Article

Perceptual and Motor Skills
2024, Vol. 131(2) 589–611

DOI: 10.1177/00315125231225579
journals.sagepub.com/home/pms

Psychometric Validation of the Tactical Assessment Instrument in Football for Use in Physical Education and with Youth Sport Teams

Carmen Barquero-Ruiz[1], David Kirk[2,3], Lourdes Meroño[4], and José L. Arias-Estero[5]

Abstract

Recently, the Tactical Assessment Instrument in Football (TAIS) was developed to address football's tactical requirements in terms of game phases, learners' roles, and organizational levels of play. In this study, we sought to evaluate the psychometric properties of this existing instrument in the Physical Education (PE) context and for youth sport teams (ages 8–12 years). Our research design involved: (a) participants playing 37 football games, (b) assessing participants via observation with the TAIS and (c) conducting psychometric analysis on the obtained assessment data. For the psychometric analysis, we carried out our work in four phases: (a) a statistical analysis of the evaluation criteria, (b) analysis of the instrument's structural dimensions, (c) internal item reliability analysis, and (d) provision of evidence for external validity. Participants were 592 children (156 girls, 436 boys; ages 8–12 years) from 74 PE classes or sport teams (from three institutional contexts - schools, community-based sports, and sports clubs) in Spain. We selected 12 tactical criteria and grouped them into four

[1]Department of Physical Education and Sport Sciences, University of Limerick, Limerick, Ireland
[2]School of Education, University of Strathclyde, Glasgow, UK
[3]School of Education, University of Queensland, St Lucia, QLD, Australia
[4]Facultad de Deporte, Universidad Católica de Murcia UCAM, Murcia, Spain
[5]Departamento de Expresión Plástica, Musical y Dinámica, Facultad de Educación, Universidad de Murcia, Murcia, Spain

Corresponding Author:
José L. Arias-Estero, Departamento de Expresión Plástica, Musical y Dinámica, Facultad de Educación, Universidad de Murcia, Campus Universitario de Espinardo, Murcia 30100, Spain.
Email: jlarias@um.es

theoretical dimensions (appropriate attack, inappropriate attack, appropriate defense, inappropriate defense) to confirm the instrument's structural dimensions. The Cronbach alpha and Omega McDonald coefficients were greater than .70. The TAIS discriminated between participants from the three institutional contexts, offering evidence of the instrument's external validity. Thus, the TAIS now has sufficient psychometric support for assessing learning of football tactics in PE and youth sport teams. This instrument facilitates an ecological assessment of youth players' understanding of football tactics through four theoretical tactical dimensions and 12 criteria involving information about four learners' roles (both attack and defense, on- and off-the-ball) and three organizational tactical levels of play (whole team, small groups of learners, and individual learners).

Keywords
sport pedagogy, game-based approaches, ecological assessment, youth football, game assessment instrument

Introduction

Teaching the tactical aspects of games has been emphasized in Physical Education (PE) after research showed positive results in the development of learners' cognitive abilities and their understanding of the games without neglecting technical execution (Barba-Martín et al., 2020; Morales-Belando et al., 2022). According to the PE curriculum standards in Spain, the development of children's knowledge of tactics in game situations is one of the main challenges of the Primary School phase of PE (Royal Decree 157/2022). In this national context, football (soccer) is one of the most socially impactful games in PE and in youth sport teams (8–12 years old), and tactics are crucial for game play (Fédération Internationale de Football Association [FIFA], 2018; Royal Decree 157/2022). However, none of the five instruments currently available for assessing tactics in football are supported by published psychometric research. This means there is insufficient knowledge of the suitability of these tools for assessing tactics, even though they were designed for this purpose (Barquero-Ruiz et al., 2022). These instruments lack any analysis of their item content (comprehension validity), structural dimensions (construct validity), internal item consistency (reliability), or external correlates (concurrent validity) that would allow others to use the instruments confidently (Barquero-Ruiz et al., 2020; Carretero-Dios & Pérez, 2007; Godbout & Gréhaigne, 2022).

These five most popular tools for tactical assessment in football in the PE and youth sports contexts are: (a) the Team Sports Assessment Procedure (TSAP; Gréhaigne et al., 1997); (b) the Game Performance Assessment Instrument (GPAI; Oslin et al., 1998); (c) the System of Tactical Assessment in Football (FUTSAT; Costa et al., 2011); (d) the Game Performance Evaluation Tool (GPET; García-López et al., 2013); and (e) the

Tactical Assessment Instrument in Football (TAIS, Barquero-Ruiz et al., 2022). While each of the first four instruments has been widely used in the context of PE and youth sport, recent investigators have named four main limitations, precipitating the creation of the fifth instrument, TAIS (Barquero-Ruiz et al., 2020; Memert & Harvey, 2008).

