Indonesian Journal of Electrical Engineering and Computer Science
Vol. 39, No. 2, August 2025, pp. 1035~1052
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v39.i2.pp1035-1052

Devising the m-learning framework for enhancing students' confidence through expert consensus

Teik Heng Sun[1,2], Muhammad Modi Lakulu[1], Noor Anida Zaria Mohd Noor[1]
[1]Faculty of Computing and Meta-Technology, Sultan Idris Education University, Perak, Malaysia
[2]Faculty of Information and Communication Technology, Tunku Abdul Rahman University, Perak, Malaysia

Article Info

Article history:

Received Sep 1, 2024
Revised Mar 19, 2025
Accepted Jul 1, 2025

Keywords:

Activity theory
Fuzzy Delphi method
Learning analytics
Mobile learning
Self-confidence

ABSTRACT

Past research has shown the relationship between self-regulated learning (SRL) and academic success. Self-regulated learners will monitor their learning, reflect on what they have learnt, adjust their learning strategies accordingly, and repeat this entire process throughout their learning. The ability to perform SRL will require the individual to have the belief and confidence in his/her capacity to succeed and accomplish the tasks. Therefore, this study aims to devise a mobile learning (m-learning) framework for enhancing the students' confidence. To achieve this, the Fuzzy Delphi method was used to validate the proposed framework where the survey questionnaire was distributed to 21 experts who are the experts in their respective fields for their consensus to be obtained. Consensus showed that "assessment data" can indicate the students' confidence when they attempt the assessment. Experts opined that "goal expectation," and "viewed lessons, chapters, or syllabus" exert the most influence on the students' confidence when they attempt their assessment. There was strong consensus from experts that "data security" is the most important element in the system infrastructure, and the "text mining technique" element can be used to evaluate the students' confidence.

Corresponding Author:

Muhammad Modi Lakulu
Faculty of Computing and Meta-Technology, Universiti Pendidikan Sultan Idris
Perak, Malaysia
Email: modi@meta.upsi.edu.my

1. INTRODUCTION

Self-confidence refers to an individual's belief in their ability to successfully complete tasks and achieve goals. Research suggests that self-confidence has a significant impact on academic achievement, and academic success, in turn, reinforces self-confidence [1], [2]. Students with higher confidence levels are more likely to engage in learning and embrace challenges. The cognitive aspect of self-confidence plays a crucial role in learning and can be assessed through self-report questionnaires or performance-based evaluations. In contrast, epistemic confidence refers to an individual's certainty regarding the accuracy, validity, and reliability of their knowledge. This form of confidence is particularly important in decision-making and problem-solving. Several tools exist for measuring general self-confidence, including the personal evaluation inventory (PEI) developed by Shrauger and Schohn [3], Agnihotri's self-confidence inventory (ASCI) [4], and the Pandey self-confidence inventory (PSCI) [5]. However, these instruments assess confidence in personal abilities rather than epistemic confidence, which relates to certainty in one's knowledge. Confidence-based marking (CBM) is an approach that encourages students to assess the accuracy of their answers by rewarding them for appropriately calibrated confidence. This method fosters self-reflection and metacognitive awareness, which are closely linked to epistemic confidence [6]. Another tool, the metacognitive awareness inventory (MAI), developed by Schraw and Dennison [7], evaluates metacognitive awareness, including self-regulation and confidence in one's knowledge. However, it is not explicitly designed to measure epistemic confidence. Despite the availability of self-report measures, assessing epistemic confidence remains challenging, as some individuals may exhibit overconfidence, while others may underestimate their knowledge. These biases highlight the need for objective and reliable assessment methods to measure epistemic confidence accurately.

This study aims to develop an m-learning framework to enhance students' confidence. The framework is based on Engeström's [8] extended activity system and is validated using the fuzzy Delphi method (FDM). Engeström's extended activity system expands upon activity theory (AT), originally introduced by Lev Vygotsky, later developed into a conceptual framework by Aleksei Leont'ev, and subsequently extended by Engeström [9]. Engeström's [9] extended activity system consists of six components, whose interactions produce the "outcome". These components are "tools," "subject," "rules," "community," "division of labour," and "object". The "subject" refers to an individual or group engaged in the activity, while the "object" represents the goal or problem space, which is transformed into the "outcome". The "subject" utilizes "tools" to achieve objectives, with the entire system governed by "rules". The "community" consists of individuals sharing a common object and contributing to the activity, while the "division of labour" defines the distribution of tasks based on roles and relationships within the community. The framework devised in this study that is based on Engeström's extended activity system is shown in Figure 1. In the devised framework, the "tools" component is revised to "technology enabler," while the "object" component is renamed "objective" to align with course learning outcomes. The framework is centered around the "confidence model," which comprises "assessment data," "activity and miscellaneous data," and "output." These components rely on "machine learning and data analysis" for assessment analytics. The elements within the "objective" component vary depending on the course implementing the framework, as each course has its own learning outcomes.

The FDM has been widely used in existing studies to determine consensus among experts and validate research elements. For instance, FDM was used to evaluate expert consensus on the usability and relevance of the school leaders' competencies in data-driven decision making (SLC3DM) model. The study pursued two primary objectives: first, to examine the validity and utility of the model's competency components; and second, to determine the overall applicability of the framework [10]. Similarly, [11] used FDM to reach a consensus on the role of the school environment and classroom management in pre-schoolers' moral development. Masnan *et al.* [11] validated and determined key items through a survey instrument analyzed by an eleven-member expert panel, aiming to establish agreement on the importance of the physical school environment and classroom management in fostering healthy moral development in preschool children. FDM has also been applied in various other studies. For example, [12] utilized FDM to identify essential elements for an online problem-based learning module in an Islamic studies course. In another study, [13] employed FDM to validate the design of a Spanish-language questionnaire on the "quality of dying in long-term care" by obtaining expert consensus through fuzzy opinions. Additionally, [14] applied FDM to determine the key factors and items to be included in a framework for evaluating the success of m-government services in Malaysia. Furthermore, [15] used FDM to validate the essential competencies that programming teachers should possess to effectively teach students the principles of programming. The popularity of FDM stems from its ability to handle uncertainties, provide stronger consensus-building among experts, offer flexibility in expressing opinions, reduce sensitivity to outlier responses, and enable quantitative analysis of fuzzy opinions.

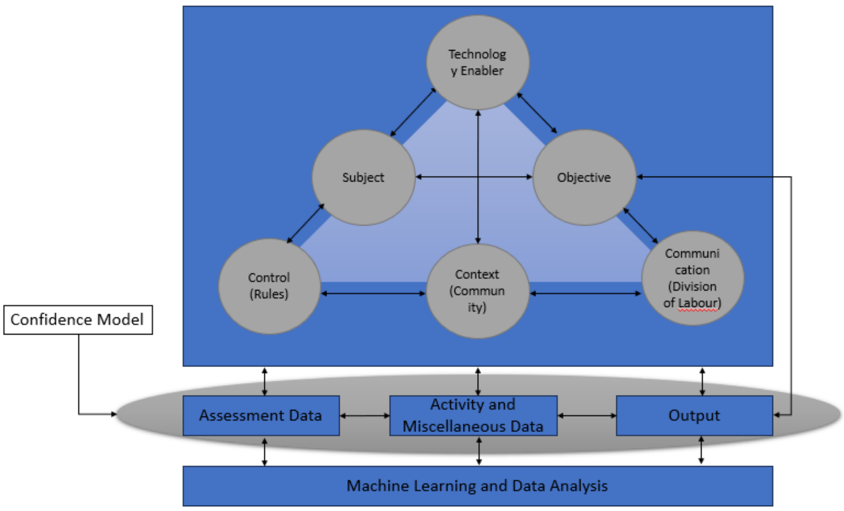

Figure 1. The devised framework

This paper is structured as follows. Section 2 provides a detailed explanation of the methodology employed in this study, outlining the key approaches and techniques used for data collection and analysis. Section 3 presents the results, highlighting the significant findings. Section 4 offers a comprehensive discussion of the findings, interpreting their relevance in the context of existing literature and suggesting possible directions for future research. Implications and potential applications are also discussed in this section. Finally, section 5 summarizes the key conclusions of this study.

2. METHOD

The development of the framework followed five distinct phases: analysis, design, development, implementation, and evaluation. To establish expert consensus and validate the framework components and elements, the FDM was employed. FDM is an enhanced version of the traditional Delphi method, integrating fuzzy set theory while maintaining its foundational principles. This approach ensures a more robust and systematic consensus-building process among experts. The conventional Delphi method, while widely used, is often time-intensive and costly due to its requirement for multiple rounds of expert feedback to reach a consensus. In contrast, FDM refines this process by leveraging fuzzy number sets, enabling a more efficient and structured decision-making framework [16]. The summary for each phase of the framework development is given in Figure 2.

