Available online at
https://mijuitm.com.my

e-ISSN: 2735-0703

Mathematical Sciences and Informatics Journal 6(2) 2025, 218 – 234.

Mathematical Sciences and Informatics Journal

Validating MEGUX: Reliability and Validity of a Mobile Educational Game UX Instrument

Sharifah Nurulhikmah Syed Yasin[1,2*], Roslina Ibrahim[1]

[1]*Faculty of Artificial Intelligence, Universiti Teknologi Malaysia, Kuala Lumpur*
[2]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Cawangan Terengganu, Kampus Kuala Terengganu*

ARTICLE INFO

Article history:
Received 18 August 2025
Revised 27 September 2025
Accepted 28 September 2025
Online first
Published 31 October 2025

Keywords:
Mobile Educational Games
User Experience (UX)
UX Evaluation Instrument
MEGUX
Mobile Learning
Reliability
Validity

DOI:
10.24191/mij.v6i2.9179

ABSTRACT

Mobile educational games (MEGs) blend gameplay and pedagogy on smartphones and tablets, yet existing UX frameworks focus on general usability, entertainment gaming, or desktop e-learning and do not fully address the mobile context or the dual aims of learning and engagement. This study evaluates the MEGUX model internal consistency and known-groups validity by comparing two programming-learning games. Following an eight-phase UX instrument development process, a set of an 88-item pool across six dimensions (Flow, Immersion, Player Context, Game Usability, Mobility, Learning) was generated. After expert review and pilot testing, MEGUX was administered to 265 undergraduates (AlgoRun, $n = 161$; Rodocodo, $n = 104$) over a one-week play period in March–April 2025. Internal consistency was assessed via Cronbach's α and known-groups validity was tested with independent-samples t-tests (with Levene's checks) and Cohen's d effect sizes. Cronbach's α values ranged from .820 (Player Context) to .899 (Flow, Learning), indicating strong reliability. Rodocodo significantly outperformed AlgoRun on 70 of 88 items ($p < .05, d = 0.23 - 0.74$), demonstrating MEGUX's sensitivity to UX differences in real-world MEGs. MEGUX is a comprehensive and robust instrument for evaluating UX in mobile educational games, offering designers and researchers a reliable tool to identify strengths and weaknesses in engagement and learning components.

1. INTRODUCTION

Global active smartphones in circulation increased from 14.2 billion units in 2020 to over 18 billion in year 2025 (Laricchia, 2023). This increase has made mobile devices a ubiquitous platform for learning. Mobile devices, especially smartphones, have reshaped the way Generation Y and later generations engage with digital educational content. They are used in both formal and informal educational settings. One of the

[1,2*] Corresponding author. *E-mail address*: nurulhikmah@uitm.edu.my
https://doi.org/10.24191/mij.v6i2.9179

emerging tools for learning through mobile devices is mobile educational games (MEGs). MEGs can provide interactive, easy access, and engaging experiences. Given their increasing use, it is essential to ensure that MEGs provide an engaging user experience (UX) and meaningful and effective learning (Rahmadi et al., 2021; Ramtohul and Khedo, 2022).

The traditional UX models and frameworks, such as ISO 9241-11, provide general standards for digital content that focus on usability. At the same time, Hassenzahl's experiential and emotional design model offers valuable guidance for digital system UX (Hassenzahl et al., 2003). Hassenzahl and Tractinsky (2006) serve as the foundation of MEGs UX, albeit with limitations in addressing the dual-purpose nature of MEGs that combine gameplay and pedagogy. To extend these foundations, existing UX approaches curated for gaming and educational technology provide further insights for specific purposes. The game Experience Questionnaire (GEQ) and Game Flow model have been widely used to evaluate engagement and enjoyment in gaming contexts (Poels et al., 2007) while the ARCS model emphasises motivation in educational environments (Savi et al., 2011; Imran, 2023). However, these models and frameworks often focus on a single gaming context, either for entertainment or educational purposes, without adequately addressing the integrated nature of mobile educational games. Thus, while general UX standards and experiential models provide the groundwork, and game-based or learning-focused models contribute domain-specific insights, there remains an underexplored area in the literature that requires the other facets of UX for MEGs, to propose a comprehensive model, to and validate a tool that can capture the digital game, learning, and mobile experience in MEGs.

The present study proposes and validates a UX model, MEGUX, tailored specifically for mobile educational games to address this gap. Accordingly, it has two primary objectives: (1) to evaluate the internal consistency of the MEGUX subscales using Cronbach's α, and (2) to demonstrate MEGUX's known-groups validity by comparing scores between the AlgoRun and Rodocodo games via independent-samples t-tests. Drawing on the UX model development methodology of Quiñones (2018), the study unfolds across eight phases that cover a review of relevant literature, construction of an initial item pool, expert evaluation of those items, a pilot study to refine the model, and validation of the instrument with real users. The remainder of this paper details the literature review, research methodology, empirical results, discussion, and concluding remarks.

2. LITERATURE REVIEW AND RELATED WORKS

2.1 Trends and growth of MEGs

Mobile devices overtook desktops as the fastest-growing computing platforms in the late 2000s, laying the groundwork for mobile educational games (MEGs). Early studies described a "radical break from traditional computer platforms," noting that by the early 2020s, there were roughly five mobile devices for every desktop and that over 70% of the world's population owned a smartphone (Li et al., 2022; Abduljawad and Ahmad, 2023; Chen and Morrell, 2025). At this level of adoption, mobile devices became more than just sophisticated gadgets and emerged as practical learning tools, becoming a dependable, widely adopted medium for educational activities.

Mobile devices such as smartphones and tablets offer new opportunities for mobile educational games to support various learning styles through their affordances. Portability and always-on connectivity make brief, on-the-go activities like micro-learning and scenario simulations available outside scheduled class sessions (Abduljawad and Ahmad, 2023; Chen and Morrell, 2025). Built-in GPS, accelerometer, and time sensors enable context-aware adaptations to cope with location and time challenges and support personalized learning options (Li et al., 2022; Abduljawad and Ahmad, 2023). These dynamic features, combined with game interactions, support self-paced study, peer collaboration, and continuous engagement.

https://doi.org/10.24191/mij.v6i2.9179

In higher-education contexts, the low cost of mobile devices and expanding campus wireless infrastructure have accelerated MEG acceptance (Abdallah et al., 2021). Mobile tools now support traditional teaching by providing on-demand access to readings, enabling assignment submission, and facilitating group work without the need for fixed computer labs (Abdallah et al., 2021). Moreover, MEGs align with lifelong and self-learning by allowing students to engage with content whenever they choose rather than within rigid schedules (Abdallah et al., 2021). Anchoring web-based e-learning as the starting point, MEGs use widespread smartphone ownership and context-aware features to provide smooth, game-like learning experiences that meet contemporary students' needs (Chen and Morrell, 2025).