There are four main problems with the existing instruments that led us to develop TAIS. First, neither the TSAP nor the GPET consider all the learners' roles (e.g., attack and defense roles, and roles on- and off-the-ball), and this omission limits learning because decisions made without possession of the ball also significantly influence decisions made with ball possession (MacPhail et al., 2008). Second, none of the first four instruments listed above consider the three organizational levels of tactical play, and this limits assessment of what learners must do on the field (Godbout & Gréhaigne, 2022). Since tactical outcomes are not simply the sum of individual actions, awareness of tactical organization in team play is critically important to the assessment of interrelated game behaviors (Godbout & Gréhaigne, 2020; Kirk, 2017). Third, the TSAP, GPAI and FUTSAT use indices or ratios that make it difficult to identify the initial player behavior that triggers a positive or negative evaluation and only provides general information about learning (Memmert & Harvey, 2008). For example, if a player has a learning outcome score of 3.34 on the GPAI Performance Index, it is difficult to identify whether this result was a consequence of Decision Making, Skill Execution or other GPAI components; it is very difficult to know what specific learner behaviors led to this score. Finally, each of these first four instruments were validated in only one of the three institutional contexts for football play (i.e., school, community-based sports, or sports clubs), and this constrains the instruments' use in the other very different contexts (Harvey & Jarrett, 2014). For example, students in PE are more likely to have experienced training in tactical perspectives than sport teams (8–12 years old), and, on youth sport teams the duration of each training session and the number of sessions per week are higher than in PE (Harvey & Jarrett, 2014; Kinnerk et al., 2018; O'Connor et al., 2018; Stonebridge & Cushion, 2018).

The fifth assessment tool listed earlier, the TAIS (see Table 1) was created to address the limitations of the other instruments by enabling the assessment of team tactics through: (a) all the learners' roles; (b) three organizational tactical levels of play; (c) eliminating masking of data through indexes; (d) including tactic and technique as components of game behaviors; (e) analyzing both attack and defense game phases; and (f) addressing tactical play in all three institutional contexts for youth. As such, the TAIS was designed sequentially through two stages. In the first stage, the tactical criteria were identified and delimitated through a review of the literature, game observation, and judgments made by an expert panel. In the second stage, item content and comprehension validities were obtained by consulting experts systematically observing games, construct validity was determined by analyzing tactics of players from different skill levels, criterion validity was determined by comparing the GPAI and TAIS, and reliability was determined through inter-rater agreement calculations. In practical terms, the TAIS consists of 25 criteria (13 attacking, 12 defending) on which player behaviors are judged as appropriate or inappropriate. These criteria are grouped

Table 1. Criteria of the Current Tactical Assessment Instrument in Football by Organizational Tactical Levels of Play (Barquero-Ruiz et al., 2022).

Organizational match level (whole team)
- Appropriate positional attack: When it allows maintaining the possession of the ball or progressing in the field without losing the ball
- Inappropriate positional attack: When it does not allow maintaining the possession of the ball or progressing in the field with the ball
- Appropriate counter-attack: When it allows progressing in the field without losing the ball or players shot to goal from a forward field area
- Inappropriate counter-attack: When it does not allow progressing in the field with the ball or players do not shot on goal from a forward area
- Appropriate amplitude: When it creates safe pass options or zones without defenders
- Inappropriate amplitude: When it does not create safe pass options or zones without defenders
- Appropriate depth: When it allows vertically movements in the field to a forward field area
-Inappropriate depth: When it does not allow vertically movements in the field to a forward field area
- Appropriate defense in zone: when players are defending a specific area and they avoid the opponents' progression with the ball or shoot
- Inappropriate defense in zone: when players are defending a specific area and they do not avoid the opponents' progression with the ball or shoot
- Appropriate individual defense: When each player marks an opponent and they avoid the opponents' progression with the ball or shoot
- Inappropriate individual defense: When each player marks an opponent and they do not avoid the opponents' progression with the ball or shoot

Partial forefront organizational level (small groups of players)
- Appropriate pass: when the ball achieves a teammate during the pass
- Inappropriate pass: when the possession of the ball is lost during the pass
- Support: When off-ball attackers make a movement generating line of pass free of defenders
- Appropriate defensive coverage: When defenders avoid the opponent progression with the ball or shoot
- Inappropriate defensive coverage: When defenders do not avoid the opponent progression with the ball or shoot

Primary organizational level (individual player)
- Appropriate dribbling: When the on-ball attacker progress in the field, keeping the control of the ball by at least three touches
- Inappropriate dribbling: When the on-ball attacker does not progress in the field, keeping the control of the ball by at least three touches
- Appropriate shoot: When it finishes between the goalposts or it results in a corner kick
- Inappropriate shoot: When it does not finish between the goalposts or it does not result in a corner kick
- Appropriate interception: When the defender deflects the path of the ball and regains the possession of the ball

