliveira, A. T. F. (2023). Assessment and evaluation in higher education. eds. Sandra, R. G. F., Marta, A., and Ana T. F. C. Basel, Switzerland: MDPI AG.
- Flanigan, A. E., and Babchuk, W. A. (2022). Digital distraction in the classroom: exploring instructor perceptions and reactions. *Teach. High. Educ.* 27, 352–370. doi: 10.1080/13562517.2020.1724937
- Foddy, W., and Foddy, W. H. (1993). Constructing questions for interviews and questionnaires: Theory and practice in social research. Cambridge, UK; Melbourne, Australia: Cambridge university press.
- Francom, G. M. (2020). Barriers to technology integration: a time-series survey study. *J. Res. Technol. Educ.* 52, 1–16. doi: 10.1080/15391523.2019.1679055
- García-Morales, V. J., Garrido-Moreno, A., and Martín-Rojas, R. (2021). The transformation of higher education after the COVID disruption: emerging challenges in an online learning scenario. *Front. Psychol.* 12:616059. doi: 10.3389/fpsyg.2021.616059
- Gilbert, L., Gale, V., Wills, G., and Warburton, B. (2009). JISC Report on E-Assessment Quality (REAQ) in UK Higher Education
- Goldin, I., Narciss, S., Foltz, P., and Bauer, M. (2017). New directions in formative feedback in interactive learning environments. *Int. J. Artif. Intell. Educ.* 27, 385–392. doi: 10.1007/s40593-016-0135-7
- Grant, J. S., and Davis, L. L. (1997). Selection and use of content experts for instrument development. *Res. Nurs. Health* 20, 269–274. doi: 10.1002/(sici)1098-240x(199706)20:3<269::res>3.0.co;2-g
- Green, S. B. (1991). How many subjects does it take to do a regression analysis. *Multivar. Behav. Res.* 26, 499–510. doi: 10.1207/s15327906mbr2603\_7
- Hamutoglu, N. B. (2021). Testing the effects of technological barriers on high school teachers' role in technology integration. *Asian J. Distance Educ.* 16, 74–89. doi: 10.5281/zenodo.4661743

- Hébert, C., Jenson, J., and Terzopoulos, T. (2021). "Access to technology is the major challenge": teacher perspectives on barriers to DGBL in K-12 classrooms. *E-Learn. Digit. Media* 18, 307–324. doi: 10.1177/2042753021995315
- Heil, J., and Ifenthaler, D. (2023). Online assessment in higher education: a systematic review. *Online Learning* 27, 187–218. doi: 10.24059/olj.v27i1.3398
- Hettiarachchi, E., Huertas, M. A., and Mor, E. (2015). The assessment life cycle: a model for analysing institutional e-assessment development\_114. *Int. J. Eng. Educ.* 31, 529–540.
- Hutchinson, D., and Chyung, S. Y. (2023). Evidence-based survey design: adding "moderately" or "somewhat" to likert scale options agree and disagree to get interval-like data. *Performance Improvement J.* 62, 17–24. doi: 10.56811/PFI-22-0012
- Kanika, S. C., and Chakraborty, P. (2020). Tools and techniques for teaching computer programming: a review. *J. Educ. Technol. Syst.* 49, 170–198. doi: 10.1177/0047239520926971
- Kaya-Capocci, S., O'Leary, M., and Costello, E. (2022). Towards a framework to support the implementation of digital formative assessment in higher education. *Educ. Sci.* 12:823. doi: 10.3390/educsci12110823
- Khan, S., and Khan, R. A. (2019). Online assessments: exploring perspectives of university students. *Educ. Inf. Technol.* 24, 661–677. doi: 10.1007/s10639-018-9797-0
- Likert, R. (1932). A technique for the measurement of attitudes. *Arch. Psychol.*
- Loureiro, P., and Gomes, M. J. (2022). "Online peer assessment for learning: Findings from higher education students," *Education Sciences*, 13:253. doi: 10.3390/educsci13030253
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nurs. Res.* 35, 382–385. doi: 10.1097/00006199-198611000-00017
- Mage, I., and Marini, F. (2023). Advancements in multivariate analysis of variance. *J. Chemom.* 37. doi: 10.1002/cem.3504
- Malatji, W. R., Van Eck, R., and Zuva, T. (2020). Understanding the usage, modifications, limitations and criticisms of technology acceptance model (TAM). *Adv. Sci. Technol. Eng. Syst. J.* 5, 113–117. doi: 10.25046/aj050612
- Marshall, B., Cardon, P., Poddar, A., and Fontenot, R. (2013). Does sample size matter in qualitative research?: a review of qualitative interviews in IS research. *J. Comput. Inf. Syst.* 54, 11–22. doi: 10.1080/08874417.2013.11645667
- Mekhiel, N. N. (2016). Introducing TAM: time-based access memory. *IEEE Access* 4, 1073–1085. doi: 10.1109/ACCESS.2016.2524447
- Mostoufi, N., and Constantinides, A. (2023). Linear and nonlinear regression analysis. *Applied Numerical Methods Chem. Eng.* eds. Navid, M., and Alkis C. 403–476.
