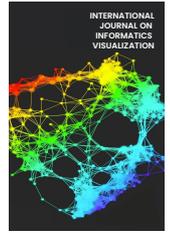




INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv



Lecopelese - a Novel Evaluation Model for Measuring Educational Aspects of Game-based Learning

Hanif Al Fatta^{a,*}, Zulisman Maksom^b, Mohd Hafiz Zakaria^b

^a Faculty of Computer Science, Universitas AMIKOM Yogyakarta, Sleman, Yogyakarta, Indonesia

^b Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya Durian Tunggal, Malaysia

Corresponding author: *hanif.a@amikom.ac.id

Abstract—This study aimed to establish a model for assessing the pedagogical quality of mobile game-based learning (GBL), which seeks to convey educational content to users. Evaluating the educational effectiveness of GBL necessitates a robust model tailored for this purpose. Current models can be improved to better address various educational challenges associated with mobile GBL. The LECOPELESE (LEarning Content, PEdagogogy and LEarning StyLE) model was developed by integrating relevant constructs identified in existing literature. To validate this model, a qualitative research approach was employed, drawing a sample from 270 undergraduate students. The analysis utilized Structural Equation Modeling (SEM) and resulted in a final model based on rigorous factor analysis. The findings indicated that the proposed model effectively measures educational quality in game-based learning. This new model includes more comprehensive constructs and items, addressing the educational aspects of game-based learning. Specifically, the model introduces a pedagogy construct to evaluate game-based learning quality, reflecting criteria for outstanding educational content and delivery through mobile applications. It assesses how effectively GBL provides real-world learning experiences. Additionally, the research highlights that the quality of pedagogy is influenced by two key factors: the GBL's ability to accommodate learners' unique characteristics (learning styles) and the quality of the learning content that adapts to learners' needs. Ultimately, the study demonstrates that both learning content and style significantly impact the pedagogy construct, suggesting that enhancing these areas can improve the overall pedagogical quality of game-based learning.

Keywords— Evaluation; learning style; learning content; usability; metrics; GQM.

Manuscript received 18 Sep. 2024; revised 8 Feb. 2025; accepted 9 Mar. 2025. Date of publication 30 Sep. 2025.
International Journal on Informatics Visualization is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



I. INTRODUCTION

Educational games or serious games has been widely used as a learning media in many areas of modern life. Educational games can be utilized to teach formal knowledge [1], health simulation and learning tools for person with disabilities [2], [3], [4] as well as training [5], [6]. Quality improvement is always a priority for application developers to meet the needs of their users. This trend also occurred in the development of educational game applications. One method that can be chosen to measure the quality of product applications is by conducting a usability evaluation [7], [8]. Usability Evaluation is widely used to measure the user experience quality of a game-based learning, serious games or educational games [9],[10], [11],[12].

However, determining suitable dimensions to measure the user experience of the educational games continues to be a challenge for researcher in usability evaluation topic. Several

models to assess usability evaluation has been proposed by several researchers [13], [14], [15], [16].

However, determining the correct dimensions to measure aspects of education from educational games continues to be a challenge for researchers in this field. This study aims to examine several dimensions of educational measurement models that have been carried out and compile a new model to enhance the existing models. Some measurement dimensions are proposed to address more usability issues related to the pedagogical aspect.

Research on measuring educational elements is relatively limited. Research by [17], proposed a dimension to measure educational element called pedagogy dimension. The literature review also revealed that research focused on learning or pedagogical dimensions is relatively less compared to research on other dimensions. This research will put more emphasis on the measurement of this pedagogical aspect of mobile game-based learning (m-GBL).

Several previous types of research have already discussed the significance of measuring learning or pedagogical aspects of m-GBL. Even though m-GBL is a game, it has special purposes, delivering learning content. Because of that, the evaluation of the learning dimension would be critical to improving the quality of learning content delivery. This study identified several papers that proposed methods, models, or frameworks to measure pedagogical aspects of m-GBL.

MEGUX, a UX model proposed by [16], declared that to measure the user experience quality of an educational games, six key components can be measured: Flow, Immersion, Game Usability, Player Context, Mobility, and Learning. Earlier research, proposed E-GUESS [18] introduced a measurement model to handle bipolarity between fun and educational, by accommodating “pedagogical issues and educational content” into the final usability aspect in E-GUESS model. In addition, research by [19] utilized Computer System Usability Questionnaire (IBM-CSUQ), focusing on three key dimensions: interface quality (INTERQUAL), information quality (INFOQUAL), and system usability (SYSUSE) to measure educational quality providing significant findings that the model can be used to measure educational experience and cognitive load. Furthermore, EDUGXQ model proposed by [20] declaring a User Experience (UX) framework for educational games (EDUGX) based on UX elements. This model proposed six UX elements: Flow, Immersion, Player Context, Game Usability, Game System and adding Learnability as a dimension to measure educational aspect of the educational games.

Research conducted by [21] suggested a framework for game-based learning comprising four critical aspects of a serious game: context, learner specification, representation, and pedagogical approach used. This framework is known as the four-dimensional framework. The context of the game and its utilization are significant to the practical usage of a serious game. This contextual factor may consist of the location where the game is employed inside a classroom or outside, the availability of technical support, and environmental support to guarantee a conducive learning environment. This aspect significantly influences the creation of a game, determining the most suitable game type.

The second aspect focused on the attributes of a specific learner or learner group. For example, factors such as age, level of expertise in using the game, learning background, learning style, and how they learned. Another aspect is the internal representation world, which emphasizes presentation mode, degree of immersion, interactivity, and fidelity within serious games or simulations. Ultimately, the pedagogical approach use aspect underlines the significance of learning theory, model, and framework. Moreover, these aspects also suggest the learning content is embedded and personalized regarding different learning approaches, and the game will take place.

Another interesting research is conducted by [22], which suggests that there is a tight relationship between learning objectives and the selection of appropriate game types. For example, if the learning outcome that will be achieved is memorization or retention, according to this study, the suitable game type is a puzzle or a drill and practice quiz. Besides, if the learning outcome is motoric, suitable game

types include shoot 'em up/beat 'em up games, fighting games, or driving/flight simulator games.

In addition, [23] proposed a framework based on the enhancement of a four-dimensional framework. This framework consists of 5 dimensions as follows: learner performance, motivation, perceptions, attitudes, and preferences. Moreover, for evaluation purposes, [23] suggests a GBL environment category along with its measurement criteria: environment, scaffolding, usability, level of social presence, and deployment. While [23] claimed that the framework is designed to be focused on the pedagogy of GBL.

Furthermore, [24] suggested learning content measurement for m-GBL comprising four components: ease of the learning content, learning content should be provided, achievement of learning outcome is assured, and the user understands the content delivered via m-GBL.