(continued)

Table 1. (continued)

- Inappropriate interception: When the defender does not deflect the path of the ball and does not regain the possession of the ball
- Appropriate tackle o charging: When the defender steals the ball or it is displaced, shaking shoulder to shoulder
- Inappropriate tackle o charging: When the defender does not steal the ball or it is not displaced, shaking shoulder to shoulder

into the three organizational tactical levels of play (Barquero-Ruiz et al., 2022, Table 1). To date, the TAIS has been designed to assess tactics from an aligned and ecological view (see discussion below), as it integrates learners' roles and interactions between the three tactical levels of play, game phases, and institutional contexts. However, it is still in need of further psychometric support, regarding its validity and reliability.

Considering tactical assessment in a PE context, alignment and ecology are both key criteria for enhancing the teaching-learning process (Morley et al., 2019). By alignment we mean ensuring that knowledge, teaching, and learning interact with and support desirable learning outcomes (Biggs, 2014). Thus, a good evaluation of tactical play must align the behaviors and knowledge assessed with the teaching-learning process that is linked to desired outcomes (Care et al., 2018). For example, if tactical decision making depends on the interaction among teammates and among opponents, tactical assessment instruments must be able to gain information about the quality of these interactions. On the other hand, regarding classroom ecology (Doyle, 2006), learners' behaviors are derived from the ecosystem formed by teachers and learners in their interactions within the learning context. The central tenet of an ecological learning approach resides in the influence of context on the learners' behaviors (Vors & Kirk, 2016). Within this framework, an ecological assessment must evaluate learners while considering the contextual demands of the task. In other words, a tactical sport assessment instrument should include information from multiple entry points and pathways, across the various elements in a game setting, considering all game phases, all learners' roles, and all interactions between the three organization levels of play, in real-game situations, and in the institutional context (Godbout & Gréhaigne, 2022; Kirk, 2017). More specifically, an assessment instrument for football tactics must include information regarding attack and defense game phases, players behaviors on- and off-the-ball, and interactions among the whole team, among small groups of players, and at individual player level.

Even though the current version of TAIS attempts to assess players' tactical knowledge and behavior from an aligned and ecological perspective, its further psychometric validation is needed (Godbout & Gréhaigne, 2022; Kirk, 2017). In the PE context, tactics can be defined as the player's capacity to make decisions regarding *which* behavior is appropriate, and *when*, *where,* and *how* it should be performed in games (Kirk, 2017). These decisions emerge from the game demands as determined by

the rules that constrain game phases, player roles, and three organizational levels of play (Godbout & Gréhaigne, 2020; Parlebas, 2001).

Concerning game phases, during an invasion game like football, there are two possible scenarios of interaction, depending on whether the team is in possession of the ball or not (attack and defense phases, Ribeiro et al., 2019). When the team has the ball, players can attack on-the-ball or off-the-ball. When the team does not have the ball, participants can defend on-the-ball or off-the-ball (Laakso et al., 2022). In both game phases, players' tactical behaviors result from the interactions at three tactical levels of play, among the whole team, among small groups of players, and at the level of an individual player (Godbout & Gréhaigne, 2022). Considering the three organizational tactical levels and the four possible player roles (two on attack and two on defense), players' appropriate behaviors become appropriate or inappropriate according to the game scenario (e.g., passing or dribbling according to the defenders' positions). Hence, tactical assessment instruments must provide information about game occurrences, and phases that characterize the four different learner roles and the three different organizational tactical levels of play (Barquero-Ruiz et al., 2022; Godbout & Gréhaigne, 2022; Kirk, 2017; Laakso et al., 2022; Ribeiro et al., 2019). Although the current version of TAIS integrates the game phases, player roles, and three tactical levels of play, it was not structured in accordance with the learning outcomes by game phase (e.g., appropriate attacking behaviors, inappropriate attacking behaviors, appropriate defensive behaviors, and inappropriate defensive behaviors), as theoretical dimensions of the construct to be assessed, because the initial validation study did not assess all the instrument's psychometric properties.

The psychometric properties of assessment tools allow for an exploration and confirmation of the instrument's underlying structural dimensions supporting an assessment construct (Mettert et al., 2020). In this case, the underlying construct, game tactics, is an abstract concept that is not directly measurable. According to Carretero-Dios and Pérez (2007), the analysis of this psychometric property is crucial to ensure that the instrument assesses exactly what it is intended to assess. The goal is to estimate the degree to which the instrument's criteria and dimensions structure the construct to be measured, and on which the interpretations of the measurements will be based (Mettert et al., 2020). The degree of an instrument's psychometric support should be considered by test users, such as teachers or researchers, when choosing an instrument, since psychometric support is scientific assurance that the tool they are using can accurately (validly) and consistently (reliably) measure the intended construct (Irwing et al., 2018). On the other hand, instruments without good psychometric properties constrain rigorous application and use (Hair et al., 2018).

