Edu Sportivo
Indonesian Journal of Physical Education
e-ISSN 2745-942X
Journal Homepage: https://journal.uir.ac.id/index.php/SPORTIVO

RESEARCH ARTICLE

Cross-cultural adaptation and psychometric validation of the Game Performance Assessment Instrument (GPAI) for team-sport performance in Indonesian physical education

[1abc]Ega Trisna Rahayu*, [1bcd]Kamal Prihandani, [1bcd]Rolly Afrinaldi, [1bcd]Rina Syafrida, & [2bcd]Farruh Ahmedov

Department of Physical Education, Faculty of Teacher Training and Education, Universitas Singaperbangsa Karawang, Bekasi, Indonesia[1]
Department of Coordination of Scientific Works, Samarkand State University, Samarqand, Uzbekistan[2]

Received 05 June 2026; Accepted 22 April 2026; Published 31 August 2026

OPEN ACCESS

ABSTRACT

Background: Assessing students' game performance requires observation-based instruments that capture tactical behaviours during authentic gameplay. Although the Game Performance Assessment Instrument (GPAI) is widely used, evidence of its linguistic and contextual adaptation for Indonesian physical education remains limited. **Objectives:** This study aims to culturally adapt the GPAI and evaluate its psychometric properties for assessing students' game-performance behaviours in team sports. **Methods:** A cross-sectional instrument-validation study was conducted with 200 junior and senior high school students. The adaptation process involved forward translation, synthesis, back translation, expert review, and pilot testing. Content validity was assessed using the Content Validity Ratio (CVR), construct validity using confirmatory factor analysis (CFA), internal consistency using Cronbach's alpha, and inter-rater reliability using the intraclass correlation coefficient (ICC). **Results:** The CFA demonstrated acceptable model fit (χ^2/df = 2.18, CFI = 0.93, RMSEA = 0.06), with standardised factor loadings ranging from 0.71 to 0.86. Cronbach's alpha coefficients ranged from 0.74 to 0.85, while ICC coefficients ranged from 0.78 to 0.87, indicating satisfactory internal consistency and inter-rater reliability. **Conclusion:** The Indonesian GPAI demonstrated preliminary evidence of acceptable validity and reliability for assessing game-performance behaviours among secondary-school students in team-sport contexts. Further validation using broader populations and independent samples is warranted.

Keywords: Game performance assessment instrument; game performance; tactical behaviour; physical education; team sports

***Corresponding Author**
ega.trisna.rahayu@staff.unsika.ac.id

10.25299/es:ijope.2026.vol7(2).22855

How to Cite: Rahayu, E. T., Prihandani, K., Afrinaldi, R., Syafrida, R., & Ahmedov, F. (2026). Cross-cultural adaptation and psychometric validation of the Game Performance Assessment Instrument (GPAI) for team-sport performance in Indonesian physical education. *Edu Sportivo: Indonesian Journal of Physical Education, 7*(2), 201-217. https://doi.org/10.25299/es:ijope.2026.vol7(2).22855

Authors' Contribution: a – Study Design; b – Data Collection; c – Statistical Analysis; d – Manuscript Preparation; e – Funds Collection

INTRODUCTION

Team sports provide an important context for physical education (PE) because students are required to coordinate technical execution, tactical decision making, and positional behaviour under dynamic game conditions. Contemporary game-centred approaches to PE emphasise the integration of tactical knowledge, decision making, and skill execution rather than isolated technical performance (Barba-Martín et al., 2020; González-Valero et al., 2024; Wang et al., 2024). Recent systematic evidence has shown growing research interest in game-centred approaches such as the Tactical Games Model (TGM), particularly in school-based team and invasion sports, although evidence concerning cognitive and decision-making outcomes remains comparatively limited (Wang et al., 2024). Accordingly, assessment during authentic gameplay is important for determining how effectively students apply tactical knowledge, make decisions, and execute skills under dynamic game constraints (Pan et al., 2023).

Traditional PE assessment has often emphasised technical skill execution through structured or isolated tasks. Although such approaches can provide useful information about motor proficiency, they may not adequately capture the tactical and decision-making demands that emerge during actual gameplay (Breed et al., 2025). As PE pedagogy increasingly incorporates representative game environments, assessment should similarly capture what students do, decide, and execute within the game context rather than focusing solely on isolated technical performance (Manso-Lorenzo et al., 2024; Silva et al., 2023). This creates an important measurement challenge because observation-based assessment requires clearly defined behavioural criteria that can be applied consistently across students, game situations, and educational settings (Barquero-Ruiz et al., 2020; Manso-Lorenzo et al., 2024; Wilkie et al., 2023).

The Game Performance Assessment Instrument (GPAI), developed by Oslin et al. (1998), provides an observation-based framework for evaluating students' performance during authentic gameplay. The GPAI was originally designed to assess multiple dimensions of game performance, including base, adjust, decision making, skill execution, support, cover, and guard. These components collectively address positioning, adaptation to changing game situations, tactical decision making, skill execution, offensive support, defensive coverage, and defensive play. The instrument has subsequently been applied in various game-based PE contexts, and recent systematic evidence confirms its continued use in school physical education (Manso-Lorenzo et al., 2024). However, the interpretation and application of GPAI behaviours depend strongly on the operational definitions, observation procedures, game context, and characteristics of the learners being assessed.

The methodological use of GPAI therefore requires particular attention to measurement quality. Earlier research raised concerns regarding the calculation and interpretation of game-performance indices, the distinction between game involvement and game performance, and observer reliability (Memmert & Harvey, 2008). More recent evidence has likewise emphasised the importance of clearly defined assessment criteria, observation units, learner roles, and contextual conditions when evaluating tactical learning and game performance (Barquero-Ruiz et al., 2020; Manso-Lorenzo et al., 2024). Consequently, evidence that supports the use of GPAI in one population or educational context cannot automatically be assumed to establish its validity in another context.

This issue is particularly relevant when an observation-based instrument is transferred across languages and educational cultures. Cross-cultural adaptation involves more than literal translation; it requires preservation of conceptual, semantic, and contextual equivalence between the original and adapted instrument (Cruchinho et

al., 2024; Zaragoza-Salcedo et al., 2023). For GPAI, this consideration is especially important because behavioural constructs such as support, cover, guard, and base are expressed through dynamic interactions among players and may be interpreted differently according to sport, game format, instructional practices, and learner characteristics (Barquero-Ruiz et al., 2020; Manso-Lorenzo et al., 2024). Therefore, an Indonesian adaptation requires empirical evaluation to determine whether the translated behavioural indicators are conceptually relevant, structurally coherent, and consistently applicable in the target PE context.