Some research [25], [26], [27], [28], particularly described learning style as one of the pedagogical aspects worth measuring. The research by [25] stated that learning style is recognized as an essential human factor affecting students' learning performance. Previous studies have shown that, by taking learning styles into account, learning systems can be of more significant benefit to students owing to the provision of personalized learning content presentation that matches the information-perceiving and processing styles of individuals. This research also proposed a personalized game-based learning approach based on the sequential/global dimension of the learning style by Felder and Silverman. Moreover, the results show that the personalized educational computer game not only promotes learning motivation but also improves the learning achievements of the students.

Another research by [29] stated the relationship between Game Genres, Learning Techniques, and Learning Styles in Educational Computer Games. Furthermore, [29] argued that developing an effective educational computer game, different game genres, learning activities and techniques, and learning styles are important issues for consideration. Rapeepisarn [29] suggests that the learning style of the learner will affect the choice of their behavior while playing a game as well as using the computer. Learning style influences the most suitable learning activities and leads to a preference for a game type that suits the learner.

In the research proposed by [26], learning dimensions also become an essential part of the m-GBL evaluation. [10] suggested that learning dimensions should be a specialty in games with educational purposes. It comprises (1) Learner: the consideration of age, level of the learner, how they learn, learning style, background, and their learning preference. (2) Curriculum: An m-GBL should contain subject content, the scope of learning defined by the curriculum, and teacher customization allowing the adaptation of learners' needs. (3) Feedback: it is used to inform the gap between the current knowledge the learner acquired and the mandatory learning content they should master. (4) Context: The specific area where the learning occurred should be an essential factor that supports the learner, or on the other hand, hinders the learning. (5) Direct experience: m-GBL should deliver the direct experience that engages and motivates players during the learning process.

II. MATERIALS AND METHOD

During the development of the model to evaluate pedagogical aspects, two essential activities are defined: (1) model construction and (2) Model Testing, as seen in Fig. 1.

A. Stage 1: Model Development

To ensure the evaluation model's acceptability in terms of research methodology, this step defines methods to generate a new model, building on the method proposed by [30] and augmented by [31]. The model construction can be divided into three steps:

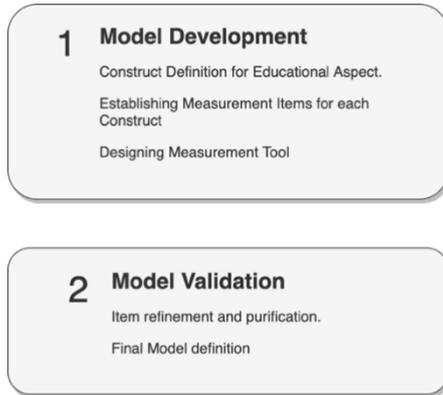


Fig. 1 Method for generating an indicator for measuring educational aspects of m-GBL

1) *Constructs definition for educational aspects:* During the process of proposing a new model, this research argued that a suitable model for evaluating educational aspects of m-GBL should comprise measurements on some constructs or latent variables. During the literature review, the suitable constructs for measuring educational aspects of m-GBL are extracted and analyzed.

2) *Establishing measurement items for each construct:* After the process of defining the constructs is completed, the next activities are to define suitable items/heuristics according to possible constructs. Each construct can have a different number of criteria, based on supporting literature. At the end of these activities, every construct is equipped with its suitable indicator/heuristics. The first part of the new model is ready.

3) *Designing the Measurement Tools:* This research employed the GQM model [32], [33], [34] as a reference for designing measurement metrics/questionnaires based on the dimensions obtained from stage A.1 and A.2. Based on the literature, the most critical constructs along with their criteria are selected. This list of criteria will not provide many benefits if it lacks the mechanism to measure them. Therefore, it is necessary to provide measurement metrics for each dimension, including all its criteria. If this activity is completed, the first version of the measurement metrics/questionnaire is ready to validate.

B. Stage 2: Model Validation

Following literature from [35], validation for proposed criteria for usability evaluation can be done in two consecutive steps: (1) item refinement and purification: in this round all items for initial version of the model is refined by analyzing its reliability and validity. This round may result in

deletion of several items that did not meet the reliability and validity requirements. (2) Final model definition: final items that fulfilled the reliability and validity test will be defined as the final items for the proposed model.

According to [17] suitable dimensions to measure usability for m-GBL are mobility, playability, interface usability, and pedagogy. The proposed model will explore the pedagogical aspects of m-GBL and provide measurement metrics for the usability evaluation system. This section will focus on the pedagogy dimension. The pedagogy dimension is regarded as a unique dimension since it will measure the most significant objective of an educational game, which is delivering educational content. The proposed model comprised two main aspects of educational usability measurement comprising Learning Style and Learning Content. Each aspect will be divided into several items as described in **Error! Reference source not found.** The rationale of each construct is also described in table 1. After defining the aspect and its related items, the next step is to create attributes/items for each construct. In the next section, all constructs for the proposed model are described as follows:

1) Pedagogy construct:

Pedagogy is a significant aspect to measure, as proposed by [36]. Pedagogy constructs represent the degree to which an application contains educational content and performs appropriate tasks to deliver educational content by following the concept of learning in the real world. The essence of creating GBL is delivering learning content. According to [23], pedagogy reflected the criteria of excellence in learning content as well as learning delivery formatted in a mobile application. This construct measures how well a GBL can present a real-world learning experience within the game. As stated by [23], in real-world learning, the learning style of the learner should be accommodated like the GBL and should be evaluated during the assessment of GBL quality. And according to [37], learning content is also a significant aspect to measure during GBL. This proposed model employed the learning style and learning content in the pedagogy construct.

TABLE I
RATIONALES FOR PEDAGOGICAL ASPECTS DEFINITION

Construct and Rationale	Literature Resources
Pedagogy Construct: The essence of creating GBL is delivering learning content. This construct reflected the criteria of excellent learning content as well as learning deliverance formatted in a mobile application. This construct measure how well a GBL is able to present real world learning experience within the game.	[23]
Learning Style Construct: Learning style represent the ability of the game-based learning to accommodate the special characteristic of their learner. Learning style constructs is proposed as a Hierarchical Component Model served as higher order component. The lower order component consist of several constructs as follows:	
Learner: The personal characteristics of learner comprising: age, particular methods how they learn, learning background, learning style and learning preferences influenced the learning process using GBL.	[21] [26] [38]
Preferences: During GBL intervention, learner and instructor preference should be considered.	[23] [36]

Construct and Rationale	Literature Resources
According to [17], there are various learning styles; therefore, different learner will need different preference, for example, preferences on educational medium. This construct also deals with personalized learning, where personalized educational games that met the learning style of learner significantly benefited the students compared with the educational games that did not meet their learning style.	[25] [29]
Learning Content Construct: Learning style construct represent the degree of the learning content quality and how flexible learning content serve their variative learner. Learning content served as higher order construct consist of some lower order constructs as follows:	
Content Quality: To ensure the quality of the learning content, m-GBL should meet several requirements for their content:	[37]
<ol style="list-style-type: none"> 1. The learning content should be available within the game. 2. Learning the content should be an easy process. 3. The learning objective of the game is achieved. 4. The content is understandable. 	
System Adaptation: Learning deliverance could be accommodated into the gameplay in several ways. First, the mode that customizes the game according to the student's learning style is identified by using a learning style questionnaire, and a mode that has an in-game adaptive system that dynamically and continuously adapts its content according to the student's interactions in the game.	[23]

