

International Journal of Learning, Teaching and Educational Research
 Vol. 25, No. 6, pp. 983-1003, June 2026
<https://doi.org/10.26803/ijlter.25.6.42>
 Received Apr 8, 2026; Revised Jun 15, 2026; Accepted Jun 18, 2026

Psychometric Validation of the Game Experience Questionnaire (GEQ) Among Adolescents

Yahya M. Khatatbeh*  and Bandar S. Alalam 
 Imam Mohammad Ibn Saud Islamic University (IMSIU),
 Riyadh, Saudi Arabia

Abstract. The growing integration of digital gaming into adolescents' everyday lives has important implications for education, learning, and student development, as game-based environments increasingly intersect with motivation, engagement, and digital behavior in school-age populations. However, research in Arab educational contexts lacks psychometrically sound instruments for assessing students' multidimensional game experiences. This study aimed to validate the Arabic version of the Game Experience Questionnaire (GEQ) among adolescents within an Arab educational context. A cross-sectional psychometric validation design was employed with 380 middle-school students. The Arabic GEQ was examined using exploratory and confirmatory factor analyses, reliability testing through Cronbach's alpha and McDonald's omega, and assessment of convergent, discriminant, and gender-based measurement invariance. The findings supported a refined four-factor, 19-item structure consisting of Competence, Immersion, Positive Affect, and Involvement. The model demonstrated acceptable fit ($\chi^2/df = 2.79$, RMSEA = 0.069, SRMR = 0.058), satisfactory reliability ($\alpha = 0.758-0.857$; $\omega = 0.767-0.886$), and acceptable construct validity. Full configural, metric, and scalar invariance across gender was established. The findings indicate that the Arabic GEQ provides a reliable, valid, and culturally appropriate measure of adolescent game experience across multiple dimensions of engagement. The study contributes to educational measurement by providing a reliable and culturally appropriate instrument for assessing adolescent game experience in Arab school contexts. The findings offer practical value for researchers, school counselors, and practitioners seeking to understand how digital gaming experiences relate to student engagement and psychosocial functioning.

Keywords: Game experience; Adolescents; Psychometric validation; Reliability; Measurement invariance; Arabic version

Citation:
 Khatatbeh, Y. M., &
 Alalam, B. S. (2026).
 Psychometric Validation
 of the Game Experience
 Questionnaire (GEQ)
 Among Adolescents.
*International Journal of
 Learning, Teaching and
 Educational Research*,
 25(6), 983-1003.
<https://doi.org/10.26803/ijlter.25.6.42>

*Corresponding Author: Yahya M. Khatatbeh; ymkhatatbh@imamu.edu.sa

1. Introduction

Digital gaming has become one of the largest global leisure industries and has important cultural and developmental implications, with approximately 2.7–3 billion active players worldwide (Stevens et al., 2021; Rebhi et al., 2023). Recent evidence continues to demonstrate the growing relevance of gaming experiences for adolescent well-being, motivation, and psychosocial development in both educational and recreational contexts (Paquin et al., 2025; Luján-Barrera et al., 2025). The academic community has transitioned from using quantitative data about how often players engage in gaming to the qualitative nature of the psychology behind what drives players' interest in gaming due to the massive increase in size and complexity of gaming systems made available through technological access and connectivity.

As such, adolescents have been identified as being particularly significant within this area due to their high levels of engagement with digital gaming environments along with their unique developmental characteristics related to excessive usage of digital games and heightened sensitivity to rewards. Adolescents are one of the fastest growing demographics in gaming studies, and therefore represent a population of interest for researchers investigating this demographic group. It is crucial to consider the importance of research in adolescent education since digital gaming has become more connected with learning and motivation processes. Thus, investigating the experiences that adolescents have with digital games may prove beneficial for educational engagement and learning environment design.

Therefore, for both theoretical models and empirical studies, a complete comprehension of the subjective psychological experience involved in playing video games (i.e., beyond just frequency of behavior) is necessary. Studies have emphasized that subjective game experiences, including immersion, emotional engagement, and perceived competence, are central predictors of player motivation, continued engagement, and psychological outcomes among adolescents and young adults (Dombrovskis et al., 2025; Luján-Barrera et al., 2025).

For example, multiple components of game experience include immersion, flow, perceived competence, affective response, tension, and challenge. A well-established measurement tool for assessing these aspects and other areas of game experience is the Game Experience Questionnaire (GEQ) (Law et al., 2018); however, in order to adapt psychological tools to be applicable across cultures, it is also required to demonstrate conceptually, metrically, and structurally similar evidence across different cultural contexts (Milfont & Fischer, 2010; Fischer et al., 2025).

The process of culturally adapting psychological measures is more involved than simply translating the instrument into another language. Cultural differences can affect the way items are interpreted, how respondents respond to those items, and ultimately how well the respondent's responses capture their underlying construct(s), potentially affecting both the factorial integrity and the measure's validity. The use of factorial analysis, reliability indices, and assessments of

measurement invariance in evaluating an instrument across new languages and cultures has been recommended as appropriate methodological approaches (Vandenberg & Lance, 2000; Chen, 2007). Failure to follow these methodologies could result in biased or non-equivalent comparisons across groups. Therefore, there remains a need for comprehensive validation studies among Arabic-speaking adolescents utilizing contextually relevant psychometrics; although the GEQ continues to be utilized increasingly within the field of international research.

While there have been prior validations of the GEQ in Arabic-speaking settings (Rebhi et al., 2023), very little has been done to investigate its factor stability for adolescents, measure invariance in gender groups, and inclusion within theoretically based associations between game experience variables. The present study extends previous validation research by examining these three aspects simultaneously. Studies by Phan et al. (2016), Rebhi et al. (2023), and Rebhi et al. (2023) examining adaptations of the game experience measures provided evidence of satisfactory factor structures and reliability, although the results also indicated significant effects of gender, age and game type on questionnaire outcomes.

1.1 Game Experience Questionnaire (GEQ)

The Game Experience Questionnaire (GEQ) was developed by (Rebhi et al., 2023) to assess players' subjective gaming experiences across emotional, cognitive, and experiential dimensions – immersion, flow, competence, positive affect, negative affect, tension, and challenge. The original core version consists of 33 items distributed across seven dimensions. Research on video games has made extensive use of the GEQ as a benchmark for evaluating players' experiences (IJsselsteijn et al., 2013; Law et al., 2018). studies have further highlighted the importance of user-centered approaches and immersive experiences in understanding player engagement, while emphasizing the need to refine assessment frameworks to accommodate evolving gaming environments and diverse user populations (Da Silva et al., 2026; Xuning & Jimin, 2026).