In Indonesian PE, team sports are an important setting for developing students' tactical understanding, decision making, and game-based competencies (Kusuma et al., 2024; Pratama et al., 2026). However, evidence concerning the systematic use and psychometric evaluation of standardised observation-based instruments for assessing students' tactical and game-performance behaviours remains limited in Indonesian school settings. Recent research has predominantly examined game-based pedagogical approaches and students' performance outcomes, whereas empirical evidence on the cross-cultural adaptation and psychometric evaluation of established game-performance observation instruments remains scarce (Manso-Lorenzo et al., 2024). In particular, it remains unclear whether the behavioural definitions and proposed measurement structure of the GPAI can be interpreted consistently and demonstrate satisfactory measurement properties among Indonesian secondary-school students. Thus, the research gap extends beyond the absence of an Indonesian-language version; empirical evidence is needed to establish whether an adapted GPAI is valid and reliable within the Indonesian educational and cultural context.

Addressing this gap is important for both measurement and pedagogical practice. A culturally adapted and psychometrically evaluated GPAI could provide PE teachers and researchers with a structured framework for observing students' tactical and game-performance behaviours during authentic team-sport activities. Such an instrument would also facilitate more systematic assessment of game performance and provide a basis for future research on game-centred PE in Indonesia. Therefore, this study aims to culturally adapt the GPAI for Indonesian secondary-school physical education and evaluate its measurement properties for assessing students' game-performance behaviours in team-sport contexts. Specifically, the study addressed three research questions:

RQ1: Does the Indonesian adaptation of the GPAI demonstrate adequate content validity and contextual relevance?

RQ2: Does the proposed measurement structure of the Indonesian GPAI demonstrate adequate structural validity?

RQ3: Does the Indonesian GPAI demonstrate adequate reliability, including internal consistency and inter-rater reliability?

METHODS

Research Design

This study employed a cross-sectional instrument-validation design to culturally adapt the Game Performance Assessment Instrument (GPAI) for Indonesian physical education and evaluate its measurement properties in secondary-school students. The study followed a sequential process comprising cross-cultural adaptation, content validation, field administration, and psychometric evaluation. The cross-cultural adaptation process involved forward translation, translation synthesis, back-translation, expert panel review, and pilot testing. Following the adaptation process, the Indonesian-adapted GPAI was administered during authentic team-sport activities to obtain field data for psychometric evaluation. The measurement evaluation examined content

validity, structural validity, convergent validity, discriminant validity, internal consistency, and inter-rater reliability. The overall procedure consisted of seven sequential stages: (i) forward translation, (ii) translation synthesis, (iii) back-translation, (iv) expert panel evaluation, (v) pilot testing, (vi) field data collection, and (vii) psychometric analysis. The stages and their respective purposes are presented in **Table 1**.

Table 1. Stages of the Indonesian GPAI Cross-Cultural Adaptation and Validation Process

Stage	Procedure	Purpose
1	Forward translation	Translate the original GPAI from English into Indonesian
2	Translation synthesis	Reconcile discrepancies between the forward translations and produce a preliminary Indonesian version
3	Back-translation	Examine semantic and conceptual equivalence with the original instrument
4	Expert panel evaluation	Assess relevance, clarity, cultural appropriateness, and content validity
5	Pilot testing	Evaluate comprehensibility, feasibility, and the clarity of the observation procedures
6	Field data collection	Assess students' game performance during team-sport activities
7	Psychometric analysis	Evaluate content validity, structural validity, convergent and discriminant validity, internal consistency, and inter-rater reliability

Participants

A total of 200 secondary-school students participated in the study, comprising 110 male students (55.0%) and 90 female students (45.0%). Of these, 105 students (52.5%) were enrolled in junior high school and 95 (47.5%) in senior high school. Students were observed during team-sport physical education activities involving football (n = 94), basketball (n = 57), and volleyball (n = 49). These sports were selected because they provide authentic team-game contexts in which tactical decision making, skill execution, positioning, support, and defensive behaviours can be observed during gameplay. The demographic and sport characteristics of the participants are presented in **Table 2**.

Participants were recruited using purposive sampling. Eligibility criteria included: (i) enrollment in a junior or senior high school physical education programme; (ii) participation in the selected team-sport activities during the data-collection period; (iii) regular participation in the observed physical education sessions; and (iv) provision of informed consent by the students and, where required, their parents or legal guardians. Students were excluded if they were unable to participate in the observed physical education activities or had incomplete observation data.

The sample size was considered adequate for the planned confirmatory factor analysis. The hypothesised measurement model comprised seven latent factors and 21 observed indicators. According to COSMIN guidance, a sample of at least seven times the number of items, with a minimum of 100 participants, can be considered adequate for factor analysis (Mokkink et al., 2018). For the present 21-indicator model, this corresponds to a benchmark of 147 participants. The final sample of 200 students exceeded this benchmark. However, because CFA sample-size requirements may also depend on model complexity, factor loadings, and estimation characteristics, the participant-to-indicator ratio was considered a supporting criterion rather than the sole basis for determining sample adequacy.

Table 2. Demographic and Sport Characteristics of the Participants

Variable	Category	n	%
Sex	Male	110	55.0
	Female	90	45.0
School Level	Junior High School	105	52.5
	Senior High School	95	47.5
Sport Observed	Football	94	47.0
	Basketball	57	28.5
	Volleyball	49	24.5
Total		**200**	**100**

Instrument

The primary instrument used in this study was the Game Performance Assessment Instrument (GPAI), originally developed by (Oslin et al., 1998). The GPAI is an observation-based instrument designed to assess students' game performance through systematic observation of behaviours occurring during authentic gameplay. Rather than focusing exclusively on isolated technical skill execution, the GPAI captures multiple dimensions of game performance, including tactical decision making, skill execution, positioning, and interactions with teammates and opponents within dynamic game situations. The original GPAI comprises seven game-performance components: base, adjust, decision making, skill execution, support, cover, and guard (Oslin et al., 1998). These components represent complementary dimensions of performance during gameplay, including maintaining appropriate positioning, adapting to changing game situations, making tactical decisions, executing selected skills, providing offensive support, providing defensive coverage, and defending against opponents.