2.2 UX Principles

According to Norman and Nielsen (1998), user experience is the entirety of a person's interactions with a company's products, services, and brand, combining functional effectiveness with simplicity and emotional satisfaction. A typical user experience meets users' exact needs without friction, delivers elegant and joyful interactions, and results from the seamless integration of engineering, marketing, industrial design, and interface design. Within this holistic conception, usability becomes a critical subset, where Jakob Nielsen describes five quality components (learnability, efficiency, memorability, error tolerance, and satisfaction) that ensure interactive systems meet users' task-oriented needs with minimal friction (Nielsen, 2012). Moreover, his ten usability heuristics offer actionable design guidelines to reduce cognitive load and support usability and engagement in digital interfaces (Nielsen, 1994a). Prior to the 10 heuristics, Nielsen applied factor analysis to multiple heuristics lists and extracted nine components that closely match the ten-heuristic model, where "help and documentation" was folded into other components (Nielsen, 1994b).

Based on Nielsen's foundation, Hassenzahl et al. (2003) present the pragmatic–hedonic separation UX, where pragmatic quality covers usefulness and ease of use, and hedonic quality, comprised of stimulation (novelty, challenge) and identity (self-expression, social meaning), drives personal growth and emotional attachment (Hassenzahl et al., 2003). They demonstrate that these dimensions operate independently and can be measured using the AttrakDiff questionnaire (Hassenzahl et al., 2003). Later, (Hassenzahl and Tractinsky, 2006) describe UX as a multidimensional experience that focus on the interaction between the user's internal state (motivation, mood), system characteristics (complexity, presentation), and context of use (social setting, voluntariness). They recommend research approaches that consider needs beyond practical functionality needs, positive emotions, and experience progression over time (Hassenzahl and Tractinsky, 2006).

For mobile educational games (MEGs), integrating the above-mentioned foundations ensures that game usability (guided by Nielsen's heuristics and five attributes) supports learning efficiency. For instance, through clear feedback, consistent controls, and error-tolerant mechanics. While hedonic stimulation and identity opportunities sustain motivation and engagement (Nielsen, 1994a; Hassenzahl et al., 2003). Together, these principles form a solid framework for designing and evaluating MEGs interfaces that balance usability, engagement, and learning outcomes.

2.3 UX Models and Frameworks for MEGs

Nagalingam et al. (2020) introduced EDUGX to structure UX assessment for online educational games through six dimensions: Flow, Immersion, Player Context, Usability, Game System, and Learnability. This framework was proposed from a comprehensive literature review and analysis of pedagogy and gameplay. The experts have examined and thoroughly validated it using real users' data. However, EDUGX does not consider mobile-specific features such as screen size, interrupted play sessions, touch screen, and device variability. Therefore, this framework may neglect the mobile platform-specific needs where the hardware and context of use are unique.

https://doi.org/10.24191/mij.v6i2.9179

Another frequent model used to evaluate MEGs is GEQ. Poels et al. (2007) and IJsselsteijn et al. (2013) offer a modular instrument (core game, post-game, and social presence) to capture players' subjective experiences across dimensions like competence, flow, and affect. GEQ was validated in an extensive survey carried out in the FUGA project, where the questionnaires were confirmed using responses from hundreds of players across various game genres. However, GEQ prioritizes the entertainment elements and outcomes over the learning features. It also lacks coverage of the on-the-go interactions and the educational needs of MEGs. Thus, its modularity is important in MEGs evaluation; it requires more components to measure educational effectiveness and mobile usability.

Meanwhile, the E-GUESS heuristic framework (da Silveira et al., 2021) extends usability evaluation for educational games by integrating pedagogical criteria with traditional usability checks. It guides designers in uncovering interface issues early, ensuring that game mechanics effectively convey educational content. Still, E-GUESS was developed without explicit attention to mobile affordances; touch controls, screen fragmentation, or varying usage contexts, thereby limiting its applicability to smartphone and tablet environments.

Several heuristic-based approaches have also been proposed specifically for mobile games. Korhonen and Koivisto (2006) proposed playability heuristics for mobile game design and evaluation. The heuristics emphasize the early detection of playability flaws in mobile games by segmenting the heuristics in the usability, mobility, and gameplay components. However, these heuristics lack assessing the learning outcomes, emotional engagement, and flow states tailored to educational gameplay. Therefore, playability heuristics leave the UX area insufficient for the UX evaluation model or framework to evaluate MEGs comprehensively.

Lastly, Shoukry et al. (2015) presented Pre-MEGa as a comprehensive checklist for designing preschool learning games, spanning screen layout, navigation, feedback, and pedagogical features across fifteen categories. The detailed guidelines in pre-MEGa cover child-centred design issues but also have the potential to overwhelm game designers. Additionally, the details may not align well with primary, secondary school, and adult learners or multiple device contexts (Kolak et al., 2021). Table 1 illustrates the main components highlighted in the previous studies reviewed. The development of the MEGUX model is grounded in these components.

Table 1. Dimensions and components of selected UX evaluation models for mobile or educational games

Framework/ Model	Dimensions / Components
EDUGX	Flow; Immersion; Player Context; Usability; Game System; Learnability
GEQ	Competence; Flow; Immersion; Tension; Challenge; Negative Affect; Positive Affect; Social Presence
E-GUESS	Usability Heuristics; Pedagogical Integration; Narrative Delivery; Educational Content Delivery
Playability Heuristics	Game Controls; UI Intuitiveness; Context Awareness; Gameplay Mechanics
Pre-MEGa	Screen Design; Navigation & Control; Ease of Use; Responsiveness; Game Design; Learning Potential; Instructions; Feedback; Difficulty; Content Delivery; Pedagogical Agent; Customization; Security; Accessibility; Value
MEGUX (proposed)	Flow; Immersion; Player Context; Usability; Game Usability; Learning; Mobility

Table 1 illustrates the overlaps and differences among existing models and frameworks, showing how recurring dimensions such as Flow, Immersion, and Usability are consistently emphasized in both game and educational contexts. These commonalities formed the foundation of MEGUX. The development of the MEGUX model was based on (Quiñones et al., 2018) methodology for UX model development. MEGs differ from online educational or entertainment-focused games in that they must integrate gameplay, pedagogy, and the unique affordances of mobile platforms. To address this, MEGUX extends beyond EDUGX by incorporating three additional components: Game Usability, Learning, and Mobility.