Consistent with Carretero-Dios and Pérez's (2007) guidelines, the development and validation of an instrument should follow six phases. The two first phases correspond to conceptually and qualitatively defining the construct to assess (tactical behaviors in football in the present study). In the case of the TAIS, these two phases were previously completed (content validity, Barquero-Ruiz et al., 2022). The remaining phases are related to the instrument's psychometric validation. In phase three, the aim of a

statistical analysis of the evaluation criteria is to examine the metric properties of the test item criteria (comprehension validity). In phase four, a study of the instrument's structural dimensions seeks to explore and confirm that the theoretical assumptions or apparent statistical coherence shown by clusters of the criteria (test items) are in line with how they are grouped (construct validity). Thus, in the present study, we explored the theoretical dimensions of appropriate attack, inappropriate attack, appropriate defense, and inappropriate defense. In phase five, an analysis of the instrument's reliability aims is to determine the internal consistency of the criteria (items) clustered in each dimension. Finally, phase six provides evidence for external validity by showing whether the instrument discriminates based on some external factor as it should if the construct has been accurately measured (concurrent validity). In this case, we predicted that there would be differences in tactical behaviors between participants from the three different institutional contexts (school, community-based sports, and sports clubs) because of the higher current interest in teaching games from a tactical perspective in PE at Primary School in comparison to youth football (Kinnerk et al., 2018; Morales-Belando et al., 2022; Royal Decree 157/2022). In addition, we predicted that there would not be differences between participants from community-based sports and sports clubs because coaches traditionally teach football in these contexts from a technical perspective (O'Connor et al., 2018; Stonebridge & Cushion, 2018) and more practice time does not guarantee tactical learning if the teaching-learning approach is not appropriate (Arias-Estero et al., 2020). Our aim in the present study was to provide further psychometric support for the TAIS for football tactical assessment in the PE context and for youth sport teams.

Method

Research Design

Our research design involved using the existing TAIS to observe participants playing in 37 football games and subjecting the data we gathered to psychometric analysis. For our psychometric analysis we followed the phases proposed by Carretero-Dios and Pérez (2005, 2007): (a) statistical analysis of the evaluative criteria (comprehension validity), (b) analysis of the instrument's dimensionality (construct validity), (c) analysis of the instrument's reliability, and (d) provision of evidence for external validity (concurrent validity). For the statistical analysis of the evaluation criteria, we obtained data from eight games (2 from schools, 3 from community-based sports, and 3 from sports clubs). For analysis of the instrument's dimensionality, we obtained data from 48 games (15 from schools, 17 from community-based sports, and 16 from sports clubs). For an exploratory factorial analysis, we obtained data from 16 games (5 from schools, 6 from community-based sports, and 5 from sports clubs) and for the confirmatory factorial analysis, we obtained data from 32 games (10 from schools, 11 from community-based sports, and 11 from sports clubs). For analysis of the instrument's reliability and provision of external validity evidence, we obtained data from 18 games

(6 from schools, 6 from community-based sports, and 6 from sports clubs). Criteria assessed are shown in Table 1.

Participants

The first author contacted participants using a standardized script and an email message which were sent to PE teachers and coaches in schools, community-based sports, and sports clubs football teams with children from the target age group. These email recipients were informed about the characteristics we sought in the participants and asked for their collaboration. Inclusion criteria for children to be recruited were: (a) learning to play football in schools, community-based sports, or sports clubs and (b) aged between 8–12 years old.

Approximately 60% of the teachers and coaches contacted responded within 30 days (see Figure 1). The first author randomly selected the children from approximately 50% of the teachers and coaches who responded, based on the number of participants needed for each phase of the psychometric study (Carretero-Dios & Pérez, 2007). Following recommendations by Carretero-Dios and Pérez (2007), each phase of study involved different participants and different numbers of participants. In total, participants were 592 children (156 girls, 436 boys, *Mage* = 10.76, *SD* = 1.45 years); they were selected from 74 PE classes or teams (23 from schools, 26 from community-based sports, and 25 from sports clubs; see Figure 1). They came from lower and upper middle-class backgrounds and had between zero- and two-years' experience in different games out