For the present study, all seven original GPAI components were retained as the conceptual framework for the Indonesian adaptation. The adapted instrument comprised 21 observed behavioural indicators, with three indicators representing each of the seven components. The indicators were translated and contextually adapted for use in Indonesian physical education while preserving the conceptual meaning of the corresponding original GPAI components. The resulting Indonesian version was subsequently evaluated for content validity, structural validity, convergent and discriminant validity, internal consistency, and inter-rater reliability. The seven GPAI components, their corresponding observed indicators, and the operational definitions used in the present study are presented in **Table 3**.

Table 3. Game-Performance Components and Behavioural Indicators of the Indonesian-adapted GPAI

GPAI component	Observed indicator	Operational definition
Base	Base 1	Returns to an appropriate position after performing a game action
	Base 2	Maintains an appropriate position relative to the game situation
	Base 3	Positions appropriately to remain ready for the next play
Adjust	Adjust 1	Modifies position in response to the movement of the ball
	Adjust 2	Adapts position to the movement of teammates or opponents
	Adjust 3	Changes positioning appropriately as the game situation changes
Decision Making	Decision making 1	Selects an appropriate tactical action
	Decision making 2	Selects an appropriate option based on the game situation
	Decision making 3	Makes an effective tactical decision under game constraints
Skill Execution	Skill execution 1	Performs the selected skill effectively
	Skill execution 2	Executes the skill with appropriate control and accuracy
	Skill execution 3	Completes the skill action appropriately for the game situation

GPAI component	Observed indicator	Operational definition
Support	Support 1	Moves to an appropriate supporting position for a teammate
	Support 2	Provides an available option during offensive play
	Support 3	Adjusts supporting movement according to the attacking situation
Cover	Cover 1	Provides defensive support for a teammate
	Cover 2	Positions appropriately to cover a teammate
	Cover 3	Adjusts defensive coverage according to the movement of opponents and teammates
Guard	Guard 1	Marks or defends an appropriate opponent
	Guard 2	Maintains an appropriate defensive position relative to an opponent
	Guard 3	Adjusts defensive action according to the opponent's movement

Note: The behavioural indicators were translated and contextually adapted from the original GPAI framework for use in Indonesian physical education.

Translation and Cultural Adaptation

The translation and cultural adaptation process followed established guidelines for cross-cultural adaptation of measurement instruments (Beaton et al., 2000). The process aimed to preserve semantic, conceptual, and contextual equivalence between the original GPAI and its Indonesian version. The original instrument was independently translated into Indonesian by two bilingual experts, followed by synthesis and back-translation by an independent bilingual translator. An expert panel of five specialists in physical education pedagogy and sport assessment evaluated the relevance, essentiality, clarity, and contextual appropriateness of the 21 behavioural indicators representing the seven GPAI components. A pilot test involving 20 secondary-school students was subsequently conducted to assess the clarity and feasibility of the observation procedures. The final Indonesian version was established after incorporating the findings from expert review and pilot testing.

Data Collection Procedures

Data were collected during regular physical education lessons involving team-sport gameplay in football, basketball, and volleyball. Students participated in modified small-sided or representative game situations appropriate to each sport. Trained observers independently recorded students' observable game behaviours using the Indonesian GPAI observation sheet and the predefined operational definitions and scoring criteria.

Two independent raters evaluated the gameplay sessions to assess inter-rater reliability. Before data collection, the observers underwent structured training and pilot scoring focused on identifying, classifying, and recording the seven GPAI components. During the main data collection, raters independently recorded observations without consulting each other to minimise scoring bias and ensure independent ratings.

Ethical Considerations

The study was conducted in accordance with established ethical principles for research involving human participants. Permission to conduct the study was obtained from the participating schools. Informed consent was obtained from all participants and their parents or legal guardians before data collection. Participation was voluntary, and all data were treated confidentially and used solely for research purposes.

Data Analysis

Data were analysed using IBM SPSS and AMOS to evaluate the measurement properties of the Indonesian GPAI. The validation process included assessments of content validity, structural validity, convergent validity, construct reliability, discriminant validity, internal consistency, and inter-rater reliability. Content validity was assessed using the Content Validity Ratio (CVR) based on ratings from five experts. Structural validity was examined using Confirmatory Factor Analysis (CFA) to test the hypothesised seven-component GPAI structure. Convergent validity was evaluated using standardised factor loadings and Average Variance Extracted (AVE), while construct reliability was assessed using Composite Reliability (CR). Discriminant validity was evaluated using the Fornell–Larcker criterion. Internal consistency was assessed using Cronbach's alpha. Inter-rater reliability between the two independent observers was assessed using a two-way random-effects Intraclass Correlation Coefficient with absolute agreement for single measures [ICC(2.1)]. The statistical procedures and evaluation criteria are summarised in **Table 4**.

Table 4. Statistical Procedures Used for Instrument Validation

Measurement Property	Statistical Method	Evaluation Criteria
Content Validity	Content Validity Ratio (CVR)	Critical value based on a five-member expert panel
Structural Validity	Confirmatory Factor Analysis (CFA)	$\chi^2/df \leq 3.0$; CFI ≥ 0.90; TLI ≥ 0.90; RMSEA ≤ 0.08; SRMR ≤ 0.08
Convergent Validity	Standardised factor loadings; AVE	Factor loading ≥ 0.50; AVE ≥ 0.50
Construct Reliability	Composite Reliability (CR)	CR ≥ 0.70
Discriminant Validity	Fornell–Larcker criterion	√AVE greater than correlations with other constructs
Internal Consistency	Cronbach's alpha (α)	α ≥ 0.70
Inter-Rater Reliability	Intraclass Correlation Coefficient, ICC(2,1)	Two-way random-effects model; absolute agreement; single measures; ICC ≥ 0.75 interpreted as good reliability and ≥ 0.90 as excellent reliability

RESULTS

Content Validity

Content validity of the Indonesian GPAI was evaluated by five experts in physical education pedagogy and sport assessment. The experts assessed the essentiality and contextual appropriateness of the 21 behavioural indicators representing the seven GPAI components. As shown in **Table 5**, the CVR values ranged from 0.60 to 1.00. All indicators met the predefined content-validity criterion and were therefore retained in the Indonesian adaptation. These findings indicate that the 21 behavioural indicators were considered sufficiently relevant and appropriate for observing students' game-performance behaviours in Indonesian secondary-school physical education.