Concerning the factorial stability of the GEQ, empirical tests have produced conflicting results. Although the seven-factor structure was first validated (IJsselsteijn et al., 2013) further systematic reviews found that model fit indices were inconsistent and that factor retention varied across research (Law et al., 2018). The results of the comparative analyses conducted by Denisova et al. (2016) indicate that the dimensions of tension and negative affect, in particular, display lower factor loadings or cross-loadings in samples of adolescents.

Additionally, in order to attain acceptable fit indices, such as CFI and RMSEA values nearing suggested criteria, cross-cultural adaptations may need item reduction or structure alteration (Rebhi et al., 2023). These findings suggest that it is of considerable importance to reexamine the dimensional structure of the GEQ across different developmental and cultural contexts, as they indicate that the theoretical framework underlying the GEQ may not be universally invariant.

The components that make up GEQ include a core questionnaire, social presence module, post-game module, and an abbreviated in-game version designed for repeated assessment during gameplay. The core GEQ consists of Likert-type items that measure immersion, flow, competence, positive affect, negative affect, tension, and challenge (Brockmyer et al., 2009; Denisova et al., 2016; Sabet et al., 2019). Although the GEQ is not intended as a diagnostic instrument, it may provide valuable contextual information for research on problematic gaming behaviors (American Psychiatric Association, 2013). The Social Presence Module assesses the psychological and behavioral dimensions of interactions with co-located players, online opponents, and in-game characters.

The Post-Game Module evaluates players' experiences following gameplay, including residual emotional states, fatigue, and the process of reorientation to everyday reality. This module has proven useful in both naturalistic gaming research and experimental studies investigating post-game effects (Damaiyanti et al., 2025; Stevens et al., 2021; Khatatbeh et al., 2025; Sinnari, 2022). Consequently, the GEQ serves as a descriptive and explanatory instrument that may contribute to clinical assessment and diagnostic processes by providing insights into players' subjective gaming experiences (Iffaturrahman & Purnamasari, 2025).

Taking into account evidence that a minority of adolescents may develop problematic or excessive gaming behaviors, with global prevalence estimates generally ranging from 1% to 3%, the accurate assessment of gaming experiences is particularly important (Kim et al., 2022; Stevens et al., 2021). Evaluating experiential dimensions such as immersion and flow may provide valuable contextual insights into patterns of intensive gaming involvement, even though the GEQ is not intended to serve as a diagnostic instrument. However, several important research gaps remain despite the widespread use of the GEQ. First, the stability of the GEQ dimensions among adolescent populations, particularly within non-Western cultural contexts, remains insufficiently understood, as previous studies have reported variations in the scale's factor structure and inconsistencies in model fit indices.

Second, comprehensive investigations of measurement invariance across gender remain limited, thereby constraining confidence in the comparability of scores between male and female participants. Third, relatively few studies have examined the psychometric stability and measurement invariance of the GEQ among adolescent populations across diverse cultural contexts. It is essential to address this gap in order to assess if the GEQ maintains its structural integrity as a theory or needs modifications for application with Arab youth. Accordingly, the present study reassessed the factor structure, reliability, construct validity, and gender equivalence of the Arabic Version of the GEQ using sequential exploratory and confirmatory model testing procedures.

2. Methods

2.1 Study design

The purpose of this cross-sectional research on the psychometric characteristics of an Arabic-language version of the Game Experience Questionnaire (GEQ) is to

assess the structural characteristics of the Arabic-language version of the GEQ for use with adolescent populations. To validate the structural properties and psychometric adequacy of the Arabic GEQ, exploratory factor analysis (EFA), confirmatory factor analysis (CFA), reliability evaluation, assessment of construct validity, and examination of measurement invariance among males and females were used.

2.2 Participants

In total, 380 middle school students participated in the study. The target population comprised Arabic-speaking adolescents enrolled in middle schools who regularly engaged in digital gaming activities. This population was selected because adolescence represents a developmental period characterized by increased engagement with digital media and immersive online environments.

A volunteer convenience sample was used to gather data for this research. Permission was obtained from participating schools to contact students via school-based distribution methods (e.g., school newsletter or e-mail) to complete a voluntary online survey. Before students began the questionnaire they were provided with information about the nature of the research, that all student responses would be kept anonymous and confidential; and they were given the opportunity to withdraw from the study at any time during the process without penalty.

The inclusion criteria required participants to meet the following conditions: (a) enrollment in middle school, (b) inclusion within the specified adolescent age range, (c) regular engagement in digital gaming activities, and (d) completion of all study measures. Participants were excluded if they did not engage in digital gaming, submitted incomplete questionnaires, or provided responses containing substantial missing data that could compromise the validity of the statistical analyses. Following the screening process, a final sample of 380 adolescents was retained for analysis.

Table 1: Frequency Distribution of Participants' Demographic Characteristics and Main Game Type (N = 380)

Variable	Category	n	%
Gender	Female	196	51.6
	Male	184	48.4
Age Group	9-11 years	14	3.7
	12-14 years	275	72.4
	14- 17 years	91	23.9
Main Game Type	Esports (Competitive Games)	146	38.4
	Single-player Games	40	10.5
	Online Multiplayer Games	194	51.1
Educational Stage	Middle School	380	100.0

The demographic characteristics of the adolescents who participated in this study are presented in Table 1 (N = 380). Gender representation for this sample was approximately equally distributed, as there were 51.6 % female participants to 48.4 % male participants. Approximately 72.4 % of all participants ranged from

ages 12-14, while 23.9 % of the total number of participants fell into the range of ages 14-17. Online multiplayer games accounted for the largest proportion (51.1%), followed by esports games (38.4%). Lastly, all participants were students who attended middle school; therefore, it can be assumed that these individuals were situated in a similar educational environment.

The participants were grouped by age and their preferred form of gaming to better describe them. The age groupings were made to reflect the different development stages of adolescents. The gaming categories were formed using participants' self-identified most often played types of games (esports/games that compete, single player, and online multiplayer) and were used as an additional way to gain a more complete understanding of the sample as well as to identify if there are significant demographic differences in the way participants experienced gaming.