Game Usability was adapted from playability heuristics (Korhonen & Koivisto, 2006). Game usability emphasizes the operational controls, intuitive interface, and game mechanics, determining the game's

https://doi.org/10.24191/mij.v6i2.9179

effectiveness. The learning component was additionally derived from educational evaluation approaches in Pre-MEGa (Shoukry et al., 2015). Finally, Mobility reflects insights from mobile playability heuristics (Korhonen & Koivisto, 2006) and Pre-MEGA (Shoukry et al., 2015). The two frameworks recognize interruptions, portability, and touch-screen interactions that critically shape user experience in mobile environments.

2.4 UX Principles in Mobile Environment

Grant (2019) investigates and categorizes mobile learning definitions into four categories, identifying the 'nomadic learner' perspective as the one that most accurately reflects the core idea of mobility. Learning is disconnected by time or place: learners engage with content anytime, anywhere, and draw upon multiple contexts and resources. This view underscores key UX requirements for mobile environments. Those are seamless access, context-sensitivity, and learner autonomy, which complement broader mobile UX goals of flexibility and personalization.

Huang (2020) offers a wide view of usability needs coping for education, commerce, healthcare, and beyond. By mapping common mobile constraints such as small displays, interruptible connectivity, and limited input modalities to corresponding UX requirements, Huang identifies a set of essential attributes. These include simplicity, consistency, visibility of system status, and error management, which designers must address to ensure effective mobile interactions. Previous studies points out that although mobile apps serve many different purposes, their UX guidelines all focus on handling device limits such as small screens, battery life, and processing power. They also adapt to changing user contexts, including where and how people use their devices.

Based on Huang's framework, Alturki and Gay (2019) systematically reviewed post-2010 guidelines on mobile applications and discovered eight fundamental attributes: satisfaction, effectiveness, efficiency, learnability, operability, attractiveness, simplicity, and usefulness. Their synthesis confirms that these pillars are universally applicable across mobile applications and serve as a ready checklist for evaluating mobile UX, regardless of content specifics.

To discuss mobile learning, Kumar et al. (2019) explore the early mobile-learning heuristics by analysing usability issues in 17 studies. The study listed the refined guidelines for mobility-related considerations (content organization, navigation, and layout) into subcomponents like touch-screen design, interruption management, and accessibility. Empirical results from mobile learning research by Mulhem and Almaiah (2021) demonstrates that thoughtfully structured course designs significantly boost learner satisfaction and performance. This structured design includes integrating mobility-cantered UX principles into the design process. Mobile learning design must not be done merely on surface-level interface tweaks so that it can improve educational impact.

Table 2. Key UX components for mobile learning and application environments

Author(s)	Component
Grant (2019)	Seamless access, context-sensitivity, learner autonomy
Huang (2020)	Simplicity, consistency, visibility of system status, error management
Alturki and Gay (2019)	Satisfaction, effectiveness, efficiency, learnability, operability, attractiveness, simplicity, usefulness
Kumar et al. (2019)	Touch-screen design, interruption management, accessibility, content organization, navigation, layout
Mulhem and Almaiah (2021)	Structured course design with mobility-centered UX principles
Krull and Duart (2017)	Adaptive design for evolving devices and usage contexts

Finally, mobile learning design strategies must remain adaptive as mobile hardware and usage patterns evolve (Krull & Duart, 2017). A dedicated mobility component in any mobile UX model ensures that designs can flexibly accommodate emerging device capabilities such as foldable screens, AR/VR, and

https://doi.org/10.24191/mij.v6i2.9179

shifting user contexts. To summarize the components, Table 2 highlights the key UX components and elements based on reviewed literatures.

2.5 Research Gap

Existing UX models and frameworks for MEGs fall short in addressing the unique demands of mobile platforms. As depicted in Fig. 1, an ideal MEG UX model sits at the relationship of Gameplay, Pedagogy, and Usability & System, all enclosed by Mobile Context and sustained Immersion. EDUGX balances education and gameplay but omits touch interaction, session interruptions, and device variability, risking overlooked usability issues on smartphones and tablets. The GEQ excels at measuring entertainment but neglects learning impact and real-world mobile contexts, limiting its relevance for MEGs. Heuristic tools like E-GUESS and mobile playability guidelines consider pedagogical and mobility factors, respectively, yet neither simultaneously evaluates screen fragmentation, interaction errors, nor learning outcomes and flow. Pre-MEGa offers a comprehensive preschool checklist but is overly detailed and poorly generalizes to older learners or multi-device scenarios. Finally, general mobile UX principles as summarised in Table 2 are never fully integrated with game-specific UX, heuristics, and educational metrics. These gaps underscore the need for a unified MEG evaluation model that combines mobile-interaction concerns (touch, context variability, performance) with robust measures of learning efficacy, engagement, flow, immersion, and usability in authentic usage environments.

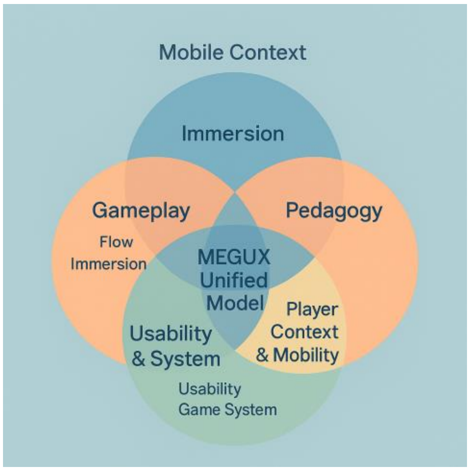

Fig. 1. Identified gaps in MEGs user experience based on previous literature

3. RESEARCH METHODOLOGY

To address the identified gap, this study adopted the eight-phase UX model development methodology proposed by Quiñones et al. (2018) as shown in Fig. 2. The proposed MEGUX model and the pilot study findings have been published and thoroughly discussed in a conference proceeding paper, "MEGUX: A UX Model for Mobile Educational Game Evaluation" (Yasin & Ibrahim, 2025). This study focuses on the final validation with real users by analysing the real user's data using reliability Cronbach's alpha and Independent-samples t-tests.

https://doi.org/10.24191/mij.v6i2.9179

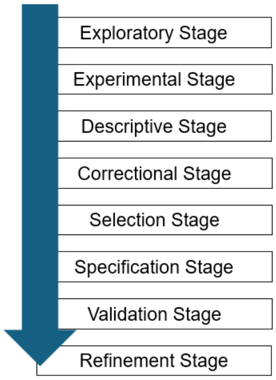

Fig. 2. UX model development methodology (Quiñones et al., 2018)

In this study, the finalised MEGUX instrument was validated with 265 participants across two distinct educational games on smartphones and tablets. The games were AlgoRun and Rodocodo. Both are MEGs specifically for computer programming learning. The respondents are the undergraduate students from Universiti Teknologi Mara Kuala Terengganu Campus. They were given one week to play the games before the MEGUX questionnaire was distributed. The 265 respondents completed answered all the MEGUX items in three controlled sessions. The collected data examines the performance and discriminative ability of MEGUX items across the two mobile educational games using inferential statistics and internal consistency metrics. Mean scores for each of the six MEGUX dimensions (Flow, Immersion, Game Usability, Player Context, Mobility, and Learning) are presented in the accompanying figures. To test whether these mean scores differ significantly between the two games, independent-samples t-tests were performed for each item, with Levene's test checking homogeneity of variances and Cohen's d computed to gauge effect sizes. Reliability was evaluated via Cronbach's alpha for each dimension.