2) Learning style construct:

The second aspect of the pedagogic dimension is Learning style. Learning style is divided into three attributes: Learner, Preferences, and Personalized Learning. This attribute is proposed to facilitate the development of items for learning style. The description and items for each attribute are presented below:

• Learner

According to [21], a significant aspect of evaluating m-GBL is learner specification. The successful game should consider several key factors, including demographics, personal knowledge of Information and Communication Technology, and users' gaming experience. An m-GBL should integrate those issues into its game development to enhance the efficiency of delivering learning content. Research by [26], support the use of learner dimensions and propose that, on the demographic item, age is a factor that should be considered during the development of game-based learning. Very similar aspects are also proposed by [38], suggesting three constructs for game-based learning evaluation: pre-game, in-game, and post-game constructs. It defined socio-demographics, including sex, nationality, age, previous experience, and skills associated with computer, game, and digital learning environments, as attributes for measurement.

• Preferences

According to [23], an evaluation framework for active GBL should consider seven categories of evaluations. They are Learner Performance, Perceptions, Preferences, GBL environment, motivation, attitude, and collaboration.

Particularly in the preference category, [23] proposed that the personal preference of learners has an impact on the effectiveness of learning, for example, the preference for the educational medium used to deliver educational content. Thus, if game-based learning is used, it should be checked whether this medium is suitable for the preferences of the learner. A study by [39] proposed the significance of personal preference, where a user can choose the difficulty level provided by the game setting. A good game requires providing this option to make sure the user has the degree of freedom to play based on their level of understanding.

Preference also covers a personalized game-based learning package. Personalized educational games that met the learning style of the learners significantly benefited the students compared to the educational games that did not meet their learning style. Research conducted by [25] supports this theory. [25] developed personalized game-based learning that suits the learner's learning style according to the learning style by Felder and Silverman. The research results showed that personalized games increased learning motivation and improved learner achievement. [29] supported the previous study by suggesting the relationship between Game Genres, Learning Techniques, and Learning Styles. Research by [29] has shown that a learner's learning style, as defined by Honey and Mumford, influences the most suitable game genre for that individual. For example, an "Activist" learner is more suited to playing multiplayer, action, or role-playing games than adventure or simulation games, which are more suitable for a "Reflector" student. From the results of this research, it is essential to design and select the genre that matches the learner's learning style to give maximum advantages during learning activities [see Table 2 for the references]. In addition, [39] explained that a good educational game should be equipped with self-learning process support. By implementing this self-learning support, users can learn the game's content independently.

TABLE II
POSSIBLE GAME GENRE-BASED ON USER'S LEARNING STYLE BASED ON [29]

Learning Style	Possible Game Genre
Activists	Multiplayer interaction, action game, role-playing game
Reflectors	Concentration game, adventure game, a simulation game
Theorists	Strategies game, a simulation game Puzzle
Pragmatists	Puzzle game, building game, constructing game, reality testing game, detective game

3) Learning Content Construct

The third aspect of the pedagogic dimension is Learning Content. This aspect deals with the quality of the learning content to meet the primary objective of game-based learning. Learning content aspect is divided into two attributes: Content Quality and Game Genre. The description of each attribute, along with its items, is described below:

• Content Quality

According to [37], the learning content aspect may consist of 4 items, they are : (1) the availability of learning content in the game, to fulfill the definition of game-based learning, it a must that intended learning content is available within the

game-based learning, (2) Learning the content should be an easy process, for its intended users, learning the content using the game should be an easy activity, it means the utilization of game-based learning provide more opportunity for the intended user to learn easily, (3) The learning objective from the game is achieved, utilization of game-based learning should not change the learning objective that should be accomplished as the learning took place in the real world, and finally (4) The content is understandable, the process of formatting learning content into game-based learning form, should guarantee that the content is still understandable to its intended user.

- *System Adaptation*

According to [28], to accommodate the learner's learning style, two approaches can be taken: adaptability and adaptivity. In the Adaptability mode, the game-based learning from the beginning of the development is explicitly customized according to a particular selected learning style [40]. Therefore, the game-based learning in this adaptability mode is equipped with an option that allows the user to customize the game according to their preference. On the contrary, adaptivity leads the game-based learning into the concept of a user adaptive system, that, through the characteristic of interaction performed by the user, the system can detect and define the learning style and set the system preference to suit that user preference, which means the user implicitly influences the system preference [41]. Finally, game genre also plays an essential role in giving user adaptation flexibility. According to [42], to deliver a particular game objective, it is necessary to select an appropriate game genre. For example, if the learning objective is about decision-making skills, then a role-playing game or a strategy game genre is suitable. On the other hand, if the learning objective is classified as memory or repetition skill, the puzzle or quiz games are appropriate [see Table 3 for reference].

TABLE III
THE RELATIONSHIP BETWEEN LEARNING OBJECTIVES AND GAMES
GENRE/TYOLOGY [22]

Learning Objective	Appropriate games/typology	Number of players
Memory/ repetition/ retention	Drill and practice, Quiz games, Puzzle games	One vs. PC
Dexterity/ spread/ precision/ motoric	Shoot'em up Beat'em up (Combat/fight games) Driving games Simulation games	One or one vs. one
Applying concepts/rules	Sport games Action games Driving games Drill and practice	One vs. PC
Decision making (strategy + problem solving)	Strategic games Adventure Games Role play games Simulation Games	One in connection with other players/ opponents
Social interaction/ values Cultures	Strategic games Role play games Simulation Games	many players
Ability to learn/self-assessment	Role-play games Simulation Games	one

After completing the items for each construct as depicted in Table 1, the next step is creating the measurement metrics. To create the measurement metric easily, the GQM model is

utilized. GQM is the goal-driven method to develop and maintain useful metrics, comprising three parts: Goals, Questions, and Metrics [32]. The initial step from GQM is to identify the goals of the measurement. All attributes shown in Table 1 will be the goals of our GQM model. From there, the question is carefully constructed, and it is necessary to ensure that all questions are measurable. Furthermore, metrics containing the essential information to answer the question is described. Following the work presented by [43], the attribute from Table 1 serves as quality characteristics, and from there, the goal and guideline can be constructed as seen in Table 4.