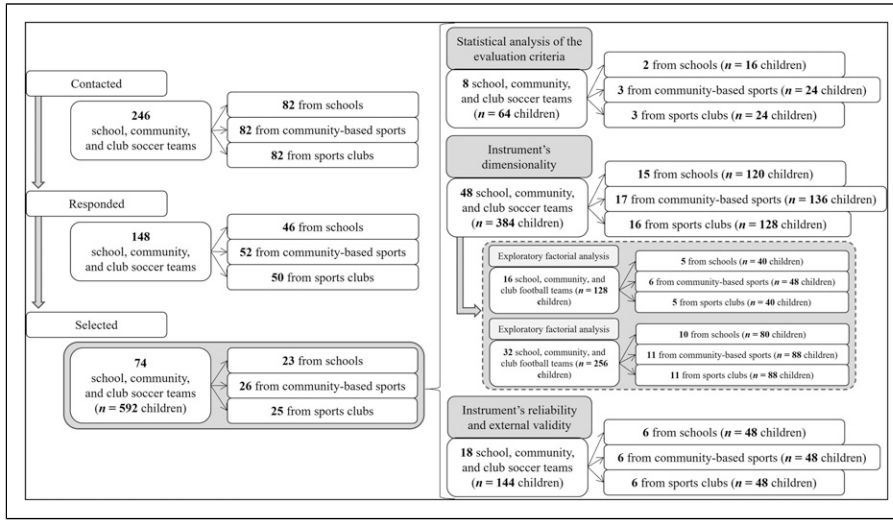

Figure 1. Participants for Each Phase of Analysis.

of PE. Their schools were coeducational, urban, private, and state, and they were both religious and non-religious. Participants received two 50–60 min PE lessons per week.

For the statistical analysis of the evaluation criteria, participants were 64 children from 8 PE classes or teams (16 children from 2 PE classes at schools, 24 children from 3 teams at community-based sports, and 24 children from 3 teams at sports clubs). For the study of the instrument's dimensionality, participants were 384 children from 48 PE classes or teams (120 from 15 PE classes at schools, 136 from 17 teams at community-based sports, and 128 from 16 teams at sports clubs). Of them, particularly for the exploratory factorial analysis, participants were 128 children from 16 PE classes or teams (40 from 5 PE classes at schools, 48 from 6 teams at community-based sports, and 40 from 5 teams at sports clubs). Regarding to the confirmatory factorial analysis, participants were 256 children from 32 PE classes or teams (80 from 10 PE classes at schools, 88 from 11 teams at community-based sports, and 88 from 11 teams at sports clubs), being at least a ratio of 10 participants for each criterion according to Kline (2016) and following the sample size recommendations (Lloret-Segura et al., 2014; Mundfrom et al., 2005; Watkins, 2018). For the study of the instrument's reliability and provision of evidence for external validity, participants were 144 children from 18 PE classes or teams (48 from 6 PE classes at schools, 48 from 6 teams at community-based sports, and 48 from 6 teams at sports clubs).

Parents signed an informed consent document prior to data collection and children assented to participate. The Research Ethics Committee of the Universidad Católica de Murcia approved the study (CE091908), which was performed in accordance with the Helsinki Declaration.

Procedure

Each participant only played one of the 37 football games of 30 min, following the country's official rules for 8–12 year-old children as criteria for inter-game consistency. In all the games the fields were 65 × 45 m, goals were 2 × 6 m, the build-out line was the offside line, and 8 versus 8 players participated in games. We arranged the football games on ten different consecutive days from 22 June to 1 July, 2021. Each day, four matches were scheduled in two-hour increments (9AM–11AM), except for the last three days in which there were three matches scheduled within 1.5 hr (9AM–10.30AM). Participants from each context played against teams from their own institutional context and were assessed only once. They were recorded during the game, with a video camera (Everio Full HD-GZ-HD7, JVC, Japan) placed five m above the ground and transversal to the side-line. The recording allowed the observation of the entire field and all the players, on- and off-the-ball.

The video records were assessed by three research assistants. They were recruited according to three inclusion criteria: (a) PE teachers at Primary School ($M = 4.30$, $SD = .10$ years of experience); (b) football coaches in teams from schools, community-based sports, or sports clubs ($M = 5.50$, $SD = .07$ years' experience); and (c) postgraduate masters students in PE with more than two years of research experience in assessing

tactical behaviors in football. They were trained in the use of the existing TAIS for 30 hr, following the instrument's guidelines (Barquero-Ruiz et al., 2022). The observers' reliability was obtained by means of an intra-observer assessment at the end of the training process. In this assessment, they observed 100 game phases of a different game from the sample included in this study, but of the same target age and contexts assessed. Subsequently, the observers assessed the same fragment again after seven days of non-observation. The calculation method for observers' reliability consisted of comparing the relationships between the two assessments of each observer separately using three reliability coefficients (Pearson's *R*, Spearman's *Rho*, and Kendall's *Tau-B*), the Kappa coefficient and the % of agreement. Observers' reliability was adequate (Pearson's $R = .95$, Spearman's *Rho* = .95, Kendall's *Tau-B* = .95, Kappa = .86, % of agreement = 83) because Pearson's *R*, Spearman's *Rho* and Kendall's *Tau-B* were greater to .90 and close to 1, Kappa coefficient was greater to .80 and % of agreement greater to 80% (Robinson & O'Donoghue, 2007; Viladrich et al., 2017).