4. RESULTS

An overview of the MEGUX instrument's validation study is provided based on data from 265 undergraduate participants who engaged with two mobile educational games (AlgoRun and Rodocodo) over one week. Descriptive statistics for the six MEGUX components (Flow, Immersion, Learning, Usability, Game Usability, Player Context, and Mobility) illustrate central tendencies and dispersion across both games. The number of items in each component is shown in Table 3. Independent-samples t-tests were conducted for each dimension to assess whether mean scores differ significantly between AlgoRun and Rodocodo. Levene's tests confirmed homogeneity of variances and Cohen's d were calculated to quantify effect sizes. Finally, Cronbach's alpha coefficients for each dimension are reported to confirm internal consistency and reliability, demonstrating the discriminative power and psychometric robustness of the MEGUX instrument in real-world mobile educational game contexts.

Table 3. Items used in each MEGUX component

Component	Sub-component	Item
	Clear goal	CG1, CG2, CG3, CG4 (4 items)
	Feedback	F1, F2, F3, F4, F5 (5 items)
Flow	Control	CT1, CT2, CT4, CT5 (4 items)
	Challenge	CL1, CL2, CL4, CL5, CL6 (5 items)
	Concentration	CN1, CN2, CN4, CN5, CN6 (5 items)

https://doi.org/10.24191/mij.v6i2.9179

Component	Sub-component	Item
Immersion	Engagement	EG1, EG2, EG3, EG4, EG5, EG6, EG7, EG8, EG9, EG12, EG13, EG14 (12 items)
	Engrossment	ER1, ER2, ER3, ER4, ER5, ER6, ER7 (7 items)
Player Context	User environment	UE1, UE5, UE6, UE7, UE8 (5 items)
	Temporal influences	TM1, TM2, TM3, TM4, TM5 (5 items)
Game Usability	Operability	OP1, OP2, OP3, OP4, OP5 (5 items)
	Understandability	UD1, UD2, UD3, UD4 (4 items)
	Attractiveness	AT1, AT2, AT3 (3 items)
Mobility	Mobile	MO1, MO2, MO3, MO4, MO5 (5 items)
	Touch specific	TS1, TS2, TS3, TS5 (4 items)
Learning	Knowledge improvement	K1, K2, K3, K4, K5 (5 items)
	Learning content	LC1, LC2, LC3, LC4, LC5 (5 items)
	Learning goal	LG1, LG2, LG3, LG4, LG5 (5 items)
Total		88 items

Table 4 summarises participant demographics for the two games. In the AlgoRun game ($N = 161$), 55.9% were male and 44.1% female; nearly all were in their first semester (98.8%). In the Rodocodo condition ($N = 104$), 45.2% were male and 54.8% female, and while 47.1% were first-semester students, the remainder were distributed across semesters 2 (6.7%), 3 (9.6%), 4 (7.7%), 5 (26.0%), and 6 (2.9%).

Table 4. Demographic information

Demographic	Category	Frequency and percentage	
		AlgoRun ($n = 161$)	Rodocodo ($n = 104$)
Gender	Male	90 (55.9%)	47 (45.2%) 137
	Female	71 (44.1%)	57 (54.8%) 128
Semester	1	159 (98.8%)	49 (47.1%)
	2	2 (1.2%)	7 (6.7%)
	3	0 (0.0%)	10 (9.6%)
	4	0 (0.0%)	8 (7.7%)
	5	0 (0.0%)	27 (26.0%)
	6	0 (0.0%)	3 (2.9%)

4.1 Reliability and Independent-Samples T-tests Results

A normality screen (skewness/kurtosis) for all 87 items showed values within the ±2 guideline (George & Mallery, 2016), except MO4, whose kurtosis was 2.062—a marginal exceedance. We retained MO4 because (i) independent-samples t-tests are robust to mild non-normality with samples >≈30 (Thomas et al., 2002; Blanca et al., 2017), (ii) reliability estimation does not require normality (Vaske et al., 2017; McNeish, 2018). Therefore, no transformations were applied, and all items were retained for subsequent analyses.

With the distributional checks satisfied and all items retained, the dataset was suitable for subsequent reliability estimation and between-group comparisons. Cronbach's α coefficients were first calculated for each MEGUX component to assess internal consistency which values ≥ 0.70 indicate acceptable reliability and ≥ 0.80 denote good to excellent coherence (Peterson, 1994; Tavakol & Dennick, 2011). Corrected item total correlations and "α if deleted" statistics accompany each overall α to verify that no individual item detracts from scale robustness. Independent-samples t-tests (with Welch's correction when Levene's test signalled unequal variances) then compared mean scores between AlgoRun (AR) and Rodocodo (RC) across the six components: Flow, Immersion, Learning, Game Usability, Player Context, and Mobility. Each comparison reports t (with df), two-tailed p-value, and Cohen's d. Fig. 3–8 present combined

https://doi.org/10.24191/mij.v6i2.9179

heatmaps of each component displaying corrected item total correlation, α-if-deleted, group means, t-values, and effect sizes, offering a concise visual summary of psychometric soundness and game-specific differences in user experience. Detailed item-level results tables are provided separately at *https://bit.ly/meguxresults*.