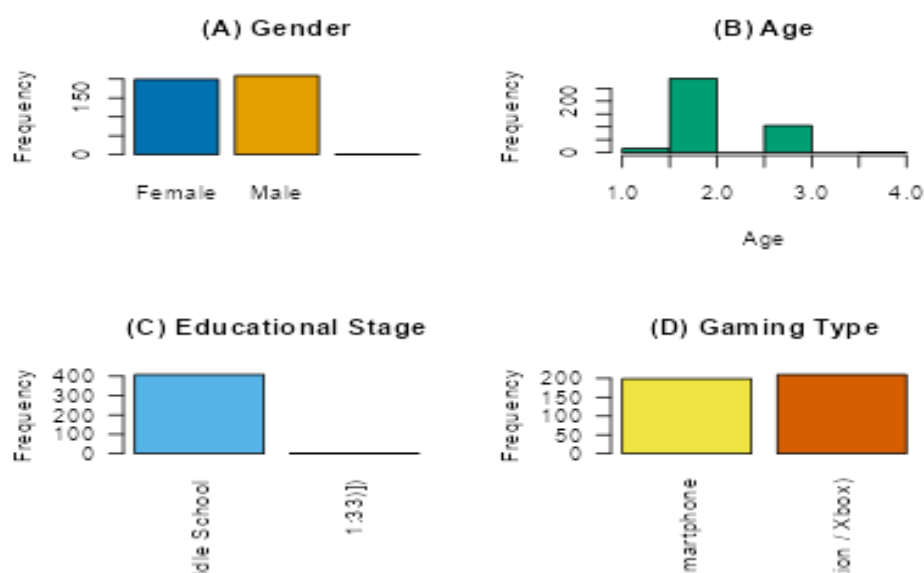


Figure 1: Demographic Characteristics and Primary Game Type Distribution Among Adolescent Participants (N = 380)

2.3 Measurement

In addition to other factors, several demographic variables were integrated in the research design and statistical analysis for this study. Additionally, these demographic variables were analyzed to provide an overall view of the sample and identify any differences that existed among participants by their respective demographics. The selected demographic indicators were gender, middle school as the educational stage, adolescent development stages as the age group, and primary game type.

Gender was used as a variable because it has been shown through numerous studies that gender influences both gaming behaviors and the psychological process that leads to gaming engagement. Middle school students also provide a unique context for understanding how education impacts where participants are located while collecting the data. Primary game type (such as competitive/esports, online multi-player and all other types of gaming) was

considered in order to identify if there are differing levels of engagement for each type of gameplay. These demographic variables provide an overall picture of who participated in this study, therefore support the interpretation of the results from the psychometrics from each group.

2.3.1 Game Experience Questionnaire (GEQ)

The Game Experience Questionnaire (GEQ) was developed by IJsselstein et al. (2013). It was designed to measure the different ways players experience emotions while playing games through both emotional experiences (affective reaction) and mental focus and attention (cognitive involvement) as well as their sense of being "in" the game (immersive engagement). The original GEQ consists of 33 items divided into seven dimensions: Immersion (six items), Flow (five items), Competence (five items), Positive Affect (five items), Negative Affect (four items), Tension/ Anger (three items), and Challenge (five items). Each item measures how much of each state the player felt when they were gaming using a five-point Likert scale with values of 0 (the least amount - not at all) to 4 (the most amount - extremely).

Scores for sub-scales are computed by averaging responses from items measuring each of the dimensions. A higher score on a particular sub-scale suggests a greater degree of endorsement of the experience-related concept being measured. Due to its multi-dimensional structure, it is typically discouraged to compute an overall (total) score. Prior validation studies have established that the original GEQ has sufficient factorial validity; as well as exploratory and confirmatory factor analysis results support its multi-dimensional nature. Reliabilities among the various sub-scales were also all at or above threshold values for reliability, supporting the notion of acceptable internal consistency. This scale is most often applied within the areas of Digital Game Research and Human-Computer Interaction to measure engagement with a digital game, emotional response to the game, level of immersion experienced while playing the game and perceptions of one's ability to play the game.

For the purpose of this study, the Game Experience Questionnaire (GEQ) was translated and culturally adapted into Arabic using established cross-cultural adaptation procedures. A forward-backward translation procedure was employed to ensure semantic and conceptual equivalence between the original and translated versions of the instrument. The original English version was translated into Arabic by bilingual specialists with expertise in psychology and digital media. Subsequently, an independent bilingual expert, who was blinded to the original instrument, performed the back-translation into English. Any discrepancies identified between the original and back-translated versions were reviewed and resolved through discussion and consensus.

Content validity was determined using an expert panel of psychologists, educators who assess student learning, and linguists in Arabic. These experts reviewed the translation to determine if the language used was clear, culturally relevant to adolescents, appropriate in terms of age, and conceptually equivalent

to the original version. Minor changes to some item wordings were made to make them more understandable, and culturally contextual.

Factor analyses, including both exploratory factor analysis and confirmatory factor analysis, were employed to evaluate the psychometric properties of the Arabic version of the instrument. The final retained version comprised 19 items distributed across four dimensions: Competence (5 items), Immersion (4 items), Positive Affect (6 items), and Involvement (4 items). All items were assessed using a five-point Likert scale ranging from 0 ("not at all") to 4 ("extremely").

The subscales of this survey were constructed by calculating an average response from all questions for each of these dimensions; the resulting averages were used as indicators of how strongly a respondent endorsed their gaming experiences in each of the respective constructs. Therefore, consistent with the multi-dimensional nature of the instrument, respondents' individual subscale scores were examined separately instead of using a total score. The detailed analysis of the statistical properties (i.e., reliability and validity) of this survey can be found below in the "Results" section.

2.4 Statistical analysis

Statistical analyses were conducted using IBM SPSS for descriptive statistics and structural equation modeling (SEM). Descriptive statistics were employed to provide a summary description of each participant's demographics and gaming characteristics. Descriptive statistics were also employed to describe participants' demographic and gameplay features. The factor structure of the Arabic version of the Game Experience Questionnaire (GEQ) was evaluated through exploratory factor analysis; principal axis factoring with promax rotation. The KMO measure and Bartlett's sphericity test were used to check whether the samples were adequately representative. Eigenvalues larger than one and parallel processing led to the retention of factors.

Then, the four-factor model was evaluated using Confirmatory Factor Analysis (CFA). To assess the model's fit, the χ^2/df , CFI, TLI, RMSEA, and SRMR indices were used. Using Cronbach's alpha and McDonald's omega coefficients, we evaluated the internal consistency. Examining discriminant validity, we used the Fornell-Larcker criteria and the Heterotrait-Monotrait (HTMT) ratio, and convergent validity, we utilized Average Variance Extracted (AVE). Statistical significance was established at $p < .05$.