After the observers' training, each assessed individually between 12 and 13 games using the existing TAIS (Table 1) and following TAIS instructions (Barquero-Ruiz et al., 2022). In short, the unit of observation was each game phase (attack phase and defense phase), represented as a row on an excel sheet. Depending on the game phase, the criteria assessed were those that corresponded to player attack or defense roles. Recording observations on the excel sheet required noting the frequency of appropriate and inappropriate behaviors in each game phase.

For the organizational match level criteria, observers recorded if the criteria were met in each game phase (0 or 1). For the partial forefront organizational level and the primary organizational level criteria, observers recorded the total number of times each criterion was met in each game phase. Observers played videotapes of the games using Virtual Dub Version 1.10.4, and they were allowed to stop and rewind the videotapes as many times as necessary to identify the appropriate or inappropriate player behavior. Each observer assessed a maximum of 30 min at each assessment session to avoid the effect of fatigue or demotivation. In total, 16,456 game phases were assessed (1872 for the statistical analysis of the evaluation criteria, 10,372 for the study of the instrument's dimensionality, and 4212 for the study of the instrument's reliability and provision of evidence for external validity). The reliability of the assessment was obtained through a final inter-observer assessment, using six randomly selected games from the study (1199 game phases). The reliability of the assessment was adequate (Pearson's $R = .93$, Spearman's *Rho* = .93, Kendall's *Tau-B* = .93, Kappa = .83, % of agreement = 80).

Data Analysis

For the statistical analysis of the evaluation criteria, we calculated descriptive statistics and the level of discrimination of each criterion using the corrected correlation coefficient between the scores of the criteria and the dimensions (both for the dimensions of theoretical relevance as well as those without). The instrument's dimensionality was analyzed after calculating the Kaiser-Meyer-Olkin (*KMO*) measure of sampling

adequacy, Bartlett's sphericity test, and the value of Mardia's coefficient to assess multivariate normality. The internal structure was studied through an exploratory factorial analysis of the main components and a confirmatory factorial analysis, using the maximum likelihood estimate. This study used two methods of oblique rotation, direct Oblimin and Promin, to obtain the maximum simplicity in the interpretation of the factorial solution (Lloret-Segura et al., 2014). Factor loadings are excellent if the values are equal to or greater than .70 (Hair et al., 2018).

The following indicators of the fit of the structural equation model was calculated: the chi-square ratio of the degree of freedom (χ^2/df), the Tucker-Lewis index (*TLI*), the comparative fit index (*CFI*), the normed fit index (*NFI*), the good fit index (*GFI*), the Root Mean Square Error of Approximation (*RMSEA*), and the Standardized Root Mean Square Residual (*SRMR,* Marsh et al., 2005). Normality and linearity statistics are appropriate if the value of X^2/df is less than 5, the incremental indices (*TLI*, *CFI*, *NFI*, *GFI*) are equal to or greater than .90, the *RMSEA* index is less than .08, and the *SRMR* index is close to zero (Marsh et al., 2005). Standardized regression weights are suitable when values are equal to .40 or higher (Hair et al., 2018). Standardized residual covariances were provided, with appropriate values considered to be those below 2.58 (Hair et al., 2018). Average variance extracted (AVE) was also computed. The AVE scores should be .50 or higher (Hair et al., 2018). Internal consistency was examined using the Cronbach alpha coefficient and the Omega McDonald coefficient. Reliability coefficients display an acceptable level of consistency with values as low as .70 (Viladrich et al., 2017).

As a part of the provision of evidence for external validity, first, we determined that the same construct was assessed across the three institutional contexts through three confirmatory factorial analyses, one confirmatory factorial analysis for each context. Tests of normal distribution and homogeneity (Kolmogorov–Smirnov and Levene's) were conducted before analysis. As data were normally distributed and homogeneous, the external validity was analyzed through a factorial ANOVA with Scheffé test to explore the differences between institutional contexts in all the dimensions (latent variables). We also calculated effect size using partial eta squared (ηp^2). We set statistical significance at $p < .05$. Data analysis was carried out with the statistical packages IBM SPSS v27 and AMOS v27.