Table 5. Content Validity of the 21 Behavioural Indicators of the Indonesian GPAI

GPAI Component	Behavioural Indicator	CVR	Decision
Base	Base 1	1.00	Retained
	Base 2	1.00	Retained
	Base 3	0.60	Retained
Adjust	Adjust 1	1.00	Retained
	Adjust 2	1.00	Retained
	Adjust 3	0.60	Retained
Decision making	Decision making 1	1.00	Retained

GPAI Component	Behavioural Indicator	CVR	Decision
	Decision making 2	1.00	Retained
	Decision making 3	1.00	Retained
Skill execution	Skill execution 1	1.00	Retained
	Skill execution 2	1.00	Retained
	Skill execution 3	1.00	Retained
Support	Support 1	1.00	Retained
	Support 2	1.00	Retained
	Support 3	0.60	Retained
Cover	Cover 1	1.00	Retained
	Cover 2	1.00	Retained
	Cover 3	0.60	Retained
Guard	Guard 1	1.00	Retained
	Guard 2	1.00	Retained
	Guard 3	1.00	Retained

Note: CVR was calculated using the Lawshe formula based on ratings from five experts. Indicators were retained when their CVR met or exceeded the predefined critical value.

Structural Validity: Confirmatory Factor Analysis

Structural validity was examined using confirmatory factor analysis (CFA) in AMOS to determine whether the observed data supported the hypothesised seven-component measurement structure of the Indonesian GPAI. The model specified seven correlated latent factors, each represented by three observed behavioural indicators, resulting in 21 observed indicators. The factors were Base, Adjust, Decision Making, Skill Execution, Support, Cover, and Guard, with each factor represented by its corresponding three indicators (Base 1-3, Adjust 1-3, Decision Making 1-3, Skill Execution 1-3, Support 1-3, Cover 1-3, and Guard 1-3). The standardized CFA model showing the relationships between the seven latent factors and their observed indicators is presented in **Figure 1**.

Model fit was evaluated using multiple complementary goodness-of-fit indices, including the chi-square/degrees of freedom ratio (χ^2/df), Comparative Fit Index (CFI), Tucker–Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardised Root Mean Square Residual (SRMR). As shown in **Table 6**, the seven-factor model demonstrated an acceptable fit to the observed data, with χ^2/df = 2.18, CFI = 0.93, TLI = 0.92, RMSEA = 0.06, and SRMR = 0.05. Overall, these indices indicate that the hypothesised seven-component measurement structure provided an adequate representation of the observed covariance structure.

Table 6. Goodness-of-Fit indices of the Indonesian GPAI Measurement Model

Fit Index	Obtained Value	Evaluation Criterion	Interpretation
χ^2/df	2.18	≤ 3.00	Acceptable fit
CFI	0.93	≥ 0.90	Acceptable fit
TLI	0.92	≥ 0.90	Acceptable fit
RMSEA	0.06	≤ 0.08	Good fit
SRMR	0.05	≤ 0.08	Good fit

Note: χ^2/df, CFI, TLI, RMSEA, and SRMR were interpreted according to the predefined criteria described in the Methods.