The primary objective of the present study was to evaluate the psychometric properties of the Arabic version of the GEQ. Accordingly, the statistical analyses were directed toward examining the instrument's factorial structure, reliability, construct validity, and measurement invariance. Consequently, the findings reported in this study are limited to psychometric validation procedures intended to assess the measurement properties of the scale within the target population.

3. Result

3.1 What are the psychometric properties of the Game Experience Questionnaire (GEQ) among adolescents?

Table 2: Reliability and Convergent Validity Indices of the GEQ Among Adolescents

Dimension	No. of Items	AVE	McDonald's ω	Cronbach's α
Competence	5	0.551	0.860	0.857
Immersion	4	0.478	0.788	0.777
Positive Affect	6	0.361	0.771	0.771
Involvement	4	0.450	0.767	0.758
Total Scale	—	—	0.886	0.861

As shown in Table 2, the retained dimensions demonstrated satisfactory reliability and convergent validity. The Arabic GEQ demonstrated satisfactory psychometric properties among adolescents, according to the results.

Table 3: Exploratory Factor Analysis Loadings for the Refined Four-Factor GEQ

Item	Competence	Immersion	Positive Affect	Involvement
C1	.78	—	—	—
C2	.81	—	—	—
C3	.75	—	—	—
C4	.73	—	—	—
C5	.84	—	—	—
I1	—	.72	—	—
I2	—	.76	—	—
I3	—	.69	—	—
I4	—	.74	—	—
PA1	—	—	.68	—
PA2	—	—	.71	—
PA3	—	—	.65	—
PA4	—	—	.73	—
PA5	—	—	.70	—
PA6	—	—	.67	—
INV1	—	—	—	.72
INV2	—	—	—	.75
INV3	—	—	—	.69
INV4	—	—	—	.77

Table 3 shows the exploratory factor analysis revealed a clear four-factor structure. Fourteen items from the original 33-item GEQ were removed due to low loadings (<0.40) or cross-loadings. The remaining 19 items were organized into four coherent dimensions: competence (5 items), immersion (4 items), positive affect (6 items), and involvement (4 items). All retained items had satisfactory loadings and conceptual consistency, which supported the structural refinement of the Arabic GEQ for adolescents.

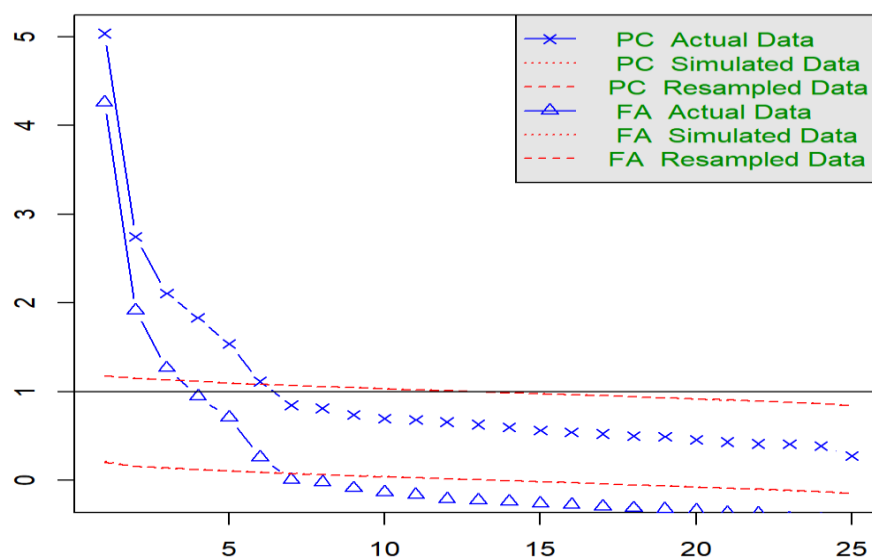


Figure 2: Scree Plot and Parallel Analysis for Determining the Number of Factors

A comparison of the observed and parallel-generated eigenvalues is shown in Figure 2 in the scree plot. According to the findings, the eigenvalues of the first four observed eigenvectors were higher than the corresponding simulated eigenvalues obtained from random data. The observed eigenvalues dropped below the simulated cutoff after factor 4.

3.2 What is the factorial structure of the Game Experience Questionnaire (GEQ) among adolescents? Does the data support the original theoretical model, or does it suggest a more appropriate alternative model?

Table 4: Confirmatory Factor Analysis Fit Indices for the Four-Factor Model

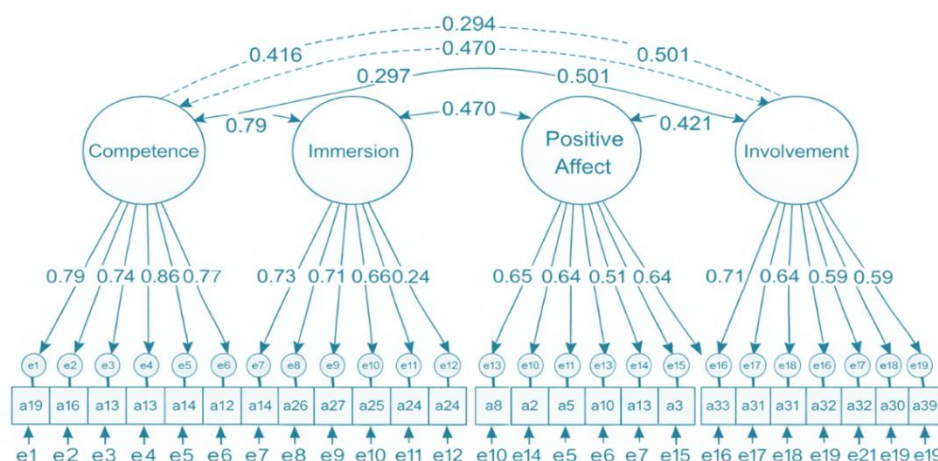
Fit Index	Value	Recommended Threshold	Interpretation
χ^2	407.170	—	Significant
df	146	—	—
χ^2/df	2.79	≤ 3.00	Acceptable
CFI	0.894	≥ 0.90	Marginally acceptable
TLI	0.876	≥ 0.90	Slightly below threshold
RMSEA	0.069	≤ 0.08	Acceptable
RMSEA 90% CI	0.061–0.077	—	Within acceptable range
SRMR	0.058	≤ 0.08	Good
GFI	0.902	≥ 0.90	Acceptable

The confirmatory factor analysis demonstrated an adequate overall fit of the four-factor model. The relative chi-square ($\chi^2/\text{df} = 2.79$) remained within the advised thresholds, indicating adequate model parsimony. The RMSEA (0.069) and SRMR (0.058) values indicated an acceptable residual fit. Despite the CFI (0.894) and TLI (0.876) being marginally below the conventional threshold of 0.90, they approached acceptable levels and were corroborated by the sufficiency of other fit indices. These findings collectively suggest that the four-factor structure adequately represents the dimensions of GEQ among adolescents.