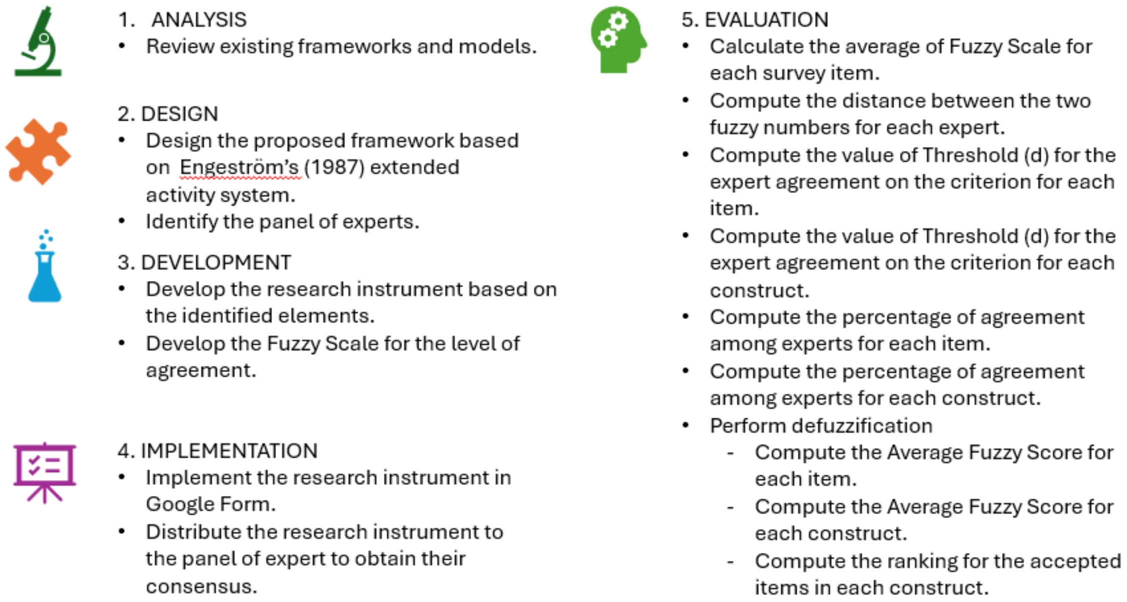

Figure 2. Phases involved in devising the framework

2.1. Analysis phase

During the analysis phase, TEN (10) existing frameworks and models were selected for thorough analysis and synthesis. The elements for the proposed framework were extracted from the existing frameworks and models. The literatures for the frameworks and models are listed below:

a) "Learning analytics framework for educational virtual worlds" [17].
b) "An assessment analytics framework (AAF) for enhancing students' progress" [18].
c) "A framework for learning analytics in Moodle for assessing course" [19].
d) "Let's talk learning analytics: a framework for implementation in relation to student retention" [20].
e) "Developing learning analytics design knowledge in the 'Middle Space': the student tuning model and align design framework for learning analytics use" [21].
f) "Assessment analytic theoretical framework based on learners' continuous learning improvement" [22].
g) "A complete validated learning analytics framework: designing issues from data preparation perspective" [23].
h) "Using design-based research to develop a mobile learning framework for assessment feedback" [24].
i) "An ontology-based framework of assessment analytics for massive learning" [25].
j) "Temporal learning analytics for computer based testing" [26].

2.2. Design phase

During the design phase, the framework was designed based on Engeström's extended activity system. Elements corresponding to each component of the framework were identified through a literature review, as outlined in Table 1 (in Appendix). A research instrument was then constructed based on these identified elements. Additionally, experts were selected to evaluate the framework. Experts are individuals with specialized knowledge and experience in a given field [27]. To ensure validity, consensus on the framework's components was sought from a panel of experts. The recommended panel size varies between 15 and 25 participants [14], [15], [28]. Selected experts were required to have a minimum of five years of experience in their respective fields and the ability to provide informed insights [29]. A total of 21 experts participated in the fuzzy Delphi study, including 17 subject matter experts, 9 information systems specialists, 2 educational technologists, and 1 educational psychologist. The composition of the expert panel is presented in Table 2. The experts were invited to complete a questionnaire distributed via the Google Forms platform. Additionally, the system recorded respondents' email addresses and the time of questionnaire submission.

Table 2. Participants of the fuzzy Delphi study

Expert no.	Job title (designation)	Field expertise	Length of service in the field of expertise
1.	Deputy Dean	Subject matter experts; information systems experts	6-10 years
2.	Associate Professor	Information systems experts	21-30 years
3.	Head of Department; Assistant Professor	Subject matter experts; information systems experts	6-10 years
4.	Deputy Dean	Subject matter experts	11-20 years
5.	Deputy Dean	Subject matter experts	11-20 years
6.	Senior Lecturer	Educational technologies	over 30 years
7.	Head of Department	Computer science	11-20 years
8.	Associate Professor	Information systems experts	11-20 years
9.	Dean	Subject matter experts	21-30 years
10.	Professor	Subject matter experts	21-30 years
11.	Senior Lecturer	Subject matter experts	11-20 years
12.	Assistant Professor	Subject matter experts; information systems experts	6-10 years
13.	Associate Professor; Head of School	Subject matter experts; information systems experts	21-30 years
14.	Head of Department; Associate Professor	Subject matter experts; information systems experts	21-30 years
15.	Associate Professor	Subject matter experts; information systems experts	11-20 years
16.	Lecturer	Subject matter experts	11-20 years
17.	Associate Professor	Subject matter experts; educational psychologist; multimedia	11-20 years
18.	Senior Lecturer	Subject matter experts	6-10 years
19.	Senior Lecturer	Subject matter experts	6-10 years
20.	Senior Lecturer	Information systems experts	21-30 years
21.	Senior Lecturer	Educational technology	6-10 years

2.3. Development phase

In this phase, a questionnaire was developed as the primary research instrument. Using questionnaires provided several advantages, including cost-effectiveness, particularly when distributed online, the ability to collect data from a large sample efficiently, consistency in question delivery, and respondent anonymity, which encourages honest and candid responses. The questionnaire employed a five-point Likert scale with response options: "strongly disagree," "disagree," "moderately agree," "agree," and "strongly agree." The five-point Likert scale was selected to reduce the likelihood of respondents over analyzing their choices. Increasing the number of scale points can introduce measurement errors due to respondent confusion [30]. Feedback was obtained from experts regarding the relevance and significance of the questionnaire.

Several studies have effectively applied the five-point Likert scale within the FDM framework [28], [31]. These studies demonstrate its applicability in mobile learning research, particularly in capturing expert consensus on various elements and competencies. In the FDM, triangular fuzzy numbers (TFNs) are employed to construct a fuzzy scale that translates linguistic variables. Fuzzy scales effectively capture human perceptions, preferences, and judgments, which are often imprecise or subjective. Rather than relying on fixed categories, fuzzy boundaries provide a more accurate representation of responses such as "slightly disagree" or "moderately agree". TFNs are defined by three values: m1, m2, and m3, where m1 represents the minimum value, m2 denotes the most reasonable or expected value, and m3 indicates the maximum value [32]. The TFNs used to generate the fuzzy scale are presented in Table 3.

Table 3. Triangular fuzzy numbers for linguistic variables

Linguistic variables	Fuzzy scale		
	m1	m2	m3
Strongly agree	0.6	0.8	1.0
Agree	0.4	0.6	0.8
Moderately agree	0.2	0.4	0.6
Disagree	0.0	0.2	0.4
Strongly disagree	0.0	0.0	0.2

2.4. Implementation phase

Google Forms was used to create the validated questionnaire. The advantages of using Google Forms included an intuitive user interface for questionnaire creation, ease of distribution, seamless access for participants on any internet-connected device, the ability for participants to save partially completed responses and continue later, downloadable responses in CSV format for statistical analysis, and built-in data validation functionality that helped reduce data entry errors. Emails containing the Google Forms link were sent to experts for their responses. A follow-up email and/or phone call was made to non-respondents after two weeks to remind them to complete the questionnaire.