Results

Evaluation Criteria

We eliminated 13 criteria from the existing TAIS (Table 1) because of low discriminatory power, negative correlation coefficients in favor of the corresponding dimension, and less than two decimal point difference in the positive correlation coefficients in favor of the corresponding dimension (appropriate and inappropriate positional attack, appropriate and inappropriate counter-attack, inappropriate amplitude, appropriate depth, appropriate defense in zone, inappropriate individual defense,

support, appropriate and inappropriate shoot, appropriate and inappropriate tackle or charging, Table 2). From a theoretical perspective, eliminating these criteria did not prevent the retention of other criteria corresponding to the four game phases, all learners' roles, and the three organizational tactical levels of play. The results of the discrimination indexes of the selected criteria and of the corrected criterion total correlation were between .51 and .79 (Table 2). This revised version of the TAIS was drawn up with 12 criteria, organized provisionally into the four theoretical dimensions (appropriate attack, inappropriate attack, appropriate defense, inappropriate defense).

Table 2. Analysis of the Discrimination Index of the Tactical Assessment Instrument in Football.

Dimensions and criteria	Mean if we eliminate the element	Variance if we eliminate the element	Corrected criterion-total correlation	Cronbach alpha if we eliminate the element
Dimension 1: Appropriate attack				
Appropriate amplitude	.66	1.39	.64	.75
Appropriate pass	.78	1.08	.68	.71
Appropriate dribbling	.68	1.61	.79	.69
Dimension 2: Inappropriate attack				
Inappropriate depth	.67	1.54	.54	.67
Inappropriate pass	.65	1.06	.57	.65
Inappropriate dribbling	.66	1.55	.52	.64
Dimension 3: Appropriate defense				
Appropriate individual defense	.66	1.41	.53	.73
Appropriate defensive coverage	.62	1.26	.60	.64
Appropriate interception	.61	1.16	.65	.69
Dimension 4: Inappropriate defense				
Inappropriate defense in zone	.75	1.59	.52	.74
Inappropriate defensive coverage	.72	1.82	.51	.71
Inappropriate interception	.69	1.31	.51	.69

Structural Dimensionality

Our study of the instrument's dimensionality confirmed the suitability of the Kaiser-Meyer-Olkin sampling (*KMO* = .80), Bartlett's sphericity test ($p < .001$), and multivariate normality (Mardia coefficient = 157.10). In the exploratory factorial analysis, the 12 criteria were statistically grouped into the four theoretical dimensions (Table 3), showing adequate variance percentages and factor loadings. The minimum value of factor loading was .71 and the maximum .86.

All the indexes of the confirmatory factorial analysis showed suitable goodness of fit (Figure 2): $\chi^2/df = 4.09$, *TLI* = .90, *CFI* = .90, *NFI* = .90, *GFI* = .90, *RMSEA* = .05 (90% confidence interval = .04–.05, *p*-close = 1), and *SRMR* = .04. Standardized regression weights were appropriate with values equal or higher to .41 (Figure 2). Standardized residual covariances were below 2.38 (.011). The AVE was equal or greater to .50 in all dimensions (appropriate attack = .53, inappropriate attack = .50, appropriate defense = .54, inappropriate defense = .50).

Internal Reliability. The four dimensions of the TAIS presented Cronbach alpha reliability coefficients between .70 and .76 (Appropriate attack: $\alpha = .71$, Inappropriate attack: $\alpha = .70$, Appropriate defense: $\alpha = .71$, Inappropriate defense: $\alpha = .76$). Omega McDonald coefficients were between .71 and .83 (Appropriate attack: $\omega = .83$, Inappropriate attack: $\omega = .72$, Appropriate defense: $\omega = .71$, Inappropriate defense: $\omega = .75$).

Table 3. Exploratory Factorial Analysis of the Four Structural Dimensions of the Tactical Assessment Instrument in Football.

Dimensions and criteria	Direct oblimin	Direct promin
Dimension 1: Appropriate attack		
Appropriate amplitude	.84	.84
Appropriate pass	.85	.85
Appropriate dribbling	.77	.77
Dimension 2: Inappropriate attack		
Inappropriate depth	.86	.86
Inappropriate pass	.79	.79
Inappropriate dribbling	.74	.74
Dimension 3: Appropriate defense		
Appropriate individual defense	.71	.71
Appropriate defensive coverage	.80	.80
Appropriate interception	.86	.86
Dimension 4: Inappropriate defense		
Inappropriate defense in zone	.79	.79
Inappropriate defensive coverage	.71	.71
Inappropriate interception	.73	.73