TABLE IV
COMPLETE GQM MODEL TO EVALUATE EDUCATIONAL DIMENSION

Quality Characteristic	Goal	Guidelines
Pedagogy	educational content availability and appropriate content delivery	It is easy to learn using this game, based on your level of ICT skill. Enjoy learning with a game-based learning format. It is easy to learn using this game, based on your level of ICT skill.
Learning Style		
Learner	Age compatibility	The game-based learning is developed for a particular age group. The game-based learning is developed with consideration of this age group's characteristics
	Familiarity with ICT	The game-based learning accommodates the level of ICT familiarity of the particular user.
	Gaming Experience	The game-based learning accommodates the level of gaming experience of the particular user
Preference	Desirable medium for delivering educational content	Game-based learning is a desirable medium based on the user's choice.
	Game-based eLearning meets user preferences in choosing the learning media	User is comfortable with the game-based learning as a learning medium
	Level of difficulty selection	The game-based learning provides an option for difficulty level selection.
	Game genre compatibility with the learner's learning style	The genre is chosen according to the user's learning style based on tables from [14]
	the game provides facilities to learn educational content independently.	The game-based learning enables users to learn the content independently when tutors are unavailable.
Learning Content		
Content Quality	Availability	The learning content should be available in the game
	Easiness	Learning through the game should be an easy process

Quality Characteristic	Goal	Guidelines
	Achievability	It is possible to achieve the learning objective by employing the game
	Understandability	The learning content is still understandable, though it was converted into the game format
System Adaptation	Adaptability	The game shows some level of customization according to the appropriate learner's learning style
	Adaptivity	The game is capable of detecting user's learning style and automatically change the game setting into a suitable learning style
	Game Genre	game genre chosen according to the appropriate learning objective based on tables from.

The guideline from Table 4 is significant in constructing the question measuring each attribute. Complete question for all measurement items is presented in Table V below:

TABLE V
LIST OF QUESTIONS BASED ON PROPOSED MODEL

Pedagogy Constructs	
Code	Question detail
ped_1	Can you understand the learning content easily through the games?
ped_2	Are you enjoying learning with a game-based learning format?
ped_3	Is it easy to learn using this game based on your level of ICT skill?
Learning Style Constructs	
Learner	
learn_1	Is there any information about the user's age restriction?
learn_2	Does the game accommodate the "look" and "feel" of the game based on its users' age category?
learn_3	Is the low ICT skill sufficient to play the game?
learn_4	Is the low Gaming experience sufficient to play the game?
Preferences	
pref_1	Do users prefer game-based learning as a medium for delivering teaching material?
pref_2	Does the game-based learning medium convince the users to deliver teaching material?
pref_3	Does the game provide the option for the user to select the level of difficulty?
pref_4	Is the game genre chosen according to the user's learning style based on tables from [14], for example: a) Multiplayer interaction, action game, and role-playing game for an active learner. b) Concentration game, adventure game, a simulation game for reflective learners c) Strategies game, a simulation game, a Puzzle for the Theorist Learner d) Puzzle game, building game, constructing game, reality testing game, detective game for a pragmatist learner
pref_5	Does game-based learning enable users to learn content independently when tutors are unavailable?
Learning Content Construct	
Content quality	

cont_1	Is the educational content available in the game?
cont_2	Is it easy to learn educational content through the game?
cont_3	Does the user achieve the learning objective by playing the game?
cont_4	Does the user understand the learning content formatted in game-based learning?
System Adaptation	
adapt_1	Does the game offer customization options to align its behavior with the user's learning style?
adapt_2	Is the game capable of detecting the user's learning style and automatically adjusting the game settings to match it?
adapt_3	Is the game genre chosen according to the appropriate learning objective based on tables from Dondi and Moretti, as: a) Shoot them up games, beat them up games, combat/fighting games b) Driving games, Simulation games to train motoric/dexterity/speed/decision skills. c) Sport Games, driving games, drill and practice games to train or learn to apply concepts/rules. d) Strategic games, adventure games, RPG (Role Playing Games), Simulation Games to train or learn Decision Making (Strategy and Problem solving) Skill.

III. RESULTS AND DISCUSSION

A. Construct of the Proposed Model

During the construction of the proposed model, the learning style construct and the learning content construct employed the Hierarchical Component Model (HCM) concept, in which they served as higher-order components (HOC) summarizing their respective lower-order component (LOC) as depicted in Figure 2. In this case, this research employed a hybrid approach by [44] and [45] to determine an indicator for the HOC. The hybrid approach suggests randomly splitting the lower-order indicators in half, using them as items or indicators for the higher-order constructs.

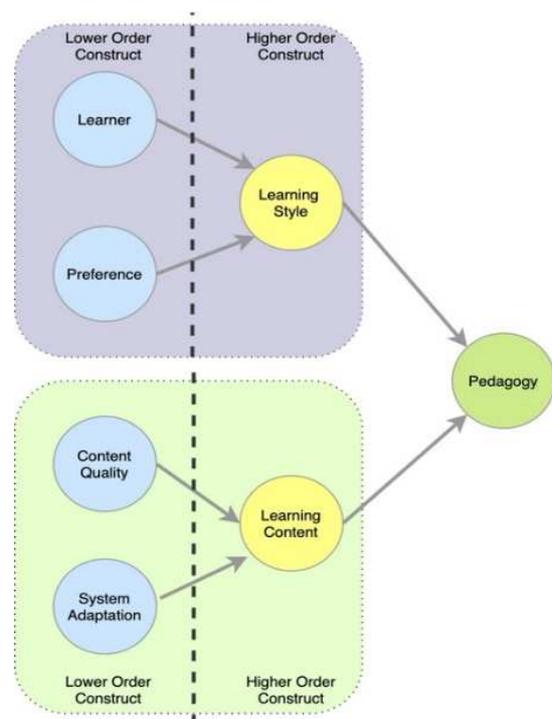


Fig. 2 Constructs for the proposed model

Following the Hybrid Approach, the items from the Lower Order Construct are split in half and added as items for the higher-order constructs. Hence, the Learning Style construct took learn_2, learn_3, pref_4, pref_5 as its items, while the Learning Content construct took cont_1, cont_4, adapt_3. The complete model showing all indicators for every construct is depicted in Figure 4. From this model, six hypotheses can be drawn:

- H1: Learner positively affects Learning Style
- H2: Preference positively affects Learning Style
- H3: Content Quality positively affects Learning Content
- H4: System Adaptation positively affects Learning Content
- H5: Learning style positively affects usability on the pedagogy aspect.
- H6: Learning Content positively affects usability on the pedagogy aspect.

Before sending the questionnaire to the subjects, a scenario was created to get the best result:

- To deal with order bias, two versions of the questionnaire were made with different random orders of 19 items.
- To tackle the reliability, boost from unnatural high correlations where participants answered consecutive questions using adjustment or anchoring, some items are displayed separately from their correlating constructs.

- And for the measurement scoring, a five-point Likert scale (1= strongly disagree, 2=disagree, 3= neutral, 4=agree, 5=strongly agree) is employed

B. Subjects of Questionnaire

This measurement instrument was purified by analyzing its validity and reliability. Following [46], the initial version contains 16 questions and a rule of thumb for factor analysis requires a minimum observation of five times the number of question items in the questionnaire, so that the number of subjects needed is at least 80 participants. Data were collected from 270 undergraduate students from the Faculty of Computer. The undergraduate students are selected for subjects for two main reasons: they are familiar with game-based learning and have insight into pedagogy concepts. The demography of the participants is depicted in Figure 3. All participant is instructed to play the selected game-based learning named “Toon Math: Endless Run and Math Games”, an endless running game to learn basic operations in math, installed on their phone, and then started to complete the questionnaire.