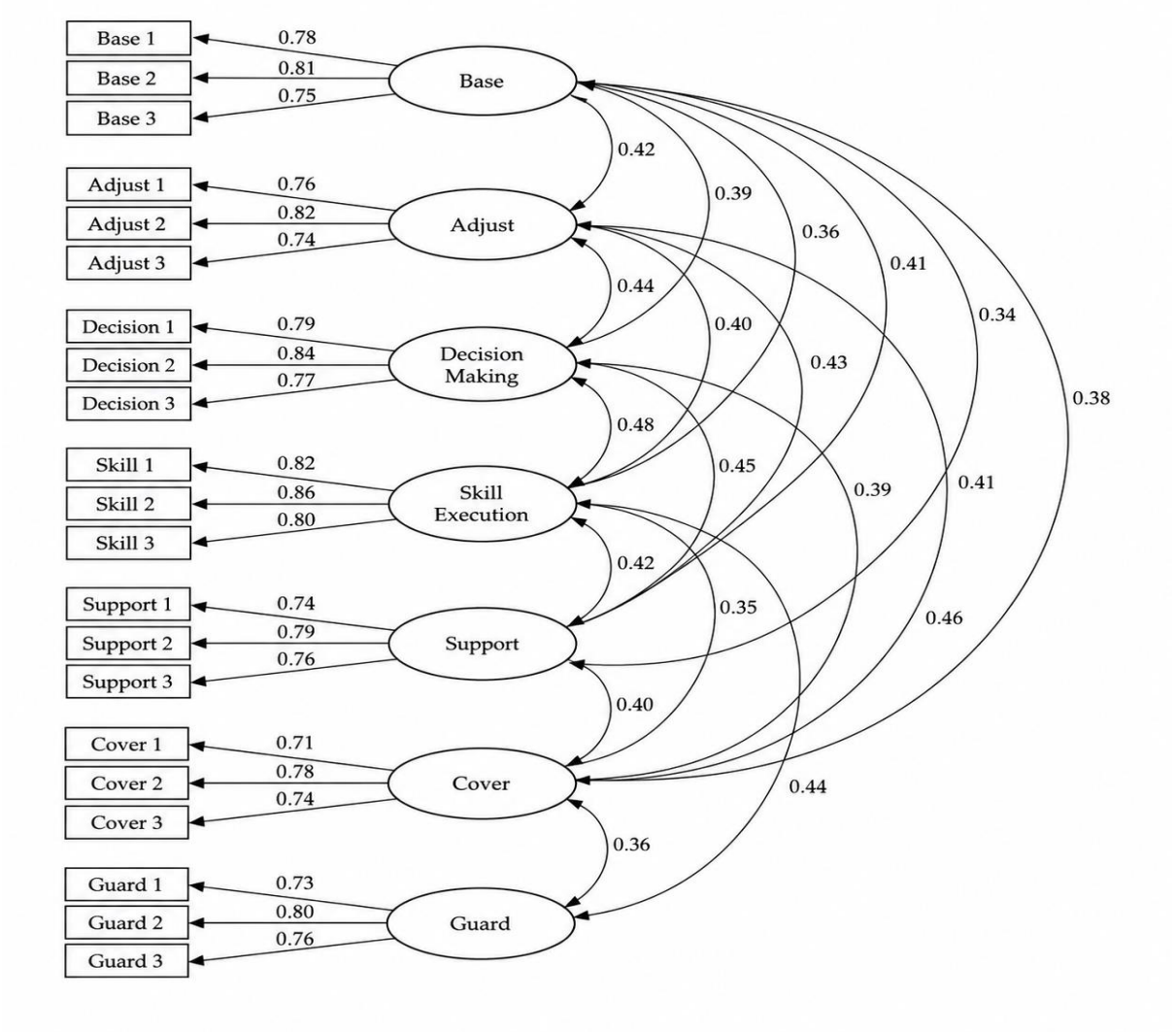

Figure 1. Confirmatory Factor Analysis Model of the Indonesian GPAI

The standardised factor loadings for all 21 observed indicators are presented in **Table 7**. The loadings ranged from 0.71 to 0.86, exceeding the predefined criterion of 0.50. All indicators therefore demonstrated substantial relationships with their respective latent factors, providing evidence in support of the hypothesised seven-component measurement structure of the Indonesian GPAI.

Table 7. Standardised Factor Loadings of the Indonesian GPAI Measurement Model

GPAI Component	Observed Indicator	Standardised Loading
Base	Base 1	0.78
	Base 2	0.81
	Base 3	0.75
Adjust	Adjust 1	0.76
	Adjust 2	0.82
	Adjust 3	0.74
Decision making	Decision making 1	0.79
	Decision making 2	0.84
	Decision making 3	0.77
Skill execution	Skill execution 1	0.82
	Skill execution 2	0.86
	Skill execution 3	0.80
Support	Support 1	0.74
	Support 2	0.79
	Support 3	0.76

GPAI Component	Observed Indicator	Standardised Loading
Cover	Cover 1	0.71
	Cover 2	0.78
	Cover 3	0.74
Guard	Guard 1	0.73
	Guard 2	0.80
	Guard 3	0.76

Note: Standardized factor loadings ≥ 0.50 were considered acceptable.

Convergent and Discriminant Validity

Convergent validity was assessed using standardised factor loadings, Average Variance Extracted (AVE), and Composite Reliability (CR) for the seven GPAI components. All 21 observed indicators showed standardised factor loadings above 0.50. As presented in **Table 8**, all seven components demonstrated AVE values exceeding 0.50 and CR values above 0.70, indicating adequate convergent validity and construct reliability.

Discriminant validity was assessed using the Fornell–Larcker criterion. The square root of AVE for each GPAI component was compared with its correlations with the other components. As shown in **Table 9**, the square root of AVE for each component exceeded all corresponding inter-component correlations. This finding indicates that the seven GPAI components demonstrated adequate discriminant validity and represented related but empirically distinguishable dimensions of game performance.

Table 8. Convergent Validity and Construct Reliability of the Indonesian GPAI

GPAI Component	AVE	CR	Convergent Validity
Base	0.61	0.83	Supported
Adjust	0.59	0.81	Supported
Decision making	0.63	0.84	Supported
Skill execution	0.67	0.86	Supported
Support	0.58	0.80	Supported
Cover	0.55	0.78	Supported
Guard	0.57	0.79	Supported

Note: AVE ≥ 0.50 and CR ≥ 0.70 were considered evidence of adequate convergent validity and construct reliability, respectively.

Table 9. Discriminant Validity of the Indonesian GPAI Using the Fornell–Larcker Criterion

GPAI Component	Base	Adjust	Decision making	Skill execution	Support	Cover	Guard
Base	**0.781**	0.42	0.39	0.36	0.41	0.34	0.38
Adjust	0.42	**0.768**	0.44	0.40	0.43	0.37	0.41
Decision making	0.39	0.44	**0.794**	0.48	0.45	0.39	0.46
Skill execution	0.36	0.40	0.48	**0.819**	0.42	0.35	0.43
Support	0.41	0.43	0.45	0.42	**0.762**	0.40	0.44
Cover	0.34	0.37	0.39	0.35	0.40	**0.742**	0.36
Guard	0.38	0.41	0.46	0.43	0.44	0.36	**0.755**

Note: Bold diagonal values represent the square root of AVE, while off-diagonal values represent correlations between latent constructs. Discriminant validity is supported when each diagonal value is greater than the corresponding inter-construct correlations.

Internal Consistency Reliability

Internal consistency reliability was assessed using Cronbach's alpha (α) for each of the seven GPAI components. As shown in **Table 10**, Cronbach's alpha coefficients ranged from 0.74 to 0.85, exceeding the predefined criterion of 0.70 for all seven components. These results indicate acceptable internal consistency among the behavioural indicators within each GPAI component.

Table 10. Internal Consistency Reliability of the Indonesian GPAI

GPAI Component	Cronbach's α	Interpretation
Base	0.81	Acceptable
Adjust	0.79	Acceptable
Decision making	0.82	Acceptable
Skill execution	0.85	Acceptable
Support	0.79	Acceptable
Cover	0.76	Acceptable
Guard	0.74	Acceptable

Note: Cronbach's alpha values ≥ 0.70 were considered indicative of acceptable internal consistency.

Inter-Rater Reliability

Inter-rater reliability was assessed using a two-way random-effects, absolute-agreement, single-measure Intraclass Correlation Coefficient (ICC[2,1]) to evaluate the agreement between the two independent observers. ICC coefficients were calculated for each of the seven GPAI components and ranged from 0.78 to 0.87, indicating good to excellent inter-rater reliability. These findings indicate that the two observers demonstrated a high level of agreement when applying the behavioural definitions and scoring procedures to students' game-performance behaviours. The ICC results for the seven GPAI components are presented in **Table 11**.

Table 11. Inter-rater Reliability of the Indonesian GPAI

GPAI Component	ICC(2,1)	Interpretation
Base	0.83	Good
Adjust	0.80	Good
Decision making	0.85	Good
Skill execution	0.87	Good
Support	0.80	Good
Cover	0.78	Good
Guard	0.81	Good

Note: ICC values were estimated using a two-way random-effects, absolute-agreement, single-measure model [ICC(2,1)]. ICC values of ≥ 0.75 were interpreted as indicating good inter-rater reliability, whereas values ≥0.90 were considered excellent.