2.5. Evaluation phase

For each survey item, the average value of the fuzzy scale for that item was calculated for the 21 experts using the (1)-(3) where n1ij, n2ij and n3ij represented the fuzzy scales of expert i for item j.

$$m_{1j} = \frac{\sum_{i=1}^{21} n_{1ij}}{21} \tag{1}$$

$$m_{2j} = \frac{\sum_{i=1}^{21} n_{2ij}}{21} \tag{2}$$

$$m_{3j} = \frac{\sum_{i=1}^{21} n_{3ij}}{21} \tag{3}$$

The level of consensus among the experts was determined by the value of the threshold (d). For each survey item j, the distance d_{ij} between the average value of the fuzzy scale for item j and the fuzzy scale of expert i was computed using the (4):

$$d_{ij}(\tilde{m}, \tilde{n}) = \sqrt{\frac{1}{3}[\left(m_{1j} - n_{1ij}\right)^2 + \left(m_{2j} - n_{2ij}\right)^2 + (m_{3j} - n_{3ij})^2]} \tag{4}$$

Therefore, the value of the threshold (d_j) for item j can be computed as:

$$d_j = \frac{\sum_{i=1}^{21} d_{ij}(\tilde{m}, \tilde{n})}{21} \tag{5}$$

Typically, a threshold (d) value less than or equal to 0.2 indicates expert agreement on a criterion. A smaller value of d signifies a stronger consensus among experts. For instance, studies in [12], [33] considered a criterion to have expert consensus if the threshold (d) was less than or equal to 0.2. Additionally, another condition that must be fulfilled is that the percentage of agreement among experts must be at least 75% [34], [35]. The percentage of agreement among experts for item j was calculated using the (6):

$$PA_j = \frac{Number\ of\ expert\ with\ \ d < 0.3}{21} \times 100\% \tag{6}$$

The (6) was used to compute the percentage of agreement among experts for each construct, where n denotes the number of items in the construct, and j represents the index of an item within the construct. The percentage of agreement for item j was denoted by PA_j.

$$PA_c = \frac{\sum_j^{j+n-1} PA_j}{n} \tag{7}$$

During defuzzification, the average fuzzy score was computed. If the obtained value was greater than the alpha-cut value (α=0.5), the measured item was accepted based on expert consensus. Conversely, if the value was less than the alpha-cut value, the measured item was rejected. The fuzzy score for an item was computed using the (8):

$$Fuzzy\ Score, A_j = \frac{1}{3} \times \left(m_{1j} + m_{2j} + m_{3j}\right) \tag{8}$$

3. RESULTS

The results presented in Table 4 indicated that the average threshold value (d) for all components was less than 0.2, while the average percentage of expert consensus exceeded 75%. Additionally, all components had an α-cut value greater than 0.5. Based on these results specifically, the "average threshold value (d)," "average percentage of expert consensus," and "average fuzzy score" it can be concluded that all proposed components achieved expert consensus and were thus accepted for the devised framework. The highest expert consensus (91%) was observed for "Subject," which had a fuzzy score of 0.674. Conversely, "machine learning and data analysis" had the lowest fuzzy score (0.590), yet it remained acceptable. The high expert consensus percentages, predominantly above 80%, indicated strong agreement regarding the relevance of these components. Furthermore, the defuzzified values, all exceeding 0.59, suggested that these components played a crucial role in decision-making. These findings confirm that all listed components are relevant and should be considered in framework.

Table 4. Expert consensus on the components of the framework

Components	Triangular fuzzy numbers		Defuzzification value	Result
	Average threshold value (d)	Average percentage of expert consensus, (PA_c) (%)	Average fuzzy score (A)	
Technology enabler	0.174	89	0.696	Accepted
Assessment data	0.174	83	0.604	Accepted
Activity and other miscellaneous data	0.141	88	0.613	Accepted
Machine learning and data analysis	0.172	78	0.590	Accepted
Context (community)	0.162	87	0.641	Accepted
Subject	0.176	91	0.674	Accepted
Control (rules)	0.176	88	0.648	Accepted
Communication (division of labour)	0.187	86	0.648	Accepted
Output	0.159	85	0.633	Accepted

3.1. Expert consensus for the elements of the "technologies enabler" component

Table 5 presents the expert consensus for the "technology enabler" component regarding whether the system infrastructure for mobile assessment analytics to enhance students' confidence relies on or requires the use of the specified elements. All elements within the "technology enabler" component received more than 75% consensus from the experts. The highest-ranked elements were "data security" (defuzzified value: 0.743) and "availability, reliability, and integrity of data" (defuzzified value: 0.733), highlighting their significance within the "technology enabler" component. The element "cloud platform and networking" (e.g., public cloud, private cloud, and/or hybrid cloud) received 76% expert consensus, which was slightly above the 75% threshold, with an average fuzzy score of 0.629. The slightly lower percentage agreement (76%) indicated that some experts provided responses outside the agreed range, contributing to variation in consensus. The average fuzzy score (0.629) suggested that while expert opinions were generally aligned, some divergence remained, indicating moderate agreement with a degree of disagreement. These results suggest that all listed elements are relevant for the "technology enabler" component.

3.2. Expert consensus for the elements of the "assessment data" component

Table 6 presents the expert consensus on the elements of "assessment data". The correctness of the submitted answer, the number of answered/unanswered questions, the grades obtained, the number of attempts for assessment, and students' response times were accepted. These elements had high expert consensus and achieved a defuzzified score above the threshold (typically 0.5 or higher). Conversely, the number of times an answer was changed, the duration required to complete the assessment, the number of times a question was viewed, and the amount of time students remained idle (i.e., not submitting an answer) were rejected. These elements had lower consensus with the percentage of expert consensus below the acceptable threshold. The most important element, ranked first, was the correctness of the submitted answer, with a defuzzified value of 0.648. Student behavior aspects (e.g., changing answers, idle time, and assessment duration) were rejected, suggesting that they were not considered significant indicators of students' confidence in the assessment. However, the number of attempts, response time, and viewing duration were accepted, indicating that these elements contributed meaningfully to student confidence evaluation. The results of the FDM emphasized the importance of response accuracy, participation levels, and engagement over behavioral changes such as modifying answers or remaining idle. These findings suggest that the evaluation of students' confidence should focus on accuracy and active participation rather than on how often students change their responses or the time taken to complete the assessment.

Table 5. Expert consensus for the elements of "technologies enabler"

Elements	Triangular fuzzy numbers		Defuzzification value	Result	Ranking
	Average threshold value (d)	Average percentage of expert consensus (PA_j) (%)	Average fuzzy score (A)		
Data security	0.1	95	0.743	Accepted	1
Availability, reliability and integrity of data	0.1	95	0.733	Accepted	2
ICT infrastructure	0.2	95	0.724	Accepted	3
Security and integrity of the assessment	0.2	90	0.714	Accepted	4
Data protection (for example, privacy of the data)	0.2	90	0.714	Accepted	4
Data storage	0.2	95	0.686	Accepted	6
A learning management system (LMS)	0.2	81	0.667	Accepted	7
Conformity to standards for interoperability between various LMS	0.2	81	0.657	Accepted	8
Cloud platform and networking (eg. public cloud, private cloud and/or hybrid cloud)	0.2	76	0.629	Accepted	9

Table 6. Expert consensus for the elements of "assessment data"

Elements	Triangular fuzzy numbers		Defuzzification value	Result	Ranking
	Average threshold value (d)	Average percentage of expert consensus (PA_j) (%)	Average fuzzy score (A)		
Correctness of the submitted answer (correct/wrong)	0.2	90	0.648	Accepted	1
Number of answered and unanswered questions	0.2	86	0.648	Accepted	1
Grades obtained on the assessments	0.2	81	0.648	Accepted	1
Number of attempts for an assessment.	0.2	86	0.638	Accepted	4
Students' response times (based on correct response)	0.1	86	0.610	Accepted	5
Number of times the answer for a question is changed	0.2	67	0.581	Rejected	
Time spent to answer a question (the answer is saved)	0.2	81	0.571	Accepted	6
Number of correct responses that changed to wrong responses	0.2	81	0.571	Accepted	6
Number of wrong responses that changed to correct responses	0.2	81	0.571	Accepted	6
Duration to complete the assessment	0.2	67	0.571	Rejected	
Number of times the question is viewed	0.2	67	0.537	Rejected	
Duration spent viewing a question (not saving the answer)	0.2	76	0.527	Accepted	9
Amount of time the students remained idle (not submitting an answer)	0.3	62	0.514	Rejected	