Table 5: Standardized Factor Loadings from Confirmatory Factor Analysis

Item	Factor	Standardized Loading
C1	Competence	.82
C2		.85
C3		.79
C4		.74
C5		.88
I1	Immersion	.76
I2		.81
I3		.73
I4		.78
PA1	Positive Affect	.70
PA2		.75
PA3		.67
PA4		.77
PA5		.72
PA6		.69
INV1	Involvement	.74
INV2		.78
INV3		.71
INV4		.80

The standardized factor loadings, as displayed in Table 5, were all significantly higher than the suggested cutoff of .50 and varied from moderately high to very high values. The results support the structural coherence of the four-factor model and prove that the indicators are reliable enough. Overall, the results support a refined four-factor structure of the GEQ within the adolescent sample.

**Figure 3: The confirmatory factor model derived from the current dataset**

The final GEQ confirmatory factor model for adolescents, consisting of four factors, is shown in Figure 3. The indicator reliability across the four latent constructs—Competence, Immersion, Positive Affect, and Involvement—was satisfactory, as all standardized factor loadings were statistically significant and ranged from moderate to high values. Although there is some conceptual overlap between the dimensions, their moderate inter-factor correlations indicate that

they are still empirically separate. All things considered, the model proves that the GEQ is multidimensional in the adolescent group and that its structure is coherent.

3.3 What are the reliability indicators and construct validity evidence (convergent and discriminant validity) of the Game Experience Questionnaire (GEQ) among adolescents?

Table 6: Intercorrelations Among the Four GEQ Dimensions

Variable	1	2	3	4
1. Competence	—			
2. Immersion	.240	—		
3. Positive Affect	.463	.387	—	
4. Involvement	.328	.440	.408	—

The results from the reliability analysis indicated an acceptable level of internal consistency on all the dimensions retained for the GEQ. The Cronbach's alpha coefficient scores ranged from .758 to .857, while McDonald's omega coefficients ranged from .767 to .886, which is above the acceptable value of .70. This indicates that all the four factors measured the latent construct.

Convergent validity was tested through the average variance extracted (AVE). The AVE values ranged from .361 to .551 on the four dimensions. Even though the AVE values for the Immersion, Positive affect, and Involvement were less than .50, the reliability coefficients surpassed .70. In accordance with Fornell and Larcker (1981), it can be said that even though AVE is less than .50, convergent validity is acceptable as long as composite reliability remains adequate.

Discriminant validity was assessed using inter-factor correlations, Fornell-Larcker criteria, and Heterotrait-Monotrait Ratio. As depicted in Table 6, there is no correlation among the factors that exceed the value of .70, thus showing the relations of factors but not being empirically similar. In addition, HTMT values are below .85 for all the dimensions, while the square roots of AVE are greater than inter-factor correlations.

Table 7: Fornell-Larcker Discriminant Validity Matrix

Factor	Competence	Immersion	Positive Affect	Involvement
Competence	0.74			
Immersion	0.52	0.69		
Positive Affect	0.48	0.55	0.60	
Involvement	0.46	0.50	0.58	0.67

Table 7 shows that, as presented in the Fornell-Larcker matrix, the square roots of AVE (diagonal values) exceed the inter-factor correlations, thereby supporting discriminant validity among the four latent constructs. The discriminant validity is supported by HTMT values less than the strict value of 0.85, which implies that the four factors are distinct from each other in spite of their moderate correlation. Overall, this analysis suggests that the GEQ shows acceptable validity and reliability among adolescents.

3.4 Does measurement invariance of the Game Experience Questionnaire (GEQ) hold across gender (male vs. female) among adolescents

In order to ascertain whether the Game Experience Questionnaire (GEQ) functions equally among male and female adolescents, a hierarchical multi-group confirmatory factor analysis (MG-CFA) was implemented. Sequentially, measurement invariance was assessed at three increasingly restrictive levels: configural invariance (testing the equivalence of the factorial structure), metric invariance (testing the equality of factor loadings), and scalar invariance (testing the equality of item intercepts). The invariance assessment criteria were widely accepted, and model comparisons were conducted based on changes in fit indices (ΔCFI , $\Delta RMSEA$, and $\Delta SRMR$).

Table 8: Hierarchical Measurement Invariance Testing of the GEQ Across Gender

Model	χ^2	df	CFI	RMSEA	SRMR	ΔCFI	$\Delta RMSEA$	$\Delta SRMR$
Configural	840.071	292	.805	.099	.081	—	—	—
Metric	860.803	307	.803	.097	.086	.002	.002	.005
Scalar	895.069	322	.796	.097	.083	.007	.000	.003

As shown in Table 8, while the absolute fit indices for the configural model did not indicate a high level of fit, there was still a rather small change in the fit indices between increasingly restrictive models. Following the current suggestions for measurement invariance testing (Chen, 2007; Vandenberg & Lance, 2000), the results suggest that there is configural, metric, and scalar invariance between gender groups. The application of equality constraints to factor loadings caused only a slight deterioration in model fit ($\Delta CFI = .002$). In addition, applying the same constraints to item intercepts led to a slight deterioration in fit ($\Delta CFI = .007$); however, it was still smaller than the threshold value (.01) needed for invariance testing. Changes in the remaining fit indices, including RMSEA and SRMR, were also very small. Thus, the structure of the four factors identified with the GEQ works equivalently for males and females.

4. Discussion

The current study aimed to psychometrically re-examine the Arabic Game Experience Questionnaire (GEQ) among adolescents. In contrast to the seven-dimensional model, Competence, Immersion, Positive Affect, and Involvement were shown to be four factors (IJsselsteijn et al., 2013). This structural refinement indicates contextual and developmental recalibration, not rejection of the original conception.