Fig. 3 Distribution of participants based on gender and age

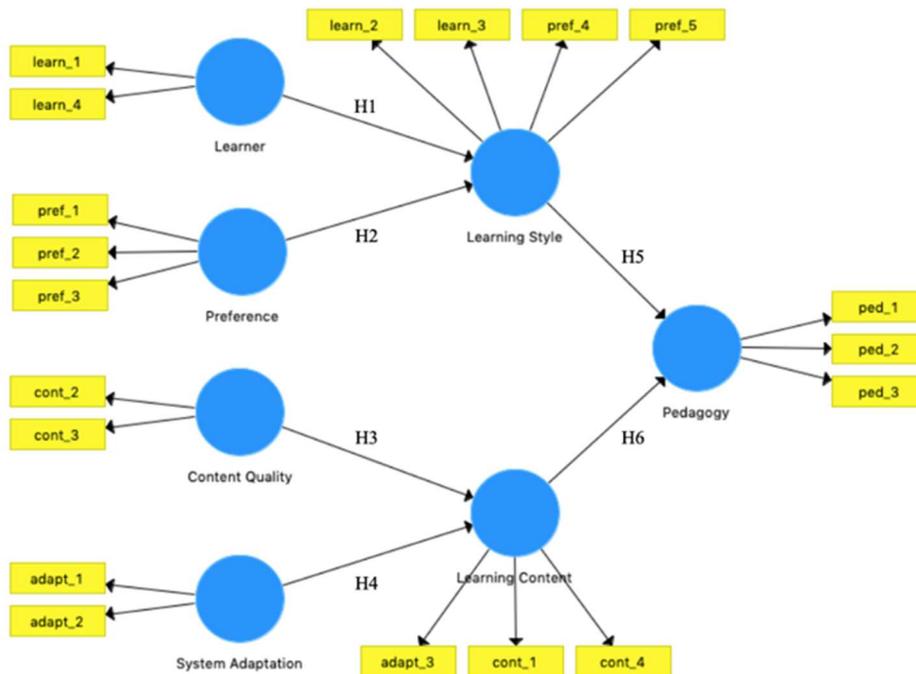


Fig. 4 Initial version of LECOPELESE model

To conduct model validation, Confirmatory Factor Analysis (CFA) is utilized to calculate the construct validity of the questionnaire. CFA indicated how well the underlying indicators explained the related construct. It means whenever the correlation of the items is relatively high, it can be concluded that the construct validity is assured [46].

1) *Face validity*: According to [47], every item used to measure proposed constructs is required to show the ability to measure the underlying construct, which is called face validity. This research provided validity by ensuring the subjects chosen were familiar with game-based learning and knowledgeable about the learning approach. The selected

subjects are students in a computer science program, familiar with different kinds of learning media, and have been exposed to pedagogy insights from their curriculum.

2) *Item validity*: This research utilized factor loading as an item validity indicator. Factor loading is an indicator showing how much particular items weigh on underlying constructs. By multiplying the factor loading times the observed variable score, the factor loading can serve as a validity coefficient, indicating how much of the score variance of observed constructs is valid [48]. According to [46], standard loading calculation should be above 0.50 and

ideally is 0.70 or higher. This high loading would indicate that the items converge on the underlying constructs and show high convergent validity. Ideally, items showing a loading value of less than 0.7 would be deleted. This resulted in three items: learn_2, pref_4, and adapt_3 being deleted. However, there are some items with loading values less than 0.7 that are not deleted. These items (marked with *) were retained because the Composite Reliability (CR) and AVE (Average Variance Extracted) values for the construct of these items are high enough for the underlying items (CR > 0.7 and AVE > 0.5) [34], as seen in Table 8.

TABLE VI
INDICATORS WITH ACCEPTED LOADING VALUES

	Content Quality	Learner	Learning Content	Learning Style	Pedagogy	Preference	System Adaptation
adapt_1							0.962
adapt_2							0.554*
cont_1			0.789				
cont_2	0.826						
cont_3	0.812						
cont_4			0.841				
learn_1		0.944					
learn_3				0.943			
learn_4		0.76					
ped_1					0.714		
ped_2					0.785		
ped_3					0.621*		
pref_1						0.838	
pref_2						0.852	
pref_3						0.57	
pref_5				0.598*			

3) *Composite reliability*: This research used Composite Reliability (CR) to determine internal consistency reliability of the constructs' items. CR value is an indicator that shows how well items/indicators describe their underlying constructs. And according to [48], the threshold value for this CR is 0.7, and the value above 0.7 is required to show how a construct is well described by its respective indicators. This research did not utilize Cronbach's alpha for two particular reasons: (1) CR does not assume that all indicator loadings are equal in the population. It is by the PLS-SEM algorithm principle, where indicators are prioritized based on their reliability during the model estimation. (2) Cronbach's alpha is sensitive to the number of items and generally underestimates internal consistency reliability [46],[48]. As presented in Table 8, the composite reliability for all observed constructs is assured by the value of the composite reliability greater than 0.7.

TABLE VII
COMPOSITE RELIABILITY AND AVE FOR ALL OBSERVED CONSTRUCTS

	Composite Reliability	Average Variance Extracted (AVE)
Content Quality	0.803	0.67
Learner	0.845	0.734
Learning Content	0.799	0.665
Learning Style*	0.759	0.623
Pedagogy*	0.751	0.504
Preference	0.804	0.584
System Adaptation*	0.75	0.616

4) *Convergent validity*

Another measurement can be calculated to measure how closely two constructs are related to each other. This measurement showed how two constructs relate or converge. This is called convergent validity and is established by comparing the observed score of one indicator in a particular construct with that of other indicators within the same construct. Convergent validity would reflect the high correlation between values of these indicators [46].

To confirm convergent validity, the Average Variance Extracted (AVE) score is established. [46] defined AVE as: "the average squared completely standardized factor loading or average communality" and is calculated for all latent variables in the measurement model. And an AVE greater than or equal to 0.5 is acceptable since a value less than 0.5 indicates more error remains, on average, in the items than the variance explained by the factor structure that was measured. This research calculated the AVE value as seen in Table 8, and the result showed that all constructs performed with an AVE value greater than 0.5, indicating all indicators for the underlying constructs are highly correlated or converged to their respective constructs.

5) *Discriminant validity*

On the other hand, discriminant validity showed opposite measurement. Discriminant validity indicates how a construction discriminates or differentiates from other constructs that it is not supposed to correlate with. Discriminant validity is defined by showing that items or indicators of a construct perform low correlation with or are "dissimilar" from other constructs [47].