3.3. Expert consensus for the elements of the "activity and other miscellaneous data" component

Table 7 presents the expert consensus on the elements of the component "activity and other miscellaneous data." The element "goal expectation" (defuzzified value: 0.648) had an average threshold value of 0.1 and received high expert consensus (95%). It was the highest-ranked accepted element, indicating strong agreement among experts on its importance. Similarly, the element "viewed lessons, chapters, or syllabus" (defuzzified value: 0.648) had the same score and ranking, suggesting equal importance in the experts' view. Slightly lower than the top-ranked elements but still considered important was the element "students' satisfaction level on the course" (defuzzified value: 0.619), which had an 86% expert consensus, reinforcing its significance. The element "activity logs" (defuzzified value: 0.581), which included metrics such as resource access and discussion participation, achieved an 81% expert consensus. While it was considered relevant, it ranked lower than "students' satisfaction level on the course" and "goal expectation". The element "recognition of affects and mood during self-assessment" (defuzzified value: 0.571) had a lower fuzzy score than other accepted elements, placing it at the lowest rank among them. However, its expert consensus was high (86%), indicating a need for consideration within the component "activity and other miscellaneous data." Conversely, the element "number of emails sent to instructor" (defuzzified value: 0.448) was rejected, as its fuzzy score (0.448) fell below the acceptable threshold. Despite a 67% expert consensus, it was deemed not significant enough for inclusion in the final evaluation. The most critical elements identified were "goal expectation" and "viewed lessons chapters, or syllabus," followed by "Students' satisfaction level on the course." Activity logs and recognition of mood were considered important but ranked lower. Email communication with the instructor was not regarded as a useful indicator. These findings aid in refining learning analytics criteria by focusing on elements that experts agreed were more relevant. Confidence appears to play a significant role in expressing one's goals and expectations, as confident individuals are more likely to communicate clearly and assertively.

Table 7. Expert consensus for the elements of "activity and other miscellaneous data"

Elements	Triangular fuzzy numbers		Defuzzification value	Result	Ranking
	Average threshold value (d)	Average percentage of expert consensus (PA_j) (%)	Average fuzzy score (A)		
Goal expectation	0.1	95	0.648	Accepted	1
Viewed lessons, chapters, or syllabus	0.2	90	0.648	Accepted	1
Students' satisfaction level on the course	0.1	86	0.619	Accepted	3
Activity logs from the following: (number of times resource accessed, date and time of access, types of recourses accessed and number of asked questions in discussion forum for activities such as chat, discussion data, temporal data, free text)	0.1	81	0.581	Accepted	4
Recognition of affects and mood during self-assessment (e.g., boredom, confusion, delight, or frustration)	0.1	86	0.571	Accepted	5
Number of emails sent to instructor	0.2	67	0.448	Rejected	

3.4. Expert consensus for the elements of the "machine learning and data analysis" component

Table 8 presents the expert consensus on the elements of "machine learning and data analysis". Out of the 10 elements within this component, two elements "process mining" (defuzzified value: 0.600) and "linguistic analysis" (defuzzified value: 0.527) failed to obtain expert consensus regarding their applicability in evaluating or determining students' confidence. The highest-ranked elements were "text mining," "association rule mining," and "affect recognition," each achieving a defuzzified score of 0.610 and an expert consensus of 81%. Lower-ranked but accepted elements included "machine learning techniques" (defuzzified value: 0.590, Rank 4), "classification techniques" (defuzzified value: 0.581, Rank 5), "descriptive statistics" (defuzzified value: 0.581, Rank 5), "inferential statistics" (defuzzified value: 0.581, Rank 5), and "speech recognition" (defuzzified value: 0.562, Rank 8). These findings indicate that text mining, association rule mining, and affect recognition are the most favorable techniques. Additionally, machine learning and statistical techniques are considered viable for evaluating students' confidence. However, process mining and linguistic analysis did not meet the required criteria for acceptance due to their percentage of expert consensus that was below the acceptable 75% threshold. The low consensus (48%) for process mining means that experts have diverse opinions some may strongly agree while others strongly disagree.

3.5. Expert consensus for the elements of the "context (community)" component

Table 9 presents the expert consensus on the "context (community)" component. Experts agreed that the "institution of higher learning (IHL)," "course director/programme director," and "employers" are key stakeholders in a system providing mobile assessment analytics to enhance students' confidence. However, the element "government/policy maker/accreditation body" failed to obtain the required consensus, as only 62% of experts supported its inclusion.

Table 8. Expert consensus for the elements of "machine learning and data analysis"

Elements	Triangular fuzzy numbers		Defuzzification value	Result	Ranking
	Average threshold value (d)	Average percentage of expert consensus (PA_j) (%)	Average fuzzy score (A)		
Text mining	0.1	81	0.610	Accepted	1
Association rule mining. [Note: A rule-based machine learning method for discovering interesting relations between variables in large databases. For example, the rule {onions, potatoes}⇒{burger} found in the sales data of a supermarket would indicate that if a customer buys onions and potatoes together, they are likely to also buy hamburger meat.]	0.2	81	0.610	Accepted	1
Affect recognition. [Note: observation of visual and auditory nonverbal cues. Nonverbal cues include facial, vocal, postural, and gestural cues.]	0.2	81	0.610	Accepted	1
Process mining	0.2	48	0.600	Rejected	
Machine learning techniques (such as particle swarm optimization-based support vector machine) or regressions technique	0.2	76	0.590	Accepted	4
Classification techniques	0.2	76	0.581	Accepted	5
Descriptive statistics [Note: statistics showing the summary of the data]	0.2	76	0.581	Accepted	5
Inferential statistics. [Note: statistics showing the conclusion and prediction.]	0.2	76	0.581	Accepted	5
Speech recognition	0.2	76	0.562	Accepted	8
Linguistic analysis	0.3	67	0.527	Rejected	

Table 9. Expert consensus for the elements of "context (community)"

Elements	Triangular fuzzy numbers		Defuzzification value	Result	Ranking
	Average threshold value (d)	Average percentage of expert consensus (PA_j) (%)	Average fuzzy score (A)		
IHL	0.2	90	0.676	Accepted	1
Course director/programme director	0.2	90	0.657	Accepted	2
Employers	0.1	81	0.590	Accepted	3
Government/policy maker/accreditation body	0.2	62	0.581	Rejected	

3.6. Expert consensus for the elements of the "subject" component

With respect to expert consensus on the services, resources, and individuals engaged in activities provided by mobile assessment analytics to enhance students' confidence, experts agreed that "teachers/instructors/lecturers" provided learning resources and e-assessment materials. Additionally, assessment analytics assisted teachers in delivering more engaging, useful, and relevant content, assessment materials, and activities. Experts also agreed that "students/learners" engaged in learning activities using various resources, including pre-recorded videos, lecture notes, useful links, and wikis. As shown in Table 10, all elements had an average threshold value of 0.2. The most important elements were "learning resources" and "students/learners," both achieving the highest defuzzified score of 0.686. "Teachers/Instructors" and "Services such as e-assessment and assessment analytics" were also considered important but ranked slightly lower. All elements were accepted, indicating their relevance in the study.

3.7. Expert consensus for the elements of the "control (rules)" component

Consensus was obtained among experts regarding the controls and rules governing the implementation of mobile assessment analytics to enhance students' confidence. Table 11 indicates a lack of agreement on the elements "policy," and "legislation and law". While experts may have partially agreed on certain aspects, "policy" received only 52% agreement, reflecting diverse expert responses. The fuzzy score value of 0.6 for "policy" suggests that while responses were not widely scattered, they were not strongly unified either. This implies a level of common understanding with some variation in expert opinions. For "legislation and law," a 71% agreement was observed, which is close to the 75% threshold, indicating that a majority of experts provided similar responses. However, since this value remained below 75%, some variation in expert opinions persisted. The fuzzy score value of 0.543 further suggests that expert responses were somewhat dispersed, though not entirely misaligned. This indicates that while general agreement existed, significant differences in individual opinions lowered the overall consensus score. Some experts may have strongly opposed certain aspects, reducing the overall consensus. Consequently, these two elements were removed from the framework. Conversely, “pedagogy” (defuzzified value: 0.705) and “user readiness/knowledge/willingness to use technology” (defuzzified value: 0.705) ranked highest, indicating their critical importance within the given context. “Goal setting” (defuzzified value: 0.667), “action” (defuzzified value: 0.657), and “grounding/boundary” (defuzzified value: 0.648) followed closely, emphasizing the significance of goal definition, action-taking, and boundary-setting in the analysed domain. “Curriculum structure” (defuzzified value: 0.648) and “skill sets by industries” (defuzzified value: 0.638) were accepted but ranked lower, indicating moderate importance. “Culture” received the lowest accepted score (0.543), suggesting that while it met the threshold, it was not considered highly critical. Overall, pedagogical and user readiness were regarded as the most crucial elements. Policy and legislation did not receive sufficient expert support, implying that they might not be as relevant within this specific educational and technological context. Additionally, goal setting, actions, and boundary definitions were deemed important considerations.