4.1 Factorial Refinement and Item Reduction

The deletion of Negative Affect, Tension, and Challenge and the decrease from 33 to 19 elements need careful interpretation. Empirical studies have indicated a lack of stability across time, for example from an adolescent to young adult population, and cross-culturally within these dimensions (Rebhi et al., 2023; Damaiyanti et al., 2025; Law et al., 2018). However, beyond being statistically related via cross-loadings on a measure of scale items, there is a possible developmental explanation for this instability. Cognitive development such as affective differentiation and control matures during adolescence (Casey et al., 2022;

Blakemore & Mills, 2014). With regard to emotionally arousing states in digital environments that provide an engaging opportunity for user involvement (e.g., through immersion), states of stress or frustration can become part of what users view as fun. Players who engage in a variety of game styles (i.e., most common types of games available in our sample) will find themselves competing against other players in competitive and/or online multiplayer formats, which can mask the distinction between feeling challenged and motivated by the task at hand versus experiencing anxiety due to competition. These findings are consistent with previous research indicating that players may experience anxiety and emotional arousal during gameplay while still perceiving the overall gaming experience as enjoyable and rewarding (Bopp et al., 2016).

If no identifiable components of negative emotions were found in the current study, it could suggest that participants integrated their experiences into their evaluations of the video game experience instead of failing to capture them adequately using psychometric measures. The revised interpretation suggests that the experiencing aspects associated with motivation for engagement and positive engaging experienced by Arab adolescents are not based on emotional valences, but rather on interactive elements of the video games. Thus, the findings suggest that researchers studying interactive media may need to modify existing psychometric instruments to suit specific cultural and developmental contexts (Richardson et al., 2025; Denisova et al., 2021).

Research has further demonstrated that affective game design elements play a central role in fostering deep player engagement and immersive experiences, reinforcing the importance of accurately measuring experiential dimensions across diverse gaming populations (Ng et al., 2026).

4.2 Model Fit and Construct Validity

Model fit was established using confirmatory factor analysis. Though CFI and TLI scores were below the standard of 0.90, RMSEA and SRMR scores had values within the recommended range. Modern standards in psychometrics emphasize the importance of assessing model fit through various indices instead of using one criterion (Putnick & Bornstein, 2016; MacCallum & Austin, 2000). Consequently, the assessment of model fit is dependent upon the overall pattern of results obtained from different indices. In the four-factor solution, there was an acceptable level of consistency exhibited by standardized factor loadings and the dimensions of theory. This is because stability of the constructs is evident from the dimensions that have been noted among adolescents and culturally diverse populations in other research (Rebhi et al., 2023; Law et al., 2018).

All reliability coefficients were found to exceed recommended criteria across all retained dimensions, indicating good internal consistency, as supported by the values obtained from Cronbach's α and McDonald's ω . However, even though AVE values for Immersion, Positive Affect, and Involvement were not found to exceed the .50 recommended criterion, reliability coefficients were found to exceed recommended criteria. Thus, this implies that the retained dimensions have good internal consistency even though their shared variances are low.

Nonetheless, results show that improvement can be made in convergent validity by refining selected items. According to Fornell and Larcker (1981), convergent validity can be considered acceptable where composite reliability exceeds recommended criteria. In addition, the observed moderate relationships among the retained factors carry theoretical significance because game play is an experience that involves complex emotions, thoughts, and motivations. Therefore, the presence of a certain amount of overlap could be attributed to the composite nature of immersive game play experience and not poor construct differentiation (Bopp et al., 2021). Overall, the findings point to the sound structure and applicability of the proposed four-factor GEQ.

4.3 Cross-Gender Measurement Invariance

This finding represents an important methodological contribution. psychometric research has to emphasize the importance of evaluating both structural validity and measurement invariance to ensure that comparisons across demographic groups reflect true differences in the underlying construct rather than measurement artifacts (Zhou et al., 2026). The fit indices were all less than or equal to .01, based on Chen's (2007) and Vandenberg and Lance's (2000) recommendations, even though the absolute CFI values were somewhat modest. Given that gaming research frequently presents gender-based findings without conducting invariance tests, this finding represents a methodological contribution. Scalar invariance provides a basis for understanding why there are latent mean differences across groups and supports previous cross-group studies.

The findings suggest that adolescent gaming experiences tend to cluster around dimensions related to engagement rather than discrete emotional categories. This provides a basis for developing an understanding of how adolescents respond developmentally to immersive digital experiences. For researchers and practitioners working with adolescents, the 19 item Arabic version of the GEQ is both more concise and culturally appropriate. Shorter scales are likely to be easier to administer within a school setting or when administering large scale surveys, but should retain adequate levels of psychometric reliability. The establishment of gender-invariant measurements will allow comparison of adolescent gamers' experience (regardless of whether they are male or female) and assess the effectiveness of interventions designed to promote healthy gaming practices.

In an academic context, the validated Arabic version of the GEQ can be useful for the researchers, school counsellors, and educational psychologists interested in understanding how students' experiences related to gaming influences their engagement with learning, motivation to learn and their interactions with technology enhanced learning environment. Lastly, the findings of this study underscore the importance of culturally sensitive psychometric refinement, developmentally informed contextualization, and methodologically rigorous approaches to the assessment of adolescent gaming experiences. Furthermore, the growing adoption of gamification strategies and digital transformation initiatives across contemporary institutions highlights the importance of robust instruments capable of assessing engagement and user experience within evolving digital ecosystems (Alsafran et al., 2026).

5. Conclusion

The current research study assessed the characteristics of the Arabic version of the Game Experience Questionnaire (GEQ) for adolescents. The results from this study provided support for the validity of the Arabic version of the GEQ. The Arabic version of the GEQ showed acceptable structural validity, reliability, and construct validity in its use with adolescent participants.

The reduction in the original seven factor model to four factors implies that there are four different aspects to adolescent gaming experience. Combining the original constructs for Competence, Immersion, Affectivity and Involvement suggest that adolescents' cognitive and emotional experiences while playing games do not represent two completely separate forms of processing. Instead, it seems that gaming experiences represent a spectrum of psychological involvement. Thus, this interpretation supports the current perspective on game engagement as a unified construct rather than as multiple disconnected cognitive and emotional experiences.

The methodological value of this study lies in its demonstration that both culturally-sensitive adaptations of scales and thorough statistical analyses are important. Additionally, the study is an excellent example of a practically useful and easy-to-use assessment tool for measuring psychological effects of digital gaming on students. Both values provide a strong base for further studies and debates about the use of the GEQ as not just another measurement instrument, but as well as a theoretical model that can be used in different contexts to understand adolescents' gaming experiences.