TABLE VIII
DISCRIMINANT VALIDITY BASED ON THE FORNELL-LARCKER CRITERION FOR ALL OBSERVED CONSTRUCT

	Content Quality	Learner	Learning Content	Learning Style	Pedagogy	Preferences	System Adaptation
Content Quality	0.819						
Learner	0.31	0.857					
Learning Content	0.566	0.304	0.815				
Learning Style	0.384	0.857	0.354	0.789			
Pedagogy	0.575	0.591	0.912	0.65	0.71		
Preference	0.325	0.259	0.445	0.299	0.472	0.764	
System Adaptation	0.256	0.114	0.294	0.177	0.299	0.246	0.785

Usually, convergent validity and discriminant validity are assessed jointly for a set of related constructs. Discriminant validity was evaluated according to the Fornell-Larcker Criterion. It is at a satisfactory level if the square root of the AVE from the constructs is greater than the correlation shared between the construct and other constructs in the model [48].

Table 6 indicates that in all observed constructs, discriminant validity is assured and recognized.

C. Analysis

From the previous sections, it is shown that the model has met the reliability and validity requirements. The next step is to test the proposed hypothesis, as follows:

TABLE VII
PATH COEFFICIENT FOR THE PROPOSED MODEL

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (O/STDEV)	P Values
Content Quality -> Learning Content	0.525	0.528	0.051	10.346	0
Learner -> Learning Style	0.836	0.839	0.045	18.672	0
Learning Content -> Pedagogy	0.779	0.78	0.048	16.156	0
Learning Style -> Pedagogy	0.375	0.372	0.046	8.111	0
Preference -> Learning Style	0.083	0.084	0.054	1.52	0.129
System Adaptation -> Learning Content	0.16	0.164	0.057	2.795	0.005

1) *Path Coefficient*: By conducting a bootstrap calculation on smartPLS, the path coefficient for the proposed model is shown in Table 7. From this table, it is demonstrated that respective path coefficient values for learning content → pedagogy and learning style → pedagogy are 0.77 and 0.37, and all constructs showed significance, indicated by their P values. A low significant value is demonstrated by the preference → learning style relationship, with a 0.083 path coefficient value and a 0.129 P value (above 0.05). See Figure 5 for a complete path coefficient representation.

while Learning Content obtains weaker results. Since the model uses multiple independent variables, it is necessary to evaluate the value of Adjusted R Square. The consecutive values for Learning Content, Learning Style, and Pedagogy are 0.339, 0.739, and 0.953. The values indicate a strong emphasis on Learning Style and Pedagogy; however, they show a weak/low value on Learning Content.

TABLE VIII
COLLINEARITY AMONG VARIABLES

	Content Quality	Learner	Learning Content	Learning Style	Pedagogy	Preference	System Adaptation
Content Quality							
Learner			1.07				
Learning Content				1.072			
Learning Style					1.143		
Pedagogy						1.143	
Preference							1.072
System Adaptation							

2) *Variance Inflation Factor (VIF)*: To identify multicollinearity issues within the proposed model, the VIF value is calculated. As shown in Table 8, all values presented are below 5, indicating that all values are accepted (no collinearity issues found). As shown in Table 9, R-Square (coefficient of determination) for the relevant construct is presented; consecutively, the values for Learning Content, Learning Style, and Pedagogy are 0.344, 0.741, and 0.941. Strong results are found on Learning Style and Pedagogy,

TABLE IX
COEFFICIENT OF DETERMINATION

	R Square	R Square Adjusted
Learning Content	0.344	0.339
Learning Style	0.741	0.739
Pedagogy	0.954	0.953

3) *Final model*: By analyzing the path coefficient and P values of the proposed model, as seen in Table 7 and Figure 6, the conclusion regarding the hypothesis can be drawn as follows:

The hypotheses showing the high impacts are:

- H1: Learner positively affects Learning Style
- H3: Content Quality positively affects Learning Content
- H6: Learning Content positively affects usability on the pedagogy aspect.
- While lower impacts are found on the following hypotheses:
- H4: System Adaptation positively affects Learning Content.
- H5: Learning style positively affects usability on the pedagogy aspect.

And finally, with 5% confidence level, H2 showed no significance, indicated by its P values and T-statistics, so H2 is rejected. This research utilized data from university students. To collect more precise data, this research suggests

that future studies can be conducted by collecting data from several respondent groups: schoolteachers, parents, and students who are the direct users of the selected game-based learning. By completing the subject from those particular groups, the validation model may result in a more accurate statistical analysis. Furthermore, the game used in data collection is a game with an endless run mode and delivers math for elementary students. It only has simple display and

gameplay. More complex game-based learning can be used to measure more usability issues and produce more in-depth and complex data analysis. And finally, since there are no exact rules on how many indicators should appear on each construct, indicators for the proposed model vary from only two items to four items. The development of the next model can add up to five indicators in each construct, to create more flexibility during the validity and reliability testing process.

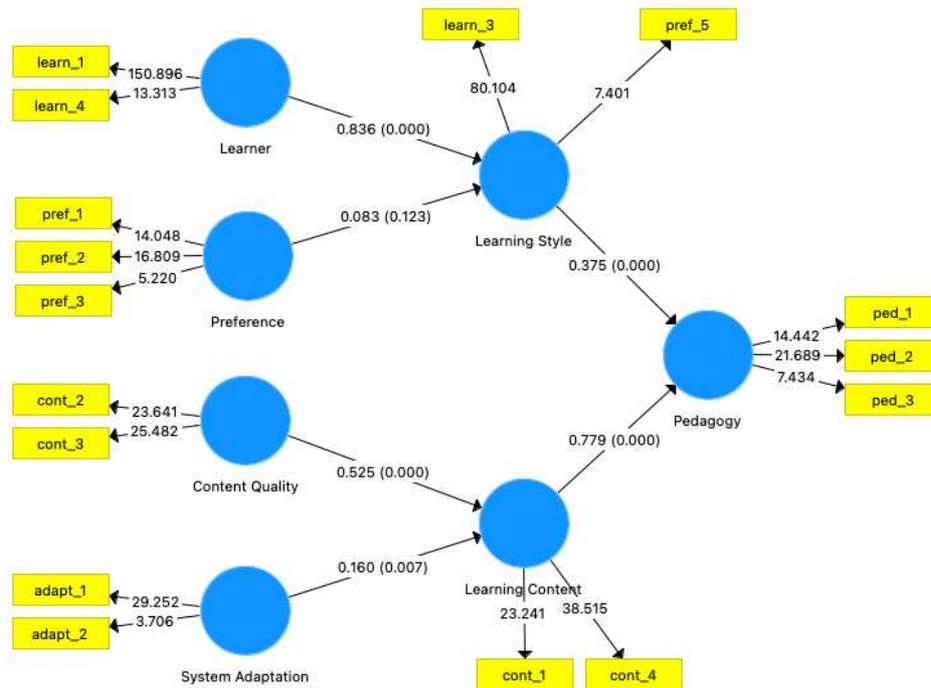


Fig. 5 Path Coefficient and P-Value for the proposed Model

Figure 5 depicts the final version of the lecopelese model. This figure showed two values for connecting constructs. The first value indicates the path coefficient, and the second value indicates the p-value. For example, the values for learning content to pedagogy and learning style to pedagogy are 0.779 (0.000) and 0.375 (0.000), respectively. The table indicates that the path coefficients for learning content to pedagogy and learning style to pedagogy are 0.77 and 0.37, respectively, with all constructs demonstrating significance as reflected in their P values. The preference for learning style shows a lower importance, with a path coefficient of 0.083 and a P value of 0.129, which is above the 0.05 threshold.