Table 10. Expert consensus for the elements of "subject"

Elements	Triangular fuzzy numbers		Defuzzification value	Result	Ranking
	Average threshold value (d)	Average percentage of expert consensus (PA_j) (%)	Average fuzzy score (A)		
Learning resources such as (pre-recorded videos, lecture notes, exercises, useful links, wiki)	0.2	95	0.686	Accepted	1
Students/learners	0.2	90	0.686	Accepted	1
Teachers/instructors/lecturers	0.2	90	0.676	Accepted	3
Services such as e-assessment and assessment analytics	0.2	90	0.648	Accepted	4

Table 11. Expert consensus for the elements of “control (rules)”

Elements	Triangular fuzzy numbers	Defuzzification value		Result	Ranking
	Average threshold value (d)	Average percentage of expert consensus (PA_j) (%)	Average fuzzy score (A)		
Pedagogy	0.2	95	0.705	Accepted	1
User readiness/knowledge/willingness to use the technology	0.2	95	0.705	Accepted	1
"Goal setting. [Note: students consider what they want to achieve and setting one or more goals accordingly.]"	0.2	95	0.667	Accepted	3
"Action. [Note: actions that can be taken by the various stakeholders (eg. policy makers, administrators, lecturers, students)]"	0.2	90	0.657	Accepted	4
"Grounding / Boundary. [Note: the scope covered. For example, the scope of information that the analytics will provide and how it relates to the educational context.]"	0.2	90	0.648	Accepted	5
Curriculum structure	0.2	86	0.648	Accepted	5
Skills sets by industries	0.2	86	0.638	Accepted	7
Data privacy/security	0.2	81	0.619	Accepted	8
Policy	0.1	52	0.600	Rejected	
Legislation and law	0.2	71	0.543	Rejected	
Culture	0.2	76	0.543	Accepted	9

3.8. Expert consensus for the “communication (division of labour)” component

Table 12 presents the expert consensus on the elements of the “communication (division of labour)” component concerning the communication medium among various stakeholders. Experts agreed that a system providing mobile assessment analytics to enhance students' confidence must facilitate communication among stakeholders, enabling forums, collaboration, and teamwork. Both elements within the “communication (division of labour)” component received high agreement among experts. “Collaboration and teamwork" achieved the highest defuzzified score (0.667), ranking first, followed by "communication and dialogue" with a score of 0.629. These results suggest that "collaboration and teamwork" is the most critical element, followed by "communication and dialogue." The high level of expert consensus validates their importance in the system.

Table 12. Expert consensus for the elements of "communication (division of labour)"

Elements	Triangular fuzzy numbers		Defuzzification value	Result	Ranking
	Average threshold value (d)	Average percentage of expert consensus (PA_j) (%)	Average fuzzy score (A)		
Collaboration and teamwork	0.2	90	0.667	Accepted	1
Communication and dialog (such as forum)	0.2	81	0.629	Accepted	2

3.9. Expert consensus for the elements of the "output" component

Table 13 presents the expert consensus on the required outputs of a system implementing mobile assessment analytics to enhance students' confidence. Experts agreed that the system should provide "achievements of learning outcomes," "feedback," and "visualization dashboard/UI/UX" as essential outputs. The system must enable learners to select the next task, either through interventions or recommendations. Additionally, users should be able to personalize the system based on their preferences and needs while reflecting on their learning progress. The system should also support interventions for remedial actions; for instance, subsequent assessment tasks should be based on a student's achievements in previous tasks. However, the "participation rates" element failed to obtain the required expert consensus and was consequently removed from the devised framework.

Among the accepted elements, "achievements of learning outcomes" received the highest ranking (defuzzified value: 0.695), indicating it was the most widely accepted component. "Feedback" (defuzzified value: 0.686) and "visualization dashboard/UI/UX" (defuzzified value: 0.657) followed, ranking second and third, respectively. "Selection of next task" and "Intervention" received a defuzzified value of 0.590 but were still accepted.

In contrast, "participation rates" was rejected due to low expert consensus (62%), which was significantly below the threshold. Despite having the same fuzzy score (0.590) as some accepted elements, it did not meet the required agreement level among experts. The decision-making process prioritized learning outcomes, feedback, and UI/UX elements while excluding participation rates due to insufficient consensus.

Table 13. Expert consensus for the elements of "output"

Items	Triangular fuzzy numbers		Defuzzification value	Result	Ranking
	Average threshold value (d)	Average percentage of expert consensus (PA_j) (%)	Average fuzzy score (a)		
Achievements of learning outcomes	0.2	90	0.695	Accepted	1
Feedback	0.2	95	0.686	Accepted	2
Visualization dashboard/UI/UX	0.2	90	0.657	Accepted	3
Personalization	0.2	76	0.619	Accepted	4
Recommendations	0.2	86	0.619	Accepted	4
Reflection	0.1	86	0.610	Accepted	6
Intervention	0.1	81	0.590	Accepted	7
Selection of next task	0.2	76	0.590	Accepted	7
Participation rates	0.2	62	0.590	Rejected	

4. DISCUSSION

The experts reached a consensus on all components of the mobile learning (m-learning) framework aimed at enhancing students' confidence. These components enable students to engage in active learning within a mobile learning environment. This approach aligns with Engeström's extended activity system, wherein learners interact with mobile technologies, learning materials, and communities while adhering to institutional rules and divisions of labour. By analyzing data from mobile interactions, institutions can optimize learning processes, personalize education, and address contradictions within the system. Several studies have applied Engeström's extended activity system in the design, implementation, and evaluation of m-learning platforms, providing a comprehensive understanding of the interactions among learners, tools, and communities [36], [37]. For instance, the Stanford mobile inquiry-based learning environment (SMILE) utilizes mobile devices to facilitate active learning by enabling students to create, share, and evaluate questions. Similarly, mobile computer-supported collaborative learning (mCSCL) environments employ mobile devices to support real-time collaborative learning activities. These platforms reflect the principles of Engeström's system by incorporating community participation and division of labour into the learning process. Furthermore, frameworks such as challenge-based learning (CBL) have been developed to engage students in real-world problem-solving through collaboration and hands-on activities. CBL emphasizes the interactions among learners, instructors, and communities, aligning with the key components of Engeström's extended activity system.

Data security received the highest level of consensus among experts for the "technology enabler" component. This finding aligns with the work of Papamitsiou and Economides [18], which examines the role of event records and logs as data types in various assessment analytics scenarios. For instance, logs that monitor and analyze students' step-by-step problem-solving processes require a robust ICT infrastructure to facilitate data storage, retrieval, and utilization. This infrastructure must ensure high availability and reliability, with security being a critical consideration. Although experts did not strongly agree on the importance of adhering to interoperability standards across different LMS, ensuring interoperability remains essential. Standardized data formats are crucial for interoperability and learning analytics. xAPI, widely used in LMS, supports interoperability and assessment tracking, playing a key role in data standardization [25]. While cloud computing including public, private, and hybrid cloud models is a well-established technology, professionals' familiarity and expertise in these domains may vary. Hybrid cloud solutions, which integrate on-premises infrastructure with public and private clouds, are evolving rapidly. Given the complexity of managing these hybrid environments, some IT professionals may still be in the process of adapting to them.

The analysis of assessment data plays a vital role in understanding learners' behavior, identifying challenges, and enhancing the design and implementation of assessments. The time a student spends viewing a question without submitting a response may indicate disengagement or difficulty with the content. By detecting periods of inactivity, educators can optimize assessments to improve student engagement. Furthermore, response times offer valuable insights into students' cognitive processes, learning behaviors, and overall performance, helping to determine their level of comprehension. Analyzing response times enables the provision of personalized feedback, thereby supporting students in improving their learning outcomes [18], [25]. The data elements from the "activity and other miscellaneous data" component have the potential to contribute to the development of personalized learning and adaptive teaching methodologies, aligning with the perspectives of Wise *et al.* [21]. Enhancing students' learning experiences necessitates a thorough understanding of their behavior and engagement within mobile learning (m-learning) environments [17].

Machine learning techniques, such as text mining, play a crucial role in identifying learners' skills and knowledge. Additionally, classification techniques have proven effective in determining text similarity in short-answer responses. This aligns with Hamiz *et al.* [22] proposed assessment analytics framework, which emphasizes continuous learning improvement and aligns with the broader concept of leveraging learning analytics for skill identification. Similarly, Papamitsiou and Economides [18] introduced an assessment analytics framework (AAF) designed to monitor and enhance student progress. The AAF comprises four primary components: input, process, output, and feedback. The input phase includes contextual information relevant to tracking and assessing learners. The process phase involves data analysis and interpretation using various methods and algorithms, including linguistic analysis, text mining, speech recognition, classification techniques, and machine learning. Findings indicate that text mining, association rule mining, and affect recognition are among the most favorable techniques, highlighting their applicability in assessing students' confidence levels by analyzing their social interactions, forum discussions, goal-setting behaviors, and assessment responses.