Overall, the Indonesian GPAI demonstrated satisfactory evidence of content validity, structural validity, convergent validity, discriminant validity, internal consistency, and inter-rater reliability. The findings support the retention of the seven-component measurement structure in the Indonesian secondary-school physical education context. Collectively, the results provide evidence supporting the measurement quality of the adapted GPAI for assessing students' game performance in team-sport settings.

DISCUSSION

Content Validity of the Indonesian GPAI

The present study provides initial evidence that the Indonesian adaptation of the GPAI adequately represents the behavioural domains required to assess students' game performance in secondary-school physical education. All 21 behavioural indicators were retained following expert evaluation, with CVR values ranging from 0.60 to 1.00. Importantly, the evaluation was conducted at the indicator level rather than only at the level of the seven components, allowing the content representation of each observable behaviour to be examined directly. This approach is particularly relevant for observational instruments because validity depends not only on whether the broad constructs are appropriate but also on whether their operational definitions adequately capture the behaviours that occur during authentic game situations (Mokkink et al., 2025). Recent research on game-based observational assessment similarly emphasises

expert evaluation of behavioural categories and operational definitions before establishing observer reliability (Wilkie et al., 2023).

The retention of all 21 indicators also supports the cultural and contextual suitability of the Indonesian version. Rather than treating translation as a purely linguistic process, the adaptation considered whether the behavioural indicators remained meaningful within Indonesian physical education settings. This is consistent with recent validation studies in physical education that have combined expert review, cultural adaptation, pilot testing, and subsequent psychometric evaluation to establish measurement equivalence (Ceylan et al., 2026; Qin et al., 2024). The findings therefore suggest that the Indonesian GPAI retains the conceptual breadth of the original instrument while providing behavioural definitions that can be interpreted within the local educational context.

Structure Structural Validity of the Seven-Component Model

The CFA results provide evidence supporting the hypothesised seven-component structure of the Indonesian GPAI. The model demonstrated acceptable fit across the reported indices, while all 21 indicators showed standardised factor loadings above 0.50. More importantly, the findings indicate that the seven components can be empirically represented as related but distinguishable dimensions of game performance. This is theoretically appropriate because effective game performance is not a single behaviour; it involves positioning, adjustment, decision making, skill execution, support, cover, and defensive actions that interact during dynamic play (Zhang et al., 2025). Recent work in physical education assessment similarly supports multidimensional approaches for capturing the complexity of students' performance and behaviour rather than reducing performance to a single global score (Manso-Lorenzo et al., 2024; Wilkie et al., 2023).

The structural findings are also relevant in light of recent research on GPAI use in physical education. Manso-Lorenzo et al. (2024) reported that GPAI has increasingly been used to examine game performance in school-based interventions, although decision making and skill execution have received greater attention than other game components. The present findings extend this literature by providing psychometric support for a seven-component structure that explicitly preserves offensive, supportive, positional, and defensive behaviours. This multidimensional structure is particularly important because recent reviews have argued that game-based assessment should capture performance in representative game contexts, where technical execution is embedded within perceptual and tactical demands (Zhang et al., 2025).

Convergent and Discriminant Validity

Evidence for convergent validity was supported by the standardised factor loadings, AVE values, and CR values. The AVE estimates exceeded the conventional 0.50 criterion and the CR coefficients exceeded 0.70 for all seven components, indicating that the indicators within each component shared sufficient variance and consistently represented their intended latent construct. Similar patterns have been reported in recent physical education instrument-validation studies, where CFA, factor loadings, AVE, and CR were jointly used to establish the quality of multidimensional measurement models (Qin et al., 2024). Thus, the current findings provide evidence that the 21 behavioural indicators are not merely descriptive categories but demonstrate coherent relationships with their respective latent dimensions.

Discriminant validity was supported by the Fornell–Larcker criterion, with the square root of AVE for each component exceeding its correlations with the other components. This distinction is important because several GPAI dimensions are conceptually related

during actual game play. For example, decision making may occur simultaneously with adjustment and skill execution, while support and cover are inherently connected to teammates' and opponents' positions. Nevertheless, the observed correlations remained sufficiently below the corresponding square roots of AVE to indicate empirical distinction among the seven constructs. Recent physical education scale-development research similarly emphasises the need to demonstrate that theoretically related dimensions remain empirically distinguishable rather than assuming that a satisfactory CFA model alone establishes discriminant validity (Wu et al., 2025).

Reliability and Inter-Rater Consistency

The internal consistency results further support the reliability of the Indonesian GPAI. Cronbach's alpha coefficients ranged from 0.74 to 0.85 across the seven components, indicating that the behavioural indicators within each component demonstrated acceptable consistency. These findings are comparable with recent validation studies in physical education in which reliability coefficients above 0.70 were considered adequate for multidimensional educational measurement (Boukari et al., 2025; Jin & Xi, 2026). However, because the GPAI is an observational rather than purely self-report instrument, internal consistency should not be interpreted as sufficient evidence of measurement quality. The consistency with which observers apply behavioural definitions is equally important.

The ICC coefficients ranging from 0.78 to 0.87 provide additional evidence that the Indonesian GPAI can be applied consistently by independent observers. This finding is particularly important because observational assessment is vulnerable to differences in observer interpretation, especially when behaviours occur rapidly and simultaneously during game play. Wilkie et al. (2023) similarly emphasised observer training and inter-observer reliability as essential stages in the validation of games-based observational assessment. A recent systematic review of game-scenario-based ball-skill assessment tools also found that inter-observer reliability was more frequently established than intra-observer reliability, reinforcing the importance of demonstrating observer agreement when such instruments are used in authentic performance environments (Zhang et al., 2025).

Implications for Physical Education Assessment

The findings have practical implications for assessment in Indonesian physical education. The Indonesian GPAI provides teachers with a structured approach for evaluating game performance beyond isolated technical execution. This is particularly consistent with the contemporary movement toward tactical and game-centered pedagogies, in which students are expected to understand not only how to execute a skill but also when, where, and why a particular action should be performed. A systematic review of tactical game-centered approaches found that such approaches can improve tactical skill, decision making, skill execution, procedural knowledge, and physical activity outcomes compared with more traditional approaches, although methodological quality across studies remains variable (Breed et al., 2025; Manninen et al., 2025). Similarly, recent reviews of tactical games in physical education indicate increasing use of game-performance assessment instruments such as GPAI, particularly in invasion and team-sport contexts (Wang et al., 2024).