IV. CONCLUSION

This research proposes a model that utilizes a pedagogy construct to determine the quality of game-based learning. This construct reflected the criteria of excellent learning content as well as learning delivery formatted in a mobile application. This construct measures how well a GBL can present a real-world learning experience within the game. This research also suggests that the pedagogy degree is influenced by two significant factors: the ability of the game-based learning to accommodate the special characteristic of their learner (learning style) and the degree of the learning content quality and how flexible learning content serves their variative learner (learning content).

Furthermore, this research showed that the learning content and learning style do have significant effects on the pedagogy construct. Enhancing or increasing the quality of learning content and learning style will improve the quality of the pedagogy element of game-based learning. In addition, this research has developed the LECOPELESE measurement scale, which can be used to measure the pedagogy factor and assess the quality of game-based learning from a pedagogy perspective.

ACKNOWLEDGMENT

We thank the Faculty of Computer Science, Universitas AMIKOM, for their assistance during the data collection process. We also extend our special thanks to Fakultas Teknologi Maklumat dan Komunikasi (FTMK), Universiti Teknikal Malaysia Melaka, for their expert help and support.

REFERENCES

- [1] H. Na and H. Sung, "Learn math through motion: A technology-enhanced embodied approach with augmented reality for geometry learning in K-12 classrooms," *Interact. Learn. Environ.*, vol. 4820, pp. 1–14, 2025, doi: 10.1080/10494820.2025.2450663.
- [2] V. Çakar and A. Karadağ, "User experiences of patients' relatives with a computer game about pressure ulcer prevention: A descriptive qualitative study," *J. Wound Manag.*, vol. 26, no. 2, pp. 58–67, 2025, doi: 10.35279/jowm2025.26.02.02.
- [3] H. M. Paiva *et al.*, "Educational computer games for burn prevention and first aid awareness," *Entertain. Comput.*, vol. 52, p. 100902, 2025, doi: 10.1016/j.entcom.2024.100902.