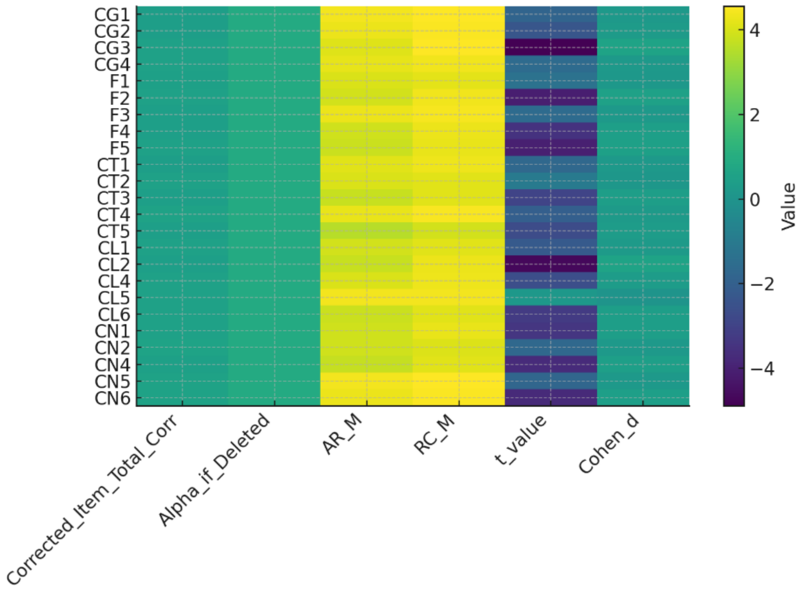

Fig. 3. Flow reliability and group differences metrics heatmap

Fig. 3 presents a heatmap of corrected item total correlations, α-if-deleted, AR vs. RC means, t-values, and Cohen's d for all 24 Flow items. Overall reliability was excellent ($\alpha = 0.899$), and deleting any single item would not improve it (α-if-deleted range: 0.893–0.898). In the t-value column, two Clear Goals items (CG2, CG3), three Feedback items (F2, F4, F5), three Control items (CT3, CT4, CT5), four Challenge items (CL1, CL2, CL4, CL6), and four Concentration items (CN1, CN3, CN4, CN6) stand out in darker hues, indicating significantly higher immersion ratings for RC (all $p < .05, d = .25 - .60$). The remaining eight items, shaded in neutral tones, showed no significant group differences ($p \geq .05, d \leq .20$), underscoring comparable flow experiences across games on those components.

https://doi.org/10.24191/mij.v6i2.9179

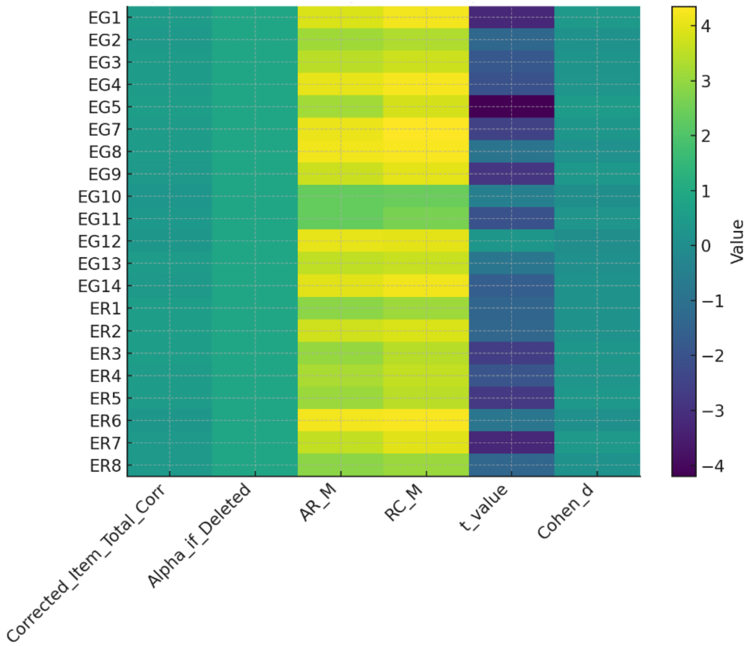

Fig. 4. Immersion reliability and group differences metrics heatmap

Fig. 4 presents a heatmap of corrected item total correlations, α-if-deleted, AR vs. RC means, t-values, and Cohen's d for all 22 Immersion items. Overall reliability was excellent ($\alpha = .876$), and deleting any single item would not improve it (α-if-deleted range: .862–.874). In the t-value column, six Engagement (EG) items (EG1, EG4, EG5, EG7, EG9, EG11) and three Engrossment (ER) items (ER3, ER5, ER7) stand out in darker hues, indicating significantly higher ratings for RC (all $p < .05, d = .26 - .53$). The remaining items, shaded in neutral tones, showed no significant group differences ($p \geq .05, d \leq .23$), underscoring comparable immersion across games on those components.

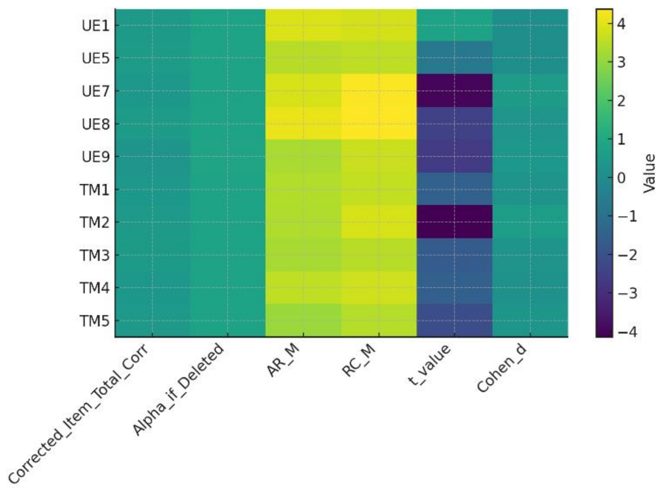

Fig. 5. Player context reliability and group differences metrics heatmap

https://doi.org/10.24191/mij.v6i2.9179

Fig. 5 presents a heatmap of corrected item–total correlations, α-if-deleted, AR vs. RC means, t-values, and Cohen's d for all 10 Player Context items. Overall reliability was strong (α = 0.820), and deleting any single item would not improve it (α-if-deleted range: 0.802–0.822). In the t-value column, three User Environment items (UE7, UE8, UE9) and two Temporal Influence items (TM2, TM5) stand out in darker hues, indicating significantly higher ratings for RC (all $p < .05$, $d = 0.31–0.63$). The remaining items, shaded in neutral tones, showed no significant group differences ($p \geq .097$, $d \leq 0.28$), underscoring comparable player-context experiences across games on those components.