6. Limitations and Recommendations

Although the study's methodology has a number of positive aspects, there are also some significant limitations that need to be taken into account. The study employed several strong psychometric techniques (i.e., exploratory and confirmatory factor analyses; reliability evaluation by means of alpha and omega coefficients; measurement invariance evaluation across genders). In addition, the translation and cultural adaptations were conducted with respect to established standards for achieving linguistic and conceptual equivalence. The study also strengthened the quality of its analytic results by employing exploratory and confirmatory factor analyses, reliability assessments, and measurement invariance tests.

Future studies will need to investigate whether the four-factor structure is stable across different age groups, different types of educational settings, and different cultural contexts by employing longitudinal and cross-cultural research methodologies. Future replication research is suggested as an area for study to increase the applicability of results in a variety of educational environments and culturally based contexts. Lowering AVE values by refining dimensions will allow for improved convergent validity. Further improvement of the measures can be obtained with the inclusion of behavioral indicators and multiple informants. Research involving interventions that target cognitive impulsivity

may also provide insight into how to promote responsible gaming behavior among adolescents.

7. Data availability statement

Upon request, the corresponding author can provide all the data that support the conclusions of this study. Data cannot be made publicly available due to ethical concerns affecting the privacy of human participants. However, it can be shared with qualified researchers upon request.

8. Ethical approval

The study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki and its subsequent amendments. All participants were informed about the purpose and procedures of the study prior to participation. Written parental consent and student assent were obtained prior to participation, and confidentiality of the collected data was strictly maintained. Participation was voluntary, and respondents were assured of their right to withdraw from the study at any stage without any consequences. Data were anonymized prior to analysis to ensure privacy and compliance with institutional and international ethical standards.

9. Funding

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2601).

10. Acknowledgments

The authors acknowledge the support provided by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University for funding this research under Grant No. IMSIU-DDRSP2601.

The authors confirm that no artificial intelligence (AI) tools were used in the writing, editing, or preparation of this paper. The manuscript represents solely the authors' original work and intellectual contributions.

11. References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). <http://doi.org/10.1176/appi.books.9780890425596>
- Alsafran, M. A., Awad, Y. A., Alsaber, A., Almutairi, S., & Al-Kandari, A. (2026). Assessing gamification and corporate sustainability: the role of embracing digital transformation and open innovation in ecosystem. *Management & Sustainability: An Arab Review*, 5(2), 427-451. <http://doi.org/10.1108/MSAR-11-2024-0196>
- Blakemore, S.-J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, 65, 187-207. <http://doi.org/10.1146/annurev-psych-010213-115202>
- Bopp, J. A., Mekler, E. D., & Opwis, K. (2016). Negative emotion, positive experience? Emotionally moving moments in digital games. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2996-3006). Association for Computing Machinery. <http://doi.org/10.1145/2858036.2858227>
- Bopp, J. A., Vornhagen, J. B., & Mekler, E. D. (2021). "My soul got a little bit cleaner": Art experience in videogames. *Proceedings of the ACM on Human-Computer Interaction*, 5(CHI PLAY), 1-19. <http://doi.org/10.1145/3474664>

- Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, 45(4), 624–634. <http://doi.org/10.1016/j.jesp.2009.02.016>
- Casey, B., Simmons, C., Somerville, L., & Baskin-Sommers, A. (2022). Making the sentencing case: Psychological and neuroscientific evidence for expanding the age of youthful offenders. *Annual Review of Criminology*, 5(1), 321–343. <http://doi.org/10.1146/annurev-criminol-030920-012711>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <http://doi.org/10.1080/10705510701301834>
- Zhou, J. Z., Rekola, H. R., Sormunen, M., & Mäki-Opas, T. (2026). Structural Validation and Measurement Invariance of the HLS-Q12 Health Literacy Instrument in Finnish Adults: Comparing Traditional and Alignment Methods. *International Journal of Public Health*. <http://doi.org/10.3389/ijph.2026.1609337>
- Colton, E., McLean, C., Anderson, A., Hanegraaf, L., & Verdejo-Garcia, A. (2026). Gamified Assessment of Cognitive Impulsivity in Eating Disorders and Mental Ill-Health: Mixed Methods Study Incorporating Lived Experience Co-Design and Evaluation. *JMIR Serious Games*, 14, e79784. <https://doi.org/10.2196/79784>
- Damaiyanti, S., Colton, E., McLean, C., Anderson, A., Hanegraaf, L., & Verdejo-Garcia, A. (2026). Gamified Assessment of Cognitive Impulsivity in Eating Disorders and Mental Ill-Health: Mixed Methods Study Incorporating Lived Experience Co-Design and Evaluation. *JMIR Serious Games*, 14, e79784. <http://doi.org/10.2196/79784>
- Baskara, A. R., Noor, M. F., & Maulida, M. A. (2025). Evaluation of user experience in the game “Night’s Reach” using the Game Experience Questionnaire (GEQ). *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (JTIULM)*, 31–42. <http://doi.org/10.20527/jtiulm.v10i1.463>
- Da Silva, F. F., Machado, L. D. S., Leal, G. C. L., & Balancieri, R. (2026). UX Design in Digital Game Development: A Systematic Literature Review and Research Agenda. *IEEEAccess*, 14, 7339–7364. <http://doi.org/10.1109/ACCESS.2026.3652449>
- Denisova, A., Bopp, J. A., Nguyen, T. D., & Mekler, E. D. (2021). “Whatever the emotional experience, it’s up to them”: Insights from designers of emotionally impactful games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Article 120). Association for Computing Machinery. <http://doi.org/10.1145/3411764.3445286>
- Denisova, A., Nordin, A. I., & Cairns, P. (2016). The convergence of player experience questionnaires. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (pp. 33–37). Association for Computing Machinery. <http://doi.org/10.1145/2967934.2968095>
- Dombrovskis, V., et al. (2025). Extending the Game Immersion Questionnaire to Online Users. *Frontiers in Psychology*. <http://doi.org/10.3389/fpsyg.2025.1473821>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <http://doi.org/10.1177/002224378101800104>
- Fischer, R., Karl, J. A., Luczak-Roesch, M., & Hartle, L. (2025). Why We Need to Rethink Measurement Invariance: The Role of Measurement Invariance for Cross-Cultural Research. *Journal of Cross-Cultural Psychology*, 59(2), 147–179. <http://doi.org/10.1177/10693971241312459>
- Iffaturrahman, M., & Purnamasari, D. A. (2025). Evaluation of the user experience of the game Aaron Lost in the Jungle using the Game Experience Questionnaire (GEQ).