The framework proposed by West *et al.* [20] underscores the significance of institutional context, transitional components, infrastructure, retention strategies, learning analytics for student persistence, and intervention and reflection. Additionally, it facilitates discussions among various stakeholders within institutions, such as academic staff, IT departments, and senior executives, to ensure that learning analytics initiatives are effectively integrated and aligned with institutional goals. In line with West *et al.*, "IHL," "course director/programme director," and "employers" were found to be the key stakeholders in this study. When formulating policies related to education, training, or workforce development, the roles of IHLs, course or programme directors, and employers are inherently interconnected. IHLs are responsible for establishing overarching policies that regulate curriculum development, ensure quality assurance, and maintain compliance with accreditation standards. Furthermore, IHLs must align their programs with national and international workforce demands. Research institutes within IHLs contribute by fostering research, securing funding, and building partnerships to enhance learning outcomes and employability. The course or programme director plays a pivotal role in designing and implementing course content that meets both academic and industry standards. Effective collaboration among faculty, students, and employers is essential to ensure curricula remain relevant and applicable. One approach to achieving this is the incorporation of practical experiences, such as internships, apprenticeships, and industry-driven projects, within academic programs. Employers contribute by offering insights into industry trends, identifying skill gaps, and outlining workforce expectations. Additionally, companies can create internship and placement opportunities, providing students with hands-on exposure to real-world work environments and requirements. Employers also serve as mentors, equipping interns with practical knowledge and industry-specific competencies. Furthermore, organizations can collaborate with IHLs and course directors to co-develop programs that produce graduates who are well-prepared for the job market.

In contrast to previous studies [19], [25], which suggested that leveraging "duration to complete the assessment" could assist educators in making informed decisions to enhance student learning experiences and outcomes, the present study found that experts rejected this metric. The time spent on an assessment is influenced by various factors beyond student confidence, making it an unreliable indicator of certainty. Some students naturally complete assessments quickly, while others take additional time to carefully process questions, independent of their confidence levels. A confident student may spend more time reviewing answers, whereas an uncertain student might complete the assessment hastily without thorough revision. Additionally, external influences such as distractions, fatigue, anxiety, or technical difficulties can affect the time taken without necessarily reflecting confidence. Overconfident students may rush and make errors, while underconfident yet knowledgeable students might take longer to ensure accuracy.

Although the number of times a question is viewed can provide insights into students' engagement and interaction with the material, it does not serve as a reliable indicator of student confidence for several reasons. Students may revisit a question out of curiosity rather than uncertainty. Likewise, confident students may review questions to ensure accuracy or reinforce their understanding. Additionally, students might return to a question to clarify wording or verify their interpretation, independent of their confidence level. Some students use revisiting as a revision strategy, while others might do so after discussing the question with peers. Conversely, avoiding question revisits does not necessarily indicate confidence, as it may stem from overconfidence, misinterpretation, or time constraints. Due to these variations in student behavior, experts have dismissed the number of times a question is viewed as a dependable metric for assessing confidence. Similarly, the duration of idle time before submitting an answer does not reliably reflect confidence. Extended idle times may indicate deep cognitive processing rather than hesitation or uncertainty. Conversely, quick submissions do not necessarily signify confidence, as they may result from guessing or impulsive answering. Other factors, such as slow internet connections, technical issues, or unfamiliarity with the testing platform, can also contribute to delays. In collaborative environments, students may intentionally delay submission to align with their peers. Therefore, idle time alone is insufficient for measuring confidence, as it is influenced by multiple external and cognitive factors. Nevertheless, research suggests that idle time can serve as a predictor of performance, as it may reflect the level of cognitive engagement in problem-solving tasks [26].

The element “number of times the answer for a question is changed” has also been rejected by experts because the number of times a student changes their answer does not necessarily indicate their confidence level. A student might change their answer upon realizing they initially misunderstood the question, which does not necessarily reflect confidence but rather clarity. Measuring confidence purely based on the number of times an answer is changed can be misleading and does not account for the complexities of decision-making in test-taking.

This study plays a crucial role in shaping a structured and evidence-based approach to enhancing students' confidence, a key factor in academic success and personal development. By obtaining and validating expert opinions, the research ensures that the proposed framework is both comprehensive and grounded in professional insights. Confidence is closely linked to students’ motivation, resilience, and overall performance, making it essential to provide educators with effective strategies to nurture this attribute. Furthermore, the study identifies specific elements for evaluating students’ confidence, ensuring that progress can be systematically measured and improved. A well-defined evaluation mechanism enables targeted interventions, helping educators and institutions refine their approaches to confidence-building. By addressing this critical aspect of student development, the study contributes significantly to fostering a supportive learning environment where students feel empowered to overcome challenges and achieve their full potential. A potential area for future research is examining the relationship between students’ behavior during assessments and their feelings toward them, as well as how these feelings influence their confidence. Additionally, the reciprocal impact of students' confidence on course activities and vice versa can be explored.

The effectiveness of the implementation of the model must also take into consideration which generation the students belong to. For example, “Baby Boomers” were born roughly between the years 1946 to 1964, “Generation X” were born roughly between the years 1965 to 1980, “Millennials” were born roughly between the years 1981 to 1996, and “Generation Z” were born roughly between the years 1997 to 2012. Since the implementation of the proposed framework will mostly involve “Generation Z” who are actively engaged in technology, the proposed framework needs to address the m-learning teaching and learning strategies for “Generation Z” mentioned in [38] such as “incorporation of social and community learning,” “ensuring internet coverage availability,” “the need for the applications to be mobile responsive,” “the use of visual learning elements,” “entertaining educational content,” and “providing online discussion”.

Based on the findings of this study, the following recommendations aim to enhance students' confidence through the integration of a mobile learning (m-learning) framework. These recommendations focus on two key aspects: (i) incorporating m-learning into the curriculum through a well-structured LMS by applying the devised framework and (ii) integrating a confidence-based intervention model using machine learning. These strategies enhance student engagement, streamline stakeholder interaction, and provide data-driven insights for continuous improvement.

IHLs should consider embedding a course within the devised mobile learning (m-learning) framework to boost students' confidence. Implementing this framework requires developing an LMS that leverages various technological components, including ICT infrastructure and cloud-based solutions such as virtual machines, databases, queues, load balancers, and content delivery networks (CDNs). The LMS must comply with interoperability standards to facilitate seamless integration with existing platforms and support different user roles, including administrators, educators, students, and industry professionals.

Security and privacy considerations are critical due to the substantial volume of personal, learning, and assessment data involved. The LMS should feature user-friendly interfaces, dashboards, and analytics to support various stakeholders. Educators should be able to create and manage diverse assessment formats, including multiple-choice, true/false, matching, short-answer, and essay-based questions. Additionally, students should have the ability to set learning objectives, track progress, and provide feedback on their course experiences. The platform must also support collaboration through interactive features such as chat and discussion forums. Finally, comprehensive dashboards should be integrated to visualize analytical insights, monitor student performance, and evaluate learning outcomes.

A machine learning-based confidence model should be implemented to assess and predict students' confidence levels. The computed confidence scores can serve as real-time intervention feedback, influencing students' engagement, learning behaviors, and course participation. By incorporating this model into the LMS, the system can dynamically adapt to students' needs, fostering a more supportive learning environment and improving overall confidence levels.

5. CONCLUSION

This study demonstrates the effective use of data from assessments and course activities to provide targeted interventions and meaningful feedback for students, educators, higher education providers, employers, and other relevant stakeholders. By leveraging a data-driven approach, the framework ensures that insights are derived from objective measures rather than relying solely on self-reported confidence levels, which are often prone to bias and inaccuracies. A key challenge in education is the discrepancy between a student's perceived confidence and their actual competence. Traditional self-evaluation methods may not always provide a reliable representation of a student’s understanding, leading to either overconfidence or self-doubt. To address this issue, the proposed framework utilizes real-time analytics to track student progress, identify learning gaps, and offer personalized feedback based on performance data.