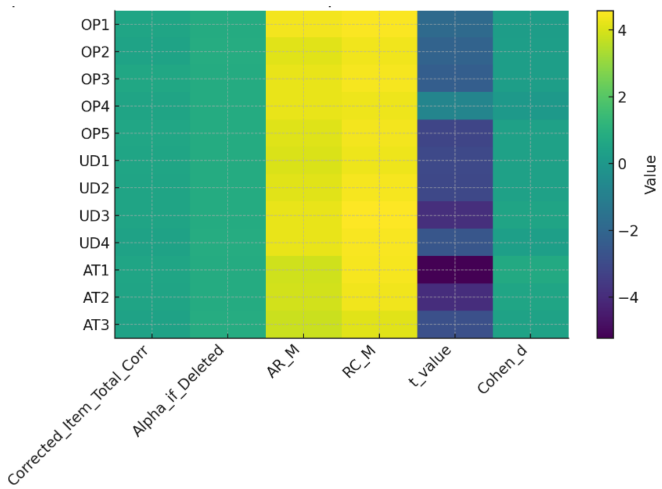

Fig. 6. Game usability reliability and group differences metrics heatmap

Fig. 6 presents a heatmap of corrected item–total correlations, α-if-deleted, AR vs. RC means, t-values, and Cohen's d for all 12 Game Usability items. Overall reliability was strong ($\alpha = 0.839$), and deleting any single item would not improve it (α-if-deleted range: 0.819–0.836). In the t-value column, three Operability items (OP2, OP3, OP5) four Understandability items (UD1, UD2, UD3, UD4) and all three Attractiveness items (AT1, AT2, AT3) stand out in darker hues, indicating significantly higher ratings for RC (all $p < .05, d = 0.28 - 0.74$). The two Operability items OP1 and OP4, shaded in neutral tones, showed no significant group differences ($p \geq .069, d \leq 0.10$), underscoring comparable usability on those components.

https://doi.org/10.24191/mij.v6i2.9179

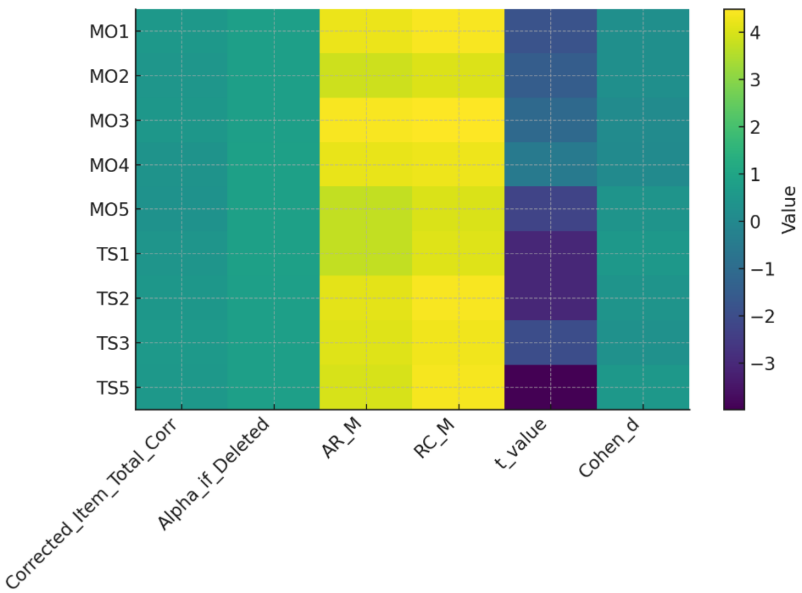

Fig. 7. Mobility reliability and group differences metrics heatmap

Fig. 7 presents a heatmap of corrected item–total correlations, α-if-deleted, AR vs. RC means, t-values, and Cohen's d for all nine Mobility items. Overall reliability was strong ($\alpha = 0.826$), and deleting any single item would not improve it (α-if-deleted range: 0.806–0.831). In the t-value column, one Mobile Functionality item (MO5) and three Touch Interaction items (TS1, TS2, TS5) stand out in darker hues, indicating significantly higher ratings for RC (all $p < .05, d = 0.41 - 0.55$). TS3 appears mid-tone, reflecting its marginal difference ($p = .050, d = 0.29$). The remaining items, shaded in neutral tones, showed no significant group differences ($p \geq .072, d \leq 0.25$), underscoring comparable general mobility features across games on those components.

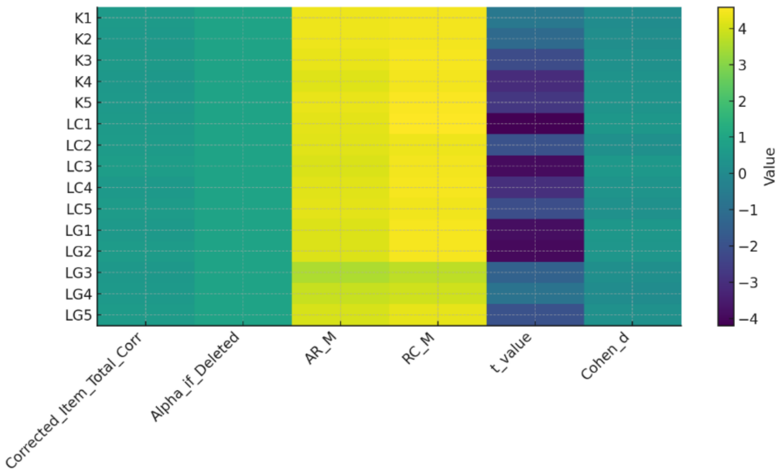

Fig. 8. Learning reliability and group differences metrics heatmap

https://doi.org/10.24191/mij.v6i2.9179

Fig. 8 presents a heatmap of corrected item–total correlations, α-if-deleted, AR vs. RC means, t-values, and Cohen's d for all 15 Learning items. Overall reliability was excellent ($\alpha = 0.899$), and deleting any single item would not improve it (α-if-deleted range: 0.887–0.899). In the t-value column, three Knowledge items (K3, K4, K5) five Learning Content items (LC1, LC2, LC3, LC4, LC5) and three Learning Goals items (LG1, LG2, LG5) stand out in darker hues, indicating significantly higher ratings for RC (all $p < .05, d = 0.23 - 0.48$). The remaining items (K1, K2, LG3, LG4) are shaded in neutral tones, showing no significant group differences ($p \geq .156, d \leq 0.19$), underscoring comparable performance on those components.

5. DISCUSSION

The present study applied the MEGUX model, which comprises six components (Flow, Immersion, Learning, Game Usability, Player Context, and Mobility), to evaluate and differentiate user experiences in two mobile educational games, AlgoRun and Rodocodo. Across all six components, internal consistency was uniformly strong ($\alpha = .820 - .899$), and no single item weakened scale reliability. Independent-samples comparisons further demonstrated that MEGUX items effectively discriminate between the two games: Rodocodo generated significantly higher ratings on the majority of items within each component (Flow: 18/24 items; Immersion: 9/22; Learning: 11/15; Game Usability: 10/12; Player Context: 5/10; Mobility: 4/9), with effect sizes ranging from small ($d \approx 0.23$) to moderate ($d \approx 0.74$).