- Journal of Applied Multimedia and Networking*, 9(1), 56–67. <http://doi.org/10.30871/jamn.v9i1.10242>
- Khatatbeh, H., Khatatbeh, Y., Pakai, A., & Khatatbeh, M. (2025). Beyond the screen: A social cognitive theory perspective on gender, online gaming, and social well-being in Jordan. *Italian Journal of Pediatrics*, 51(1), 324. <http://doi.org/10.1186/s13052-025-02162-w>
- Kim, H. S., Son, G., Roh, E.-B., Ahn, W.-Y., Kim, J., Shin, S.-H., Chey, J., & Choi, K.-H. (2022). Prevalence of gaming disorder: A meta-analysis. *Addictive Behaviors*, 126, 107183. <http://doi.org/10.1016/j.addbeh.2021.107183>
- Law, E. L.-C., Brühlmann, F., & Mekler, E. D. (2018). Systematic review and validation of the Game Experience Questionnaire (GEQ): Implications for citation and reporting practice. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play* (pp. 257–270). Association for Computing Machinery. <http://doi.org/10.1145/3242671.3242683>
- Luján-Barrera, A., Cervera-Ortiz, L., & Chóliz, M. (2025). In-game need satisfaction, frustration, and gaming addiction patterns across subgroups of adolescents through structural equation modeling: Cross-sectional and instrument validation study of the Youth Gaming Experience Scales. *JMIR Serious Games*, 13, e63612. <http://doi.org/10.2196/63612>
- MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, 51(1), 201–226. <https://doi.org/10.1146/annurev.psych.51.1.201>
- Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3(1), 111–130. <http://doi.org/10.21500/20112084.857>
- Ng, Y.Y., Nathan, R.J. & Khong, C.W. Affective game design elements for deep player engagement in indie games. *Univ Access Inf Soc* 25, 15 (2026). <http://doi.org/10.1007/s10209-025-01291-7>
- Paquin, V., Tremblay, J., Geoffroy, M. C., & colleagues. (2025). Understanding the game experiences and mental health of youth: Protocol for the Game-in-Action Quebec cohort study. *BMJ Open*, 15(9), e103685. <http://doi.org/10.1136/bmjopen-2025-103685>
- Phan, M. H., Keebler, J. R., & Chaparro, B. S. (2016). The development and validation of the Game User Experience Satisfaction Scale (GUESS). *Human Factors*, 58(8), 1217–1247. <https://doi.org/10.1177/0018720816669646>
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90. <http://doi.org/10.1016/j.dr.2016.06.004>
- Putra, G. A., & Kurniawati, M. (2026). Enhancing Gamer Loyalty: Role of Immersiveness and Parasocial Relationship with NPC Characters in Narrative Games. *Advances in Business & Industrial Marketing Research*, 4(2), 152–166. <http://doi.org/10.60079/abim.v4i2.794>
- Rebhi, M., Ben Aissa, M., Tannoubi, A., Saidane, M., Guelmami, N., Puce, L., Chen, W., Chalghaf, N., Azaiez, F., & Zghibi, M. (2023). Reliability and validity of the Arabic version of the Game Experience Questionnaire: Pilot questionnaire study. *JMIR Formative Research*, 7, e42584. <http://doi.org/10.2196/42584>
- Richardson, G. B., Bates, D. G., McLaughlin, L. E., McGee, N., Tse, W. W.-Y., & Lai, M. H. (2025). Are higher-order constructs in evolutionary psychology attributable to omitted cross-loading bias? An exploratory structural equation modeling approach. *Human Nature*, 36(2), 257–280. <https://doi.org/10.1007/s12110-025-09497-7>
- Sabet, S. S., Griwodz, C., & Möller, S. (2019). Influence of primacy, recency and peak effects on the game experience questionnaire. In *Proceedings of the 11th ACM Workshop on*

- Immersive Mixed and Virtual Environment Systems* (pp. 22–27). Association for Computing Machinery. <http://doi.org/10.1145/3304113.3326113>
- Sinnari, D. (2022). *The effect of cultural narrative and audiovisual design on player immersion: An experimental approach*. SSRN. <http://doi.org/10.2139/ssrn.5482183>
- Stevens, M. W., Dorstyn, D., Delfabbro, P. H., & King, D. L. (2021). Global prevalence of gaming disorder: A systematic review and meta-analysis. *Australian & New Zealand Journal of Psychiatry*, 55(6), 553–568. <http://doi.org/10.1177/0004867420962851>
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4–70. <http://doi.org/10.1177/109442810031002>
- Xuning, Y., & Jimin, Z. (2026). Effects of Gamification and Immersive Technology on Museum Learning: A User-Centered Empirical Study. *International Journal of Research and Innovation in Social Science (IJRISS)*, 10(2). <http://doi.org/10.47772/IJRISS.2026.10200350>

Appendix 1

Appendix A: Final Structure of the Arabic GEQ (Adolescent Sample)

Item No.	Dimension	Item Statement
1	Competence	I felt skillful
2		I was good at it
3		I felt competent
4		I was fast at reaching the game's targets
5		I felt successful
6	Immersion	I was fully occupied with the game
7		I lost track of time
8		I forgot about everything around me
9		I was deeply concentrated in the game
10	Positive Affect	I felt content
11		I felt happy
12		It felt good
13		I thought it was fun
14		I enjoyed it
15		I felt satisfied
16	Involvement	I was interested in the game's story
17		I felt imaginative
18		I felt that I could explore things
19		It felt like a rich experience

The Arabic version consists of 19 items of the Game Experience Questionnaire (GEQ), with 5 questions pertaining to competence, 4 to immersion, 6 to positive affect, and 4 to involvement. From 0 (not at all) to 4 (very much), a 5-point Likert scale is used for each item; higher values indicate a stronger endorsement of the described experiential state. At the subscale level, scoring takes place. A mean score is obtained for each dimension by adding up the responses to all of the items in that dimension and then dividing that total by the number of items in that dimension. This method guarantees that subscales with different item counts can be compared with one another.

The improved four-factor version does not use reverse-coding on any items. The subscale scores that come out of it are on a scale from 0 to 4, with higher scores indicating more psychological involvement, stronger positive affect, deeper immersion, and perceived competence. Since each dimension stands for a unique experiential construct, it is not recommended to calculate an overall total score, in keeping with the instrument's multidimensional nature. Subscale profiles rather than overall scores should be the center of attention when making interpretations.