Through the implementation of an analytic dashboard, students receive timely, actionable insights that guide their learning journey by highlighting strengths and areas requiring improvement. This empowers learners to take a proactive approach in refining their skills and knowledge. Additionally, educators can utilize these insights to adapt teaching strategies, refine course materials, and create a more responsive and dynamic learning environment. Beyond individual learning improvements, this framework benefits a broader ecosystem, including higher education institutions and employers. By harnessing data-driven feedback, universities can enhance curriculum design, align educational outcomes with industry needs, and better prepare graduates for workforce demands. Employers, in turn, gain access to more accurate indicators of student capabilities, facilitating informed recruitment and workforce development decisions. Ultimately, this study underscores the transformative potential of data analytics in education. By replacing subjective self-assessments with objective, data-backed evaluations, the framework fosters a more accurate, efficient, and adaptive learning experience that benefits all stakeholders involved.

FUNDING INFORMATION

The authors wish to extend their gratitude to Universiti Pendidikan Sultan Idris, Tanjung Malim, Malaysia for assisting this research with the GPU grant (code: 2022-0033-107-01).

AUTHOR CONTRIBUTIONS STATEMENT

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Teik Heng Sun	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	
Muhammad Modi Lakulu	✓	✓	✓	✓			✓	✓		✓	✓	✓	✓	✓
Noor Anida Zaria Mohd Noor	✓			✓						✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author, Muhammad Modi Lakulu. The data, which contain information that could compromise the privacy of research participants, are not publicly available due to certain restrictions.

REFERENCES

[1] Z. Oktafiani and Y. Yusri, “The relationship of self confidence to students learning achievement,” *Counseling and Humanities Review*, vol. 1, no. 1, pp. 20–26, Apr. 2021, doi: 10.24036/000411chr2021.
[2] L. Stankov, J. Lee, W. Luo, and D. J. Hogan, “Confidence: a better predictor of academic achievement than self-efficacy, self-concept and anxiety?,” *Learning and Individual Differences*, vol. 22, no. 6, pp. 747–758, Dec. 2012, doi: 10.1016/j.lindif.2012.05.013.
[3] J. S. Shrauger and M. Schohn, “Self-confidence in college students: conceptualization, measurement, and behavioral implications,” *Assessment*, vol. 2, no. 3, pp. 255–278, Sep. 1995, doi: 10.1177/1073191195002003006.
[4] R. Agnihotri, “Manual for agnihotri’s self-confidence inventory (ASCI),” *Psychological Corporation, Agra*, pp. 2–3, 1987.
[5] D. D. Pandey, “Manual of self-confidence inventory for school students (PSCI),” *Agra: National Psychological Corporation*, 2007.
[6] T. Gardner-Medwin and N. Curtin, “Certainty-based marking (CBM) for reflective learning and proper knowledge assessment,” *Certainty-based marking (CBM) for reflective learning and proper knowledge assessment*, pp. 1-7, Jan. 2007.
[7] G. Schraw and R. S. Dennison, “Assessing metacognitive awareness,” *Contemporary Educational Psychology*, vol. 19, no. 4, pp. 460–475, Oct. 1994, doi: 10.1006/ceps.1994.1033.
[8] Y. Engeström, *Learning by expanding: an activity-theoretical approach to developmental research*, 2nd ed. Cambridge University Press, 2014. doi: 10.1017/CBO9781139814744.
[9] T. Burner and B. Svendsen, “Activity theory—lev vygotsky, aleksei leont’ev, yrjö engeström,” 2020, pp. 311–322, doi: 10.1007/978-3-030-43620-9_21.
[10] F. Shamsuddin and A. Z. Abdul Razak, “Expert consensus on the usability of the SLC3DM model: a fuzzy Delphi method application,” *Management Research Journal*, no. 13, pp. 45–55, Mar. 2024, doi: 10.37134/mrj.vol13.1.4.2024.
[11] A. H. Masnan, N. K. Ismail, W. Usop, and D. M. Mat Zain, “Predicting preschool moral development with school climate: Fuzzy Delphi approach,” *International Journal of Evaluation and Research in Education (IJERE)*, vol. 13, no. 3, p. 1537, Jun. 2024, doi: 10.11591/ijere.v13i3.27777.
[12] A. F. Mohamed Yusoff, A. Hashim, N. Muhamad, and W. N. Wan Hamat, “Application of fuzzy delphi technique towards designing and developing the elements for the e-PBM PI-Poli module,” *Asian Journal of University Education*, vol. 17, no. 1, p. 292, Mar. 2021, doi: 10.24191/ajue.v17i1.12625.
[13] A. F. Roldán López de Hierro, M. Sánchez, D. Puente-Fernández, R. Montoya-Juárez, and C. Roldán, “A fuzzy delphi consensus methodology based on a fuzzy ranking,” *Mathematics*, vol. 9, no. 18, p. 2323, Sep. 2021, doi: 10.3390/math9182323.
[14] N. D. Azeez and M. M. Lakulu, “Evaluation framework of mgovernment services success in Malaysia,” *Journal of Theoretical and Applied Information Technology*, vol. 96, no. 24, pp. 8194–8226, 2018.
[15] N. M. Husain, M. M. Lakulu, and S. Sarkawi, “Modelling of pck competency of programming teachers through fuzzy delphi method (Fdm),” May 2018, pp. 346–357, doi: 10.15405/epsbs.2018.05.28.
[16] A. Ishikawa, M. Amagasa, T. Shiga, G. Tomizawa, R. Tatsuta, and H. Mieno, “The max-min Delphi method and fuzzy Delphi method via fuzzy integration,” *Fuzzy Sets and Systems*, vol. 55, no. 3, pp. 241–253, May 1993, doi: 10.1016/0165-0114(93)90251-C.
[17] B. Fernández-Gallego, M. Lama, J. C. Vidal, and M. Mucientes, “Learning analytics framework for educational virtual worlds,” *Procedia Comput Science*, vol. 25, pp. 443–447, 2013, doi: 10.1016/j.procs.2013.11.056.
[18] Z. Papamitsiou and A. A. Economides, “An assessment analytics framework (AAF) for enhancing students’ progress,” in *Formative Assessment, Learning Data Analytics and Gamification*, Elsevier, 2016, pp. 117–133, doi: 10.1016/B978-0-12-803637-2.00007-5.