5.1 Effectiveness of MEGUX in Capturing MEG Experience

The robustness of Cronbach's α across components confirms that MEGUX items coherently measure their intended components. Heatmap visualizations (Fig. 3-8) highlighted which items strongly drive scale reliability and which reveal pronounced game-related differences. For example, flow subcomponents (Clear Goals, Feedback, Control, Challenge, Concentration) and Learning facets (Knowledge, Content, Goals) displayed both high item–total correlations and substantial RC > AR contrasts, indicating that MEGUX sensitively captures the aspects of game engagement and educational impact. Similarly, Game Usability and Player Context items pinpointed specific operational and contextual features where Rodocodo outperformed, while Mobility items identified areas (touch interaction metrics TS1, TS2, TS5) critical to mobile-specific UX. These findings are consistent with previous studies stressing the importance of Flow and Immersion in sustaining user engagement (Poels et al., 2007; Hassenzahl & Tractinsky, 2006). The findings also align with EDUGX (Nagalingam et al., 2020), emphasizing Player Context's role in learning experiences. However, MEGUX extends this by explicitly integrating mobile-specific features (Korhonen & Koivisto, 2006).

5.2 Discrimination between games

The consistent pattern of higher Rodocodo ratings across most items underscores MEGUX's capacity to detect real differences in design quality and pedagogical efficacy. Effect sizes in key areas, such as Attractiveness (d up to 0.74), Learning Content clarity (d up to 0.48), and Touch Interaction (d up to 0.55) illustrate that MEGUX goes beyond simple usability to encompass emotional engagement and learning outcomes, making it a comprehensive tool for MEG evaluation. This result is partly in line with the E-GUESS framework (da Silveira et al., 2021), which highlights integrating pedagogical features with usability checks. However, unlike E-GUESS, MEGUX could differentiate between two MEGs by considering mobility factors, which supports earlier calls to address screen fragmentation and interruptions in mobile gameplay (Sophonhiranrak, 2021). In terms of educational impact, MEGUX aligns with Pre-MEGa (Shoukry et al., 2015) in recognizing the importance of structured learning content but avoids the excessive detail that could overwhelm designers and makes the UX model more complex and content-oriented (Kolak et al., 2021).

https://doi.org/10.24191/mij.v6i2.9179

5.3 Limitations and Future Directions

Regardless of the strength, several limitations were spotted. First, certain items in the model showed lower item–total correlations (UE9, MO5) or marginal group differences (TS3, LG3), suggesting that these items may require further theoretical refinement. Second, the number of game samples is limited to two MEG, and a specific sample population, which may constrain generalizability. This limitation echoes concerns raised in prior literature that many UX evaluation tools, including pre-MEGa and E-GUESS, were either developed for narrow age groups or lacked broad validation across contexts (Shoukry et al., 2015; da Silveira et al., 2021). For the future search, the UX model should focus on the adaptability across learner demographics and the scalability of the future mobile device context.

6. CONCLUSION

This study proposes and validates a user experience model, MEGUX. It is a comprehensive UX evaluation model specifically tailored for MEGs. To overcome the limitations in existing UX frameworks and models, MEGUX integrates six critical components: Flow, Immersion, Game Usability, Player Context, Mobility, and Learning. The model offers a multi-component tool that captures educational value and mobile-specific interaction. Empirical validation with 265 undergraduate participants across two MEGs demonstrates strong internal consistency ($\alpha = .820 - .899$) and robust discriminative ability across all components. The results shows that MEGUX effectively identifies differences in gameplay, engagement, and learning outcomes between two distinct MEGs, confirming its capacity to assess UX in authentic mobile learning contexts. This study contributes a mobile-aware UX instrument that bridges gaming and pedagogy. Simultaneously, the model also addresses the challenges of mobile interfaces and technology. The MEGUX model facilitates MEGs researchers and designers to precisely measure the important UX component for their games. Future work should expand its validation across varied MEGs, user groups, and educational domains to ensure broader applicability and continued refinement.

7. ACKNOWLEDGEMENTS/FUNDING

The authors would like to acknowledge the support of Universiti Teknologi Mara (UiTM), Cawangan Negeri Terengganu, Kampus Kuala Terengganu and Faculty of Computer and Mathematical Sciences, Universiti Teknologi Malaysia and Faculty of Artificial Intelligence for providing the facilities and financial support on this research.

8. CONFLICT OF INTEREST STATEMENT

The authors agree that this research was conducted in the absence of any self-benefits, commercial or financial conflicts and declare the absence of conflicting interests with the funders.

9. AUTHORS' CONTRIBUTIONS

Sharifah Nurulhikmah Syed Yasin conducted research, performed the analysis, and wrote the manuscript. Roslina Ibrahim conceptualised the research framework, provided critical guidance throughout the study, contributed to the methodology refinement, and reviewed and approved the final manuscript for submission.

https://doi.org/10.24191/mij.v6i2.9179

REFERENCES

Abdallah, N., Abdallah, O., & Bohra, O. M. (2021). Factors affecting mobile learning acceptance in higher education: An empirical study. *International Journal of Advanced Computer Science and Applications, 12*(4), 664–671. https://thesai.org/Publications/ViewPaper?Code=IJACSA&Issue=4&SerialNo=82&Volume=12

Abduljawad, M., & Ahmad, A. (2023). An analysis of mobile learning (M-learning) in education. *Multicultural Education, 9*(2), 145–152. https://zenodo.org/records/7665894

Al Mulhem, A., & Almaiah, M. A. (2021). A conceptual model to investigate the role of mobile game applications in education during the COVID-19 pandemic. *Electronics, 10*(17), 2106. https://doi.org/10.3390/electronics10172106

Alturki, R., & Gay, V. (2019). Usability attributes for mobile applications: A systematic review. In A. O. Fapojuwo, S. Misra, A. M. A. Haqiq, & I. Awan (Eds.), *Recent Trends and Advances in Wireless and IoT-enabled Networks* (pp. 53–62). Springer. https://link.springer.com/chapter/10.1007/978-3-319-90802-1_5

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema, 29*(4), 552–557. https://doi.org/10.7334/psicothema2016.383

Chen, M., & Morrell, E. (2025). Exploring the purpose and development of academic games: An analysis of games reported at the Foundations of Digital Games Conference (2007–2024). In *Proceedings of the 20th International Conference on the Foundations of Digital Games (FDG 2025)* (pp. 1–11). ACM. https://doi.org/10.1145/3723498.3723823

da Silveira, A. C., Martins, R. X., & Vieira, E. A. O. (2021). E-GUESS: Usability evaluation for educational games. *RIED-Revista Iberoamericana de Educación a Distancia, 24*(1), 245–263. https://doi.org/10.5944/ried.24.1.27690

Grant, M. M. (2019). Difficulties in defining mobile learning: Analysis, design characteristics, and implications. *Educational Technology Research and Development, 67*(2), 361–388. https://doi.org/10.1007/s11423-018-09641-4

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In J. Ziegler & G. Szwillus (Eds.), *Mensch & Computer 2003: Interaktion in Bewegung* (pp. 187–196). B. G. Teubner.