- O'Brien, R. G., and Muller, K. E. (1993). Unified power analysis for t-tests through multivariate hypotheses. *Applied Analysis Variance Behav. Sci.*, 297–344.
- Paiva, J. C., Leal, J. P., and Figueira, Á. (2022). Automated assessment in computer science education: a state-of-the-art review. *ACM Trans. Comput. Educ.* 22, 1–40. doi: 10.1145/3513140
- Palmer, M. Online Assessment. (2023).
- Pressey, S. L. (1926). A simple apparatus which gives tests and scores-and teaches. *Sch. Soc.* 23, 373–376.
- Rajaram, K. (2023). "Assessment, assessment rubrics and feedback" in Learning intelligence: Innovative and digital transformative learning strategies: Cultural and social engineering perspectives (Springer), 367–426.
- Ramdhani, N. (2009). Model Perilaku Penggunaan Tik "Nr2007" Pengembangan Dari Technology Acceptance Model (Tam). *Bulet. Psikol.* 17, 17–27. doi: 10.22146/bpsi.11478
- Rao, N. J., and Banerjee, S. (2023). Classroom assessment in higher education. *Higher Education for the Future* 10, 11–30. doi: 10.1177/23476311221143231
- Rothwell, W. J. (2023). "Assessment and feedback" in Transformational Coaching for Effective Leadership (New York, NY, USA: Productivity Press), 335–347.
- Ryan, T., Henderson, M., Ryan, K., and Kennedy, G. (2021). "Designing learner-centred text-based feedback: a rapid review and qualitative synthesis," *Assess. Eval. High. Educ.*, 46, 894–912.
- Scherer, R., Howard, S. K., Tondeur, J., and Siddiq, F. (2021). Profiling teachers' readiness for online teaching and learning in higher education: who's ready? *Comput. Human Behav.* 118:106675. doi: 10.1016/j.chb.2020.106675
- Sotiriadou, P., Brouwers, J., and Le, T.-A. (2014). Choosing a qualitative data analysis tool: a comparison of NVivo and Leximancer. *Ann. Leis. Res.* 17, 218–234. doi: 10.1080/11745398.2014.902292
- Stubenrauch, C. J., and Lithgow, T. (2019). The TAM: a translocation and assembly module of the  $\beta$ -barrel assembly machinery in bacterial outer membranes. *EcoSal Plus* 8, 10–1128. doi: 10.1128/ecosalplus.esp-0036-2018
- Subheesh, N. P. (2023). "Leveraging Assessment Methodologies in Engineering Education: The Way Forward," SSRN, Paper 4335548, 2023. O. ANOVA, "Analysis of Variance (ANOVA)," Group, 1:3.
- Terziz, V., and Economides, A. A. (2011). The acceptance and use of computer based assessment. *Comput. Educ.* 56, 1032–1044. doi: 10.1016/j.compedu.2010.11.017
- Tinoca, L., Pereira, A., and Oliveira, I. (2014). "A conceptual framework for e-assessment in higher education: authenticity, consistency, transparency, and practicability" in Handbook of research on transnational higher education. ed. Igi, Global (Hershey, PA, USA: IGI Global), 652–673.
- Torrisi-Steele, G., Junior, G., Morrison, C., and Venema, S. (2023). Feedback on written assessments in universities: a topography of literature. *Int. J. Adult Educ. Technol.* 14, 1–15. doi: 10.4018/IJAET.316540
- Vayre, E., and Vonthron, A.-M. (2019). Relational and psychological factors affecting exam participation and student achievement in online college courses. *Internet High. Educ.* 43:100671. doi: 10.1016/j.iheduc.2018.07.001
- Venkatesh, V., and Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manag. Sci.* 46, 186–204. doi: 10.1287/mnsc.46.2.186.11926
- Wahas, Y. M. A., and Syed, A. J. A. (2024). E-assessment challenges during e-learning in higher education: a case study. *Educ. Inf. Technol.* 29, 14431–14450. doi: 10.1007/s10639-023-12421-0
- Wongvorachan, T., Bulut, O., Tsai, Y.-S., and Lindner, M. A. (2022). "Improving student feedback literacy in e-assessments: A framework for the higher education context," *Trends High. Educ.*, 1, 16–29.