[19] S. Yassine, S. Kadry, and M.-A. Sicilia, "A framework for learning analytics in moodle for assessing course outcomes," in *2016 IEEE Global Engineering Education Conference (EDUCON)*, IEEE, Apr. 2016, pp. 261–266. doi: 10.1109/EDUCON.2016.7474563.
[20] D. West, D. Heath, and H. Huijser, "Let's talk learning analytics: a framework for implementation in relation to student retention," *Online Learning*, vol. 20, no. 2, Dec. 2015, doi: 10.24059/olj.v20i2.792.
[21] A. F. Wise, J. M. Vytasek, S. Hausknecht, and Y. Zhao, "Developing learning analytics design knowledge in the 'middle space': the student tuning model and align design framework for learning analytics use," *Online Learning*, vol. 20, no. 2, Jan. 2016, doi: 10.24059/olj.v20i2.783.
[22] M. Hamiz, M. Bakri, N. Kamaruddin, and A. Mohamed, "Assessment analytic theoretical framework based on learners' continuous learning improvement," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 11, no. 2, p. 682, Aug. 2018, doi: 10.11591/ijeecs.v11.i2.pp682-687.
[23] A. Tlili, F. Essalmi, M. Jemni, Kinshuk, and N.-S. Chen, "A complete validated learning analytics framework," *International Journal of Information and Communication Technology Education*, vol. 14, no. 2, pp. 1–16, Apr. 2018, doi: 10.4018/IJICTE.2018040101.
[24] M. Bikanga Ada, "Using design-based research to develop a mobile learning framework for assessment feedback," *Research and Practice in Technology Enhanced Learning*, vol. 13, no. 1, p. 3, Dec. 2018, doi: 10.1186/s41039-018-0070-3.
[25] A. Nouira, L. Cheniti-Belcadhi, and R. Braham, "An ontology-based framework of assessment analytics for massive learning," *Computer Applications in Engineering Education*, vol. 27, no. 6, pp. 1343–1360, Nov. 2019, doi: 10.1002/cae.22155.
[26] Z. K. Papamitsiou, V. Terzis, and A. A. Economides, "Temporal learning analytics for computer based testing," in *ACM International Conference Proceeding Series*, New York, NY, USA: ACM, Mar. 2014, pp. 31–35. doi: 10.1145/2567574.2567609.
[27] B. Ludwig, "Predicting the future: Have you considered using the Delphi methodology?," *The Journal of Extension*, vol. 35, no. 5, p. 15, 1997.
[28] S. A. Abdullah *et al.*, "Determining elements in mobile learning implemention among instructors in vocational colleges: a fuzzy Delphi method," *IEEE Access*, vol. 9, pp. 150839–150845, 2021, doi: 10.1109/ACCESS.2021.3121703.
[29] N. Yusof, N. L. Hashim, and A. Hussain, "A review of fuzzy Delphi method application in human-computer interaction studies," *In AIP Conference proceedings*, 2022, p. 040026. doi: 10.1063/5.0094417.
[30] Lei Chang, "A psychometric evaluation of 4-point and 6-point likert-type scales in relation to reliability and validity," *Applied Psychological Measurement*, vol. 18, no. 3, pp. 205–215, Sep. 1994, doi: 10.1177/014662169401800302.
[31] F. N. Pazilah, H. Hashim, and M. M. Yunus, "Refining the TPACK framework: a fuzzy Delphi approach to 21st-century competency and self-efficacy constructs," *Cogent Education*, vol. 11, no. 1, Dec. 2024, doi: 10.1080/2331186X.2024.2428886.
[32] M. Lawnik and A. Banasik, "Delphi method supported by forecasting software," *Information*, vol. 11, no. 2, p. 65, Jan. 2020, doi: 10.3390/info11020065.
[33] S. Ahmad, M. M. Ali, and K. Salehuddin, "The content validation of Malaysian dyslexia accommodating screening test (MYDAST) through the application of fuzzy Delphi method (FDM)," *International Journal of Academic Research in Progressive Education and Development*, vol. 11, no. 2, Apr. 2022, doi: 10.6007/IJARPED/v11-i2/13161.
[34] H. Chu and G. Hwang, "A Delphi-based approach to developing expert systems with the cooperation of multiple experts," *Expert Systems with Applications*, vol. 34, no. 4, pp. 2826–2840, May 2008, doi: 10.1016/j.eswa.2007.05.034.
[35] J. W. Murry and J. O. Hammons, "Delphi: A versatile methodology for conducting qualitative research," *Review of Higher Education*, vol. 18, no. 4, pp. 423–436, 1995, doi: 10.1353/rhe.1995.0008.
[36] M. I. Santally, D. Cooshna-Naik, and N. Conruyt, "A model for the transformation of the Mauritian classroom based on the Living Lab concept," in *2014 IST-Africa Conference Proceedings*, IEEE, May 2014, pp. 1–10, doi: 10.1109/ISTAFRICA.2014.6880603.
[37] L. Uden, "Activity theory for designing mobile learning," *International Journal of Mobile Learning and Organisation*, vol. 1, no. 1, p. 81, 2007, doi: 10.1504/IJMLO.2007.011190.
[38] L. L. D. Mandau and M. M. Lakulu, "A critical review on m-learning for generation Z," *Journal of ICT in Education*, vol. 9, no. 3, pp. 111–128, 2022, doi: 10.37134/jictie.vol9.sp.1.11.2022.

APPENDIX

Table 1. Identified elements and the related literature

Component	Elements	Related literature
Technology Enabler	ICT infrastructure	[17], [18], [20], [24]
	Data/database/other data sources	[17], [18], [19], [22], [23], [25], [26]
	Data availability, reliability integrity	[20], [23]
	Ease of use and interoperability between various LMS	[17], [20], [23], [25]
	Learning management systems	[17], [19], [25], [26]
Assessment data	Amount of time the students remained idle (not submitting an answer)	[18], [25], [26]
	Students' response times (based on correct response)	[18], [25], [26]
	Duration to complete the assessment	[19], [25]
	Number of attempts for an assessment	[25]
	Number of answered and unanswered questions	[25]
	Number of questions answered correctly, and number of questions answered wrongly	[25], [26]
	Number of times the question is viewed	[26]
	Number of times the answer to a question is changed	[26]
	Duration spent viewing a question (not saving the answer)	[26]
	Time spent to answer a question (the answer is saved)	[26]
	Grades obtained on the assessments	[19], [25]

Table 1. Identified elements and the related literature *(Continue…)*

Component	Elements	Related literature
Activity and miscellaneous data	Viewed lesson, chapter or syllabus	[17], [19]
	Recognition of affects and mood during self-assessment (eg, boredom, confusion, delight, or frustration)	[18]
	Activity logs (number of times resource accessed, date and time of access, types of recourses accessed and number of asked questions in discussion forum for activities such as chat, discussion data, temporal data, free text)	[17], [18], [19], [21]
	Students satisfaction level	[18], [21]
	Goal expectation	[26]
	Number of emails sent to instructor	[19]
Machine learning and data analysis	Linguistic analysis	[18]
	Text mining	[18], [22]
	Speech recognition	[18]
	Classification techniques	[18]
	Association rule mining	[18]
	Process mining	[18]
	Machine learning techniques (such as particle swarm optimization-based support vector machine) or regressions technique	[18], [22]
	Affect recognition	[18]
	Statistics	[20], [22]
Context (community)	Government/policy maker/accreditation body	[18], [19], [20], [22], [23], [24]
	Institution of higher learning (IHL)	[18], [20], [22], [23]
	Course director/programme director	[20], [22], [23]
	Employers	[17], [19], [20], [22], [23], [24]
Subject	Students/learners	[17], [18], [19], [20], [21], [22], [23], [24], [26]
	Teachers/instructors/lecturers	[17], [18], [19], [20], [21], [22], [23], [24]
	Services such as e-assessment and assessment analytics	[17], [18], [19], [22], [23], [25]
	Learning resources such as (pre-recorded videos, lecture notes, exercises, useful links, wiki)	[17], [18], [19], [21], [22]
Control (Rules)	Grounding/boundary	[20], [21], [23], [24]
	Goal setting	[19], [20], [21], [23], [24]
	Action	[17], [21], [22], [25]
	Pedagogy	[18], [20], [21], [24]
	Curriculum structure	[18], [22], [24]
	Skill sets by industries	[18], [20], [22]
	Data privacy/security	[23]
	Legislation and law	[23]
	Culture	[24]
	Policy	[24]
	User readiness/knowledge/willingness to use the technology	[23], [24]
Communication (Division of labour)	Communication and dialog	[18], [23], [24]
	Collaboration and teamwork	[18], [24]
Output	Achievements of learning outcomes	[19], [26]
	Feedback	[18], [20], [21], [24]
	Visualization dashboard / UI/UX	[17], [18], [19], [20], [23]
	Intervention	[17], [18], [20], [21], [22], [23], [24]
	Selection of next task.	[18]
	Personalization	[23], [24]
	Recommendations	[23]
	Participation rates	[18]
	Reflection	[20], [21]

BIOGRAPHIES OF AUTHORS

Teik Heng Sun is an Assistant Professor in the Faculty of Information and Communication Technology at Universiti Tunku Abdul Rahman, Malaysia and is currently doing his Ph.D. study at the Faculty of Computing and Meta-Technology at the Sultan Idris Education University, Malaysia. He received his M.Sc. in telecommunication and information systems from the University of Essex, United Kingdom in 1997 and his BS (computer science) from Campbell University, United States in 1995. His current research interests include educational and web technologies. He can be contacted at email: thsun@utar.edu.my.

Muhammad Modi Lakulu is an Associate Professor in the Faculty of Computing and Meta-Technology at the Sultan Idris Education University, Malaysia. From 2013-2019, he was the Head of the Computing Department and from 2019-2021, he was Deputy Dean (Research and Innovation) and currently he is the Director of the Quality Management Centre at Sultan Idris Education University. Moreover, He received his Ph.D. degree in computer science (knowledge management) from the Universiti Putra Malaysia (UPM), in 2012, his M.Sc. in software engineering from the University of Bradford, UK in 2002 and his B.Sc. in computer science from Universiti Teknologi Malaysia, in 1998. His research focuses on educational technology, information systems, and AI. His research works have been published in journals, books, and conferences. He can be contacted at email: modi@meta.upsi.edu.my.

Ts. Dr. Noor Anida Zaria Mohd Noor is a senior lecturer in the Faculty of Computing and Meta-Technology at Sultan Idris Education University, Malaysia. She served as Deputy Director at Academic Development Centre from 2023 to 2024and as Deputy Director at the Entrepreneurship Development and Graduate Employability Centre from 2020 to 2022 at Sultan Idris Education University. She received her Ph.D. degree in Knowledge Integration in 2012, M.Sc. (Information Technology) in 2004 and Bachelor of Information Technology (Hons) in 2001 from Universiti Utara Malaysia. Her current research interests are related to Knowledge Integration, education, and Information Systems. She can be contacted at email: anidazaria@meta.upsi.edu.my.