The practical value of the instrument is further strengthened by evidence that contemporary assessment increasingly needs to capture performance within representative game environments. Zhang et al. (2025) noted that game-scenario assessment can integrate technical performance with perceptual-cognitive processes, including decision making and tactical awareness. More specifically, Steger et al. (2026)

demonstrated that video-supported tactical analysis could improve students' positioning and correct ball-demand behaviours, illustrating how assessment of tactical behaviours can be connected directly to learning processes. Consequently, the Indonesian GPAI should not be viewed merely as a scoring instrument; it can also provide teachers with structured information for formative assessment, feedback, and instructional decision making. Nevertheless, future studies should examine its measurement invariance across different sports, school levels, genders, and regions of Indonesia, as well as its sensitivity to changes following specific game-centered pedagogical interventions.

Limitations and Future Research

Despite the encouraging psychometric findings, several limitations should be considered when interpreting the results. First, the validation was conducted among Indonesian secondary-school students; therefore, the seven-component structure and reliability estimates should not be assumed to generalise directly to elementary students, university students, or athletes. Second, the present study focused primarily on content validity, structural validity, convergent and discriminant validity, internal consistency, and inter-rater reliability. Although these findings provide substantial initial evidence for the measurement quality of the Indonesian GPAI, further evidence is needed to establish whether the instrument operates equivalently across relevant subgroups. Future studies should therefore examine measurement invariance across sex, age, school type, geographic region, and different levels of playing experience. Such analyses would determine whether comparisons of GPAI scores across groups are psychometrically justified rather than simply assuming equivalence.

Future research should also investigate the generalisability and sensitivity of the Indonesian GPAI across different team-sport contexts and instructional approaches. In particular, longitudinal and intervention-based studies could examine whether the instrument is sensitive to changes in students' tactical performance following game-centered or tactical-learning interventions. Research should also compare its functioning across invasion games, net/wall games, and other game categories, as the behavioural demands associated with positioning, decision making, support, cover, and guarding may differ across sports. Recent research on game-based assessment has emphasised the importance of representative game contexts and the need to examine assessment tools across different game situations and populations (Wang et al., 2024; Zhang et al., 2025). Accordingly, future studies should move beyond establishing whether the Indonesian GPAI is valid and reliable toward examining how sensitive, generalisable, and instructionally useful the instrument is in authentic physical education practice.

CONCLUSION

This study provides evidence supporting the psychometric quality of the Indonesian adaptation of the Game Performance Assessment Instrument (GPAI) for assessing students' game-performance behaviours in secondary-school physical education. Content validity was supported across the 21 behavioural indicators, while CFA supported the hypothesised seven-component structure with acceptable model fit and standardised factor loadings. The findings also demonstrated adequate convergent and discriminant validity, internal consistency, and inter-rater reliability.

The Indonesian GPAI therefore provides a structured observational framework for assessing students' game-performance behaviours, including decision making, skill execution, and tactical behaviours during team-sport activities. Although the findings support its use in the Indonesian secondary-school context, further research should

examine the instrument across different age groups, sports, educational settings, and larger samples, including measurement invariance and technology-assisted assessment.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to the five expert reviewers, participating schools and students, and the independent observers who contributed to the adaptation and validation of the Indonesian version of the Game Performance Assessment Instrument (GPAI). Their expertise, participation, and cooperation were essential to the completion of this study.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest that could have influenced the design, conduct, analysis, or reporting of this study.

AI DISCLOSURE STATEMENT

The authors used artificial intelligence (AI)-based tools solely to support language editing, grammar checking, and manuscript refinement. AI tools were not used to generate, analyse, or interpret the research data. All methodological decisions, statistical analyses, interpretations, and final content were reviewed and approved by the authors, who take full responsibility for the accuracy and integrity of the manuscript.

REFERENCES

Barba-Martín, R. A., Bores-García, D., Hortigüela-Alcalá, D., & González-Calvo, G. (2020). The Application of the Teaching Games for Understanding in Physical Education. Systematic Review of the Last Six Years. *International Journal of Environmental Research and Public Health*, *17*(9), 1–16. https://doi.org/10.3390/ijerph17093330

Barquero-Ruiz, C., Arias-Estero, J. L., & Kirk, D. (2020). Assessment for Tactical Learning in Games: A Systematic Review. *European Physical Education Review*, *26*(4), 827–847. https://doi.org/10.1177/1356336X19889649

Beaton, D., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Boukari, S., Guelmami, N., Chalghaf, N., Ceylan, H. İ., Stefanica, V., Mekni, R., Dhahbi, W., & Dergaa, I. (2025). Cross-Cultural Adaptation and Psychometric Evaluation of the Engaged Teachers Scale among Arabic-Speaking Physical Education Teachers. *Frontiers in Education*, *10*, 1–8. https://doi.org/10.3389/feduc.2025.1634393

Breed, R., Lindsay, R., Kittel, A., & Spittle, M. (2025). Content and Quality of Comparative Tactical Game-Centered Approaches in Physical Education: A Systematic Review. *Review of Educational Research*, *95*(2), 293–336. https://doi.org/10.3102/00346543241227236

Ceylan, M., Ceylan, K., & Yoncalık, O. (2026). The Adaptation and Validation of the Achievement Emotions Questionnaire for Physical Education-Elementary School. *International Journal of Assessment Tools in Education, 13*(3), 674–688. https://doi.org/10.21449/ijate.1730488

Cruchinho, P., López-franco, M. D., Capelas, M. L., Almeida, S., Bennett, P. M., Miranda, M., Teixeira, G., Nunes, E., Lucas, P., Gaspar, F., & Behalf, O. (2024). Translation, Cross-Cultural Adaptation, and Validation of Measurement Instruments: A Practical Guideline for Novice Researchers. *Journal of Multidisciplinary Healthcare*, *17*, 2701–2728. https://doi.org/10.2147/JMDH.S419714