Hassenzahl, M., & Tractinsky, N. (2006). User experience—A research agenda. *Behaviour & Information Technology, 25*(2), 91–97. https://doi.org/10.1080/01449290500330331

Huang, Z. (2020). Mobile usability: Review, classifications and future directions. In A. Marcus & W. Wang (Eds.), *Design, User Experience, and Usability. Design for Contemporary Interactive Environments (DUXU 2020)* (pp. 270–276). Springer. https://doi.org/10.1007/978-3-030-49760-6_19

IJsselsteijn, W. A., de Kort, Y. A. W., & Poels, K. (2013). *The Game Experience Questionnaire (GEQ)*. Technische Universiteit Eindhoven. https://pure.tue.nl/ws/files/21666907/Game_Experience_Questionnaire_English.pdf

Imran, H. (2023). An empirical investigation of the different levels of gamification in an introductory programming course. *Journal of Educational Computing Research, 61*(4), 847–874. https://doi.org/10.1177/07356331221144074

https://doi.org/10.24191/mij.v6i2.9179

Kolak, J., Norgate, S. H., Monaghan, P., & Taylor, G. (2021). Developing evaluation tools for assessing the educational potential of apps for preschool children in the UK. *Journal of Children and Media, 15*(3), 410–430. https://doi.org/10.1080/17482798.2020.1844776

Korhonen, H., & Koivisto, E. M. I. (2006). Playability heuristics for mobile games. In *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '06)* (pp. 9–16). ACM. https://doi.org/10.1145/1152215.1152218

Krull, G., & Duart, J. M. (2017). Research trends in mobile learning in higher education: A systematic review of articles (2011–2015). *The International Review of Research in Open and Distributed Learning, 18*(7), 1–23. https://doi.org/10.19173/irrodl.v18i7.2893

Kumar, B. A., Goundar, M. S., & Chand, S. S. (2019). Usability guideline for mobile learning applications: An update. *Education and Information Technologies, 24*(6), 3537–3553. https://doi.org/10.1007/s10639-019-09937-9

Laricchia, F. (2023). Forecast number of mobile devices worldwide from 2020 to 2025. *Statista*. https://www.statista.com/statistics/245501/multiple-mobile-device-ownership-worldwide/

Li, Y., Xu, Z., Hao, Y., Xiao, P., & Liu, J. (2022). Psychosocial impacts of mobile game on K-12 students and trend exploration for future educational mobile games. *Frontiers in Education, 7*, 843090. https://doi.org/10.3389/feduc.2022.843090

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health, 23*, 151–169. https://doi.org/10.1146/annurev.publhealth.23.100901.140546

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods, 23*(3), 412–433. https://doi.org/10.1037/met0000144

Nagalingam, V., Ibrahim, R., & Yusoff, R. C. M. (2020). EDUGXQ: User experience instrument for educational games' evaluation. *International Journal of Advanced Computer Science and Applications, 11*(1), 562–569. https://doi.org/10.14569/IJACSA.2020.0110170

Nielsen, J. (1994a). Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)* (pp. 152–158). ACM. https://doi.org/10.1145/191666.191729

Nielsen, J. (1994b). 10 usability heuristics for user interface design. *Nielsen Norman Group*. https://www.nngroup.com/articles/ten-usability-heuristics/

Nielsen, J. (2012). Usability 101: Introduction to usability. *Nielsen Norman Group*. https://www.nngroup.com/articles/usability-101-introduction-to-usability/

Norman, D., & Nielsen, J. (1998). The definition of user experience (UX). *Nielsen Norman Group*. https://www.nngroup.com/articles/definition-user-experience/

Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research, 21*(2), 381–391. https://doi.org/10.1086/209405

Poels, K., de Kort, Y. A. W., & IJsselsteijn, W. A. (2007). *The Game Experience Questionnaire: The fun of gaming – Measuring the human experience of media enjoyment* (FUGA D3.3 project deliverable). Eindhoven University of Technology. https://pure.tue.nl/ws/files/21666952/Fuga_d3.3.pdf

https://doi.org/10.24191/mij.v6i2.9179

Quiñones, D., Rusu, C., & Rusu, V. (2018). A methodology to develop usability/user experience heuristics. *Computer Standards & Interfaces, 59*, 109–129. https://doi.org/10.1016/j.csi.2018.03.002

Rahmadi, I. F., Lavicza, Z., & Houghton, T. (2021). Towards user-generated microgames for supporting learning: An investigative exploration. *Contemporary Educational Technology, 13*(3), ep299. https://doi.org/10.30935/cedtech/10785

Ramtohul, A., & Khedo, K. K. (2022). An engaging user experience framework for mobile augmented reality. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4174927

Savi, R., von Wangenheim, C. G., & Borgatto, A. F. (2011). A model for the evaluation of educational games for teaching software engineering. In *2011 25th Brazilian Symposium on Software Engineering* (pp. 194–203). IEEE. https://doi.org/10.1109/SBCARS.2011.24

Shoukry, L., Sturm, C., & Galal-Edeen, G. H. (2015). Pre-MEGa: A proposed framework for the design and evaluation of preschoolers' mobile educational games. In K. Elleithy et al. (Eds.), *Innovations and Advances in Computing, Informatics, Systems Sciences, Networking and Engineering* (Lecture Notes in Electrical Engineering, Vol. 313, pp. 385–390). Springer. https://doi.org/10.1007/978-3-319-06773-5_52

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Vaske, J. J., Beaman, J., & Sponarski, C. C. (2017). Rethinking internal consistency in Cronbach's alpha. *Leisure Sciences, 39*(2), 163–173. https://doi.org/10.1080/01490400.2015.1127189

Yasin, S. N. S., & Ibrahim, R. (2025). MEGUX: A UX model for mobile educational game evaluation. In *ICHORA 2025 – 7th International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings*. IEEE.

https://doi.org/10.24191/mij.v6i2.9179