- [4] P. Stark *et al.*, “Acceptability study of a co-designed educational game about dementia for children: The Kids Dementia Game,” *PLoS One*, vol. 20, no. 5, pp. 1–10, 2025, doi: 10.1371/journal.pone.0320782.
- [5] M. N. Katsantonis, A. Manikas, N. Partarakis, and I. Mavridis, “Yellow training: A collaborative tCOFELET tabletop serious game for cybersecurity awareness and training,” *Int. J. Inf. Secur.*, vol. 24, no. 4, 2025, doi: 10.1007/s10207-025-01086-0.
- [6] A. Lago-Ferreiro, M. Á. Gómez-González, and J. C. López-Ardao, “A new serious game (e-SoundWay) for learning English phonetics,” *Multimodal Technol. Interact.*, vol. 9, no. 6, pp. 1–27, 2025, doi:10.3390/mti9060054.
- [7] M. Gonzalez Capdevila, K. A. Pistili Rodrigues, T. A. S. Furlan, and T. Granollers, “Quantifying heuristic evaluation,” *Comput. Stand. Interfaces*, vol. 92, p. 103891, 2025, doi: 10.1016/j.csi.2024.103891.
- [8] Z. Liu and Y. Yang, “Fundamental usability evaluation in visualization: Integrating Nielsen’s principles and the analytic hierarchy process,” *Int. J. Hum.-Comput. Interact.*, vol. 0, no. 0, pp. 1–29, 2025, doi: 10.1080/10447318.2025.2467455.
- [9] E. S. Abd Majid *et al.*, “Heuristic evaluation for virtual reality for paediatric cancer patients: Perceptions of healthcare professionals,” in *Proc. IEEE 9th Int. Conf. Serious Games Appl. Health (SeGAH)*, 2021, pp. 1–6, doi: 10.1109/segah52098.2021.9551892.
- [10] F. Lehnert, S. Doublet, and G. Sim, “Designing usability/user experience heuristics to evaluate e-assessments administered to children,” *Comput. Stand. Interfaces*, vol. 92, p. 103933, 2025, doi:10.1016/j.csi.2024.103933.
- [11] P. Gutiérrez-Sánchez *et al.*, “Evaluating usability for an AR-enabled escape room authoring ecosystem,” *Entertain. Comput.*, vol. 55, p. 100982, 2025, doi: 10.1016/j.entcom.2025.100982.
- [12] M. K. Othman, R. Mat, and N. Zulkipli, “WVREA heuristics: A comprehensive framework for evaluating usability in wearable virtual reality educational applications,” *Educ. Inf. Technol.*, vol. 30, no. 10, 2025, doi: 10.1007/s10639-024-13234-5.
- [13] F. Coelho and A. M. Abreu, “Systemic gamification theory (SGT): A holistic model for inclusive gamified digital learning,” *Multimodal Technol. Interact.*, vol. 9, no. 7, pp. 1–26, 2025, doi:10.3390/mti9070070.
- [14] P. Wang *et al.*, “Frameworks, methods and indicators for evaluating serious games in nursing education: A scoping review,” *Nurse Educ. Pract.*, vol. 87, p. 104448, 2025, doi: 10.1016/j.nepr.2025.104448.
- [15] E. Frost *et al.*, “HEaring and Llistening eXperience (HELIX): Evaluation of a co-designed serious game for auditory-cognitive training,” *Q. J. Exp. Psychol.*, 2025, doi:10.1177/17470218251324168.
- [16] S. N. S. Yasin and R. Ibrahim, “MEGUX: A UX model for mobile educational game evaluation,” in *Proc. 7th Int. Congr. Human-Computer Interaction, Optimization and Robotic Applications (ICHORA)*, 2025, pp. 1–7, doi:10.1109/ichora65333.2025.11017212.
- [17] H. Al Fatta, Z. Maksom, and M. H. Zakaria, “Systematic literature review on usability evaluation model of educational games: Playability, pedagogy, and mobility aspects,” *J. Theor. Appl. Inf. Technol.*, vol. 96, no. 14, pp. 4677–4689, 2018.
- [18] A. C. da Silveira, R. X. Martins, and E. A. O. Vieira, “E-Guess: Usability evaluation for educational games,” *RIED-Rev. Iberoam. Educ. a Distancia*, vol. 24, no. 1, pp. 245–263, 2021, doi:10.5944/ried.24.1.27690.
- [19] S. Criollo-C *et al.*, “Enhancing the educational model using mixed reality technologies with Meta Quest 3: A usability analysis using IBM-CSUQ,” *IEEE Access*, vol. 13, pp. 56930–56945, 2025, doi:10.1109/access.2025.3555930.
- [20] V. Nagalingam, R. Ibrahim, and R. C. M. Yusoff, “EDUGXQ: User experience instrument for educational games’ evaluation,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 1, pp. 562–569, 2020, doi:10.14569/ijacsa.2020.0110170.
- [21] S. de Freitas and S. Jarvis, “A framework for developing serious games to meet learner needs,” in *Proc. Interservice/Industry Training, Simulation, and Education Conf. (IITSEC)*, 2006, pp. 1–11.
- [22] C. Dondi and M. Moretti, “A methodological proposal for learning games selection and quality assessment,” *Br. J. Educ. Technol.*, vol. 38, no. 3, pp. 502–512, 2007, doi: 10.1111/j.1467-8535.2007.00713.x.
- [23] T. Connolly, M. Stansfield, and T. Hainey, “Development of a general framework for evaluating games-based learning,” in *Proc. 2nd Eur. Conf. Games-Based Learn.*, 2008, pp. 105–114.
- [24] S. B. Zaibon and N. Shiratuddin, “Heuristics evaluation strategy for mobile game-based learning,” in *Proc. 6th IEEE Int. Conf. Wireless, Mobile, and Ubiquitous Technol. Educ. (WMUTE)*, 2010, pp. 127–131, doi: 10.1109/WMUTE.2010.27.
- [25] G.-J. Hwang, H.-Y. Sung, C.-M. Hung, I. Huang, and C.-C. Tsai, “Development of a personalized educational computer game based on students’ learning styles,” *Educ. Technol. Res. Dev.*, vol. 60, no. 4, pp. 623–638, 2012, doi: 10.1007/s11423-012-9241-x.
- [26] O. Ak, “A game scale to evaluate educational computer games,” *Procedia Soc. Behav. Sci.*, vol. 46, pp. 2477–2481, 2012, doi:10.1016/j.sbspro.2012.05.506.
- [27] S. Arnab *et al.*, “The development approach of a pedagogically-driven serious game to support relationship and sex education (RSE) within a classroom setting,” *Comput. Educ.*, vol. 69, pp. 15–30, 2013, doi: 10.1016/j.compedu.2013.06.013.
- [28] M. Soflano, T. M. Connolly, and T. Hainey, “An application of adaptive games-based learning based on learning style to teach SQL,” *Comput. Educ.*, vol. 86, pp. 192–211, 2015, doi:10.1016/j.compedu.2015.03.015.
- [29] K. Rapeepisarn, K. W. Wong, C. C. Fung, and M. S. Khine, “The relationship between game genres, learning techniques and learning styles in educational computer games,” in *Technol. E-Learn. Digit. Entertain., Lecture Notes Comput. Sci.*, vol. 5093, Berlin, Heidelberg: Springer, 2008, pp. 497–508, doi: 10.1007/978-3-540-69736-7.
- [30] S. Hermawati and G. Lawson, “A user-centric methodology to establish usability heuristics for specific domains,” in *Contemp. Ergonomics Hum. Factors 2015*, 2015, pp. 80–85.
- [31] B. Lechner, S. Petter, A. Fruhling, and H. Siy, “The chicken and the pig: User involvement in developing usability heuristics,” in *Proc. 19th Americas Conf. Inf. Syst. (AMCIS)*, 2013, pp. 3263–3270.
- [32] V. R. Basili, *Software Modeling and Measurement: The Goal/Question/Metrics Paradigm*, 1992.
- [33] N. M. Minhas, T. R. Koppula, K. Petersen, and J. Börstler, “Using goal-question-metric to compare research and practice perspectives on regression testing,” *J. Softw. Evol. Process.*, vol. 35, no. 2, pp. 1–26, 2023, doi: 10.1002/smr.2506.
- [34] R. A. Alnanih, “The impact of the GQM framework on software engineering exam outcomes,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 10, pp. 613–622, 2024, doi: 10.14569/IJACSA.2024.0151063.
- [35] E. T. Loiacono, R. T. Watson, and D. L. Goodhue, “WebQual: A measure of website quality,” *Mark. Theory Appl.*, vol. 13, no. 3, pp. 432–438, 2002.
- [36] H. Mohamed and A. Jaafar, “Development and potential analysis of heuristic evaluation for educational computer game (PHEG),” in *Proc. 5th Int. Conf. Comput. Sci. Convergence Inf. Technol. (ICCIT)*, 2010, pp. 222–227, doi: 10.1109/ICCIT.2010.5711061.
- [37] S. B. Zaibon, “User testing on game usability, mobility, playability, and learning content of mobile game-based learning,” *J. Teknol.*, vol. 77, no. 29, pp. 131–139, Dec. 2015, doi: 10.11113/jtv77.6848.
- [38] I. Mayer, “Towards a comprehensive methodology for the research and evaluation of serious games,” *Procedia Comput. Sci.*, vol. 15, pp. 233–247, 2012, doi: 10.1016/j.procs.2012.10.075.
- [39] H. Mohamed-Omar, R. Yusoff, and A. Jaafar, “Quantitative analysis in a heuristic evaluation for usability of educational computer game (UsaECG),” in *Proc. Int. Conf. Inf. Retrieval & Knowl. Manag.*, 2012, pp. 187–192, doi: 10.1109/InfRKM.2012.6205033.
- [40] K. Bontcheva, “Behaviour: Some results from the evaluation of a dynamic hypertext system,” in *Lecture Notes Comput. Sci.*, pp. 69–78, 2002, doi: 10.1007/3-540-47952-X_9.
- [41] S. Lawless *et al.*, “Adaptive educational hypermedia systems in technology enhanced learning,” in *Proc. ACM Conf.*, 2010, p. 73, doi: 10.1145/1867651.1867672.
- [42] K. Kuutti, “Activity theory as a potential framework for human-computer interaction research,” in *Context and Consciousness: Activity Theory and Human-Computer Interaction*, MIT Press, 1996, pp. 17–44, doi: 10.7551/mitpress/2137.003.0006.
- [43] A. Hussain and M. Kutar, “Usability metric framework for mobile phone application,” in *Proc. 10th Annu. PostGrad. Symp. Convergence Telecommun., Netw. Broadcast.*, 2009, pp. 1–5, doi:10.1145/1497308.1497412.
- [44] B. Wilson and J. Henseler, “Modeling reflective higher-order constructs using three approaches with PLS path modeling: A Monte Carlo comparison,” 2007.
- [45] C. Crocetta *et al.*, “Higher-order PLS-PM approach for different types of constructs,” *Soc. Indic. Res.*, vol. 154, no. 2, pp. 725–754, 2021, doi: 10.1007/s11205-020-02563-w.
- [46] J. F. Hair Jr., W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis*, 8th ed., Cengage Learning, 2019.
- [47] A. Bhattacharjee, *Social Science Research: Principles, Methods, and Practices*, 2012.
- [48] J. F. Hair, G. T. M. Hult, and C. M. Ringle, *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, 3rd ed., SAGE Publications, 2021, doi: 10.1007/978-3-030-80519-7.