González-Valero, G., Ubago-Jiménez, J. L., Melguizo-Ibáñez, E., & Fernández-García, R. (2024). Application of the Teaching Games for Understanding Model to Improve Decision-Making in Sport Learning: A Systematic Review and Meta-Analysis. *BMC Psychology*, *12*(1), 781. https://doi.org/10.1186/s40359-024-02307-2

Jin, Z., & Xi, C. (2026). Physical Education Course Satisfaction Scale for Chinese College Students: Development and Psychometric Properties. *Frontiers in Public Health*, *13*, 1–11. https://doi.org/10.3389/fpubh.2025.1717552

Kusuma, D. W. C. W., Satrianingsih, B., Kusuma, L. S. W., & Gustian, U. (2024). Implementation of Teaching Games for Understanding to Stimulate Game Understanding of Elementary School Students. *Jurnal Pendidikan Jasmani dan Olahraga*, *9*(59), 255–261. https://doi.org/10.17509/jpjo.v9i2.67267

Manninen, Mika, Magrum, Eric, Campbell, Sara, & Belton, Sarahjane. (2025). The Effect of Game-Based Approaches on Decision-Making, Knowledge, and Motor Skill: A Systematic Review and a Multilevel Meta-Analysis. *European Physical Education Review*, *31*(1), 18–32. https://doi.org/10.1177/1356336X241245305

Manso-Lorenzo, V., Evangelio, C., Ruiz-Tendero, G., & González-Víllora, S. (2024). Game Performance Assessment Instrument in Physical Education: A Systematic Review from 2015 to 2024. *Physical Education and Sport Pedagogy*, *23*, 1–16. https://doi.org/10.1080/17408989.2024.2432313

Memmert, D., & Harvey, S. (2008). The Game Performance Assessment Instrument (GPAI): Some Concerns and Solutions for Further Development. *Journal of Teaching in Physical Education*, *27*(2), 220–240. https://doi.org/10.1123/jtpe.27.2.220

Mokkink, L. B., de Vet, H. C. W., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN Risk of Bias Checklist for Systematic Reviews of Patient-Reported Outcome Measures. *Quality of Life Research*, *27*(5), 1171–1179. https://doi.org/10.1007/s11136-017-1765-4

Mokkink, L. B., Herbelet, S., Tuinman, P. R., & Terwee, C. B. (2025). Content Validity: Judging the Relevance, Comprehensiveness, and Comprehensibility of an Outcome Measurement Instrument - A COSMIN Perspective. *Journal of Clinical Epidemiology, 185*, 111879. https://doi.org/10.1016/j.jclinepi.2025.111879

Oslin, J. L., Griffin, L. L., & Mitchell, S. A. (1998). The Game Performance Assessment Instrument (GPAI): Development and Preliminary Validation. *Journal of Teaching in Physical Education, 17*(2), 231-243. https://doi.org/10.1123/jtpe.17.2.231

Pan, Y. H., Huang, C. H., & Hsu, W. T. (2023). A Comparison of the Learning Effects between TGfu-SE and TGfu on Learning Motivation, Sport Enjoyment, Responsibility, and Game Performance in Physical Education. *Frontiers In Psychology, 14*, 1165064. https://doi.org/10.3389/fpsyg.2023.1165064

Pratama, S. A., Susila, L., & Clintock, E. M. (2026). Game-Based Learning Approach to Enhance Tactical Understanding and Motivation in School Football Physical Education. *INSPIREE: Indonesian Sport Innovation Review*, *7*(1), 23–32. https://doi.org/10.53905/inspiree.v7i01.163

Qin, L., King Yan Ho, W., Xie, Y., Wang, J., Cheng, L., Liu, Z., Pan, L., & Khoo, S. (2024). The Development and Validation of a Scale to Assess Perception of Physical Education Among University Students in China. *Sage Open*, *14*(1), 21582440241228910. https://doi.org/10.1177/21582440241228910

Silva, A., Ferraz, R., Branquinho, L., Dias, T., Teixeira, J. E., & Marinho, D. A. (2023). Effects of Applying a Multivariate Training Program on Physical Fitness and Tactical Performance in a Team Sport Taught during Physical Education Classes. *Frontiers in Sports and Active Living, 5*, 1291342. https://doi.org/10.3389/fspor.2023.1291342

Steger, M., Roth, A.-C., Beege, M., & Reinhold, F. (2026). Game Tactical Learning for Primary School Students With a Digital Video Analysis Tool. *Journal of Teaching in Physical Education*, *45*(3), 608–619. https://doi.org/10.1123/jtpe.2024-0421

Wang, J., Soon, C. C., Samsudin, S., Wang, C., Gao, Z., & Xie, Q. (2024). Tactical Games Model in Physical Education: A Systematic Review. *PLoS ONE*, *19*(11), 1–29. https://doi.org/10.1371/journal.pone.0311321

Wilkie, B., Jordan, A., Foulkes, J., Woods, C. T., Davids, K., & Rudd, J. (2023). Examining the Validity, Reliability and Feasibility of Capturing Children's Physical Literacy Through Games-Based Assessment in Physical Education. *Frontiers in Sports and Active Living, 5*, 1188364. https://doi.org/10.3389/fspor.2023.1188364

Wu, Y., Wang, C., Liu, H., Zhang, C., Jin, Y., Hu, T., & Min, H. (2025). Development and Validation of a Core Literacy Scale for Physical Education Among Chinese University Students. *Behavioral Sciences*, *15*(2), 1–11. https://doi.org/10.3390/ bs15121655

Zaragoza-Salcedo, A., Oroviogoicoechea, C., Saracíbar-Razquin, M. I., & Osácar, E. (2023). The Significance of Exploring Conceptual Equivalence within the Process of the Cross-Cultural Adaptation of Tools: The Case of the Patient's Perception of Feeling Known by Their Nurses Scale. *Journal of Nursing Scholarship: An Official Publication of Sigma Theta Tau International Honor Society of Nursing, 55*(6), 1268–1279. https://doi.org/10.1111/jnu.12910

Zhang, D., Shi, P., Jin, T., & Zhang, K. (2025). Tools for Assessing Ball Skills Based on Game Scenarios: A Systematic Review and Related Insights. *BMC Sports Science, Medicine and Rehabilitation*, *17*(1), 36. https://doi.org/10.1186/s13102-025-01077-7