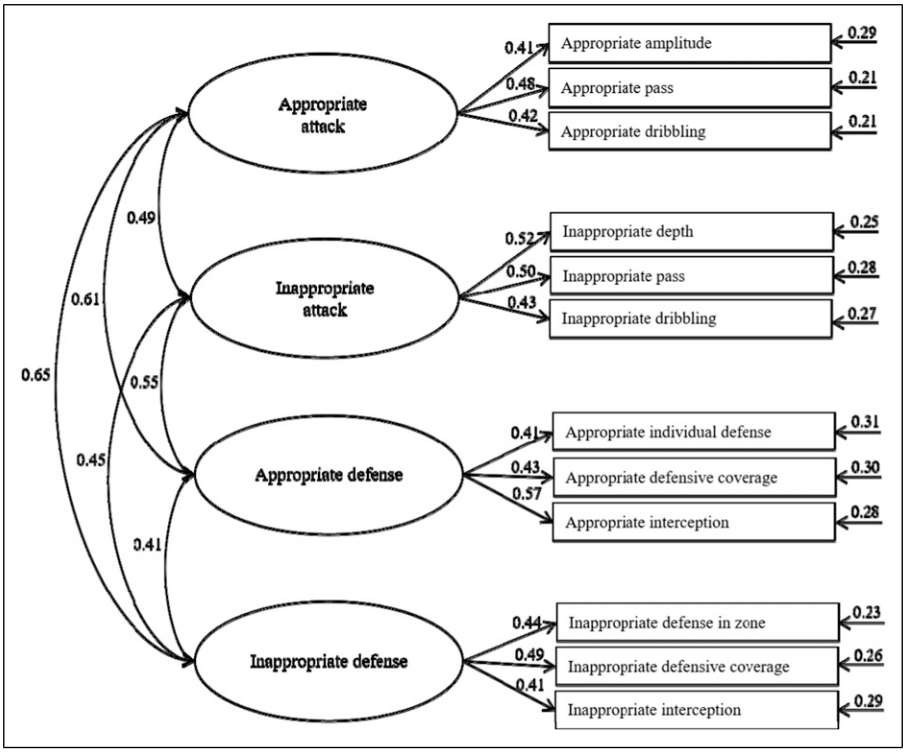

Figure 2. Model of Structural Equations and Standardized Regression Weights for Structural Dimensions of the Tactical Assessment Instrument in Football.

External Validity

First, the three confirmatory factorial analyses with the participants of each one of the three institutional contexts, respectively, showed suitable goodness of fit parameters ($\chi^2/df \leq 4.69$, $TLI \geq .81$, $CFI \geq .85$, $NFI \geq .84$, $GFI \geq .85$, $RMSEA \leq .07$ [90% confidence interval = .04–.05, p-close = 1], $SRMR \leq .05$, Table 4) and appropriate standardized regression weights (≥.40) standardized residual covariances (≤2.58 [.010]), and AVE (≥.50, Table 4).

Second, the factorial ANOVA showed statistical differences between participants of the three contexts in all the dimensions ($p < .05$, Table 5). Participants from the school context showed higher values in comparison to participants from community-based sports and sports clubs in appropriate attack ($F = 7.78$, $p < .001$, $\eta p^2 = .01$) and appropriate defense ($F = 71.41$, $p = .001$, $\eta p^2 = .05$, Table 5). In contrast, participants from the school context showed lower values in comparison to participants from community-based sports and sports clubs in inappropriate attack ($F = 7.92$, $p < .001$,

Table 4. Results of Confirmatory Factorial Analyses With Participants From Each Institutional Context.

Statistics	School context	Community-based sport context	Sports club context
χ^2/df	4.07	4.03	4.69
TLI	.81	.83	.82
CFI	.87	.88	.85
NFI	.86	.86	.84
GFI	.87	.87	.85
*RMSEA (*90% confidence interval)	.06 (.04–.05, *p*-close = 1)	.06 (.04–.05, *p*-close = 1)	.07 (.04–.05, *p*-close = 1)
SRMR	.04	.04	.05
Standardized regression weights	≥.45	≥.40	≥.45
Standardized residual covariances	≤ 2.58 (.010)	≤ 2.55 (.010)	≤ 2.44 (.012)
Average variance extracted	≥.50	≥.50	≥.50

ηp^2 = .05), and inappropriate defense (F = 12.05, p < .001, ηp^2 = .01, Table 5). However, there were no statistical differences between participants from the community-based sports and sports clubs in appropriate attack (p = 1.000), inappropriate attack (p = 1.000), and inappropriate defense dimensions (p = 1.000, Table 5).

Discussion

Our aim in this study was to provide psychometric support for the TAIS for football tactical assessment in a PE context and for youth sport teams. Consequently, we selected 12 criteria grouped into four theoretical dimensions after a statistical analysis of the evaluation criteria (appropriate attack, inappropriate attack, appropriate defense, inappropriate defense). We found empirical evidence that the internal structure of the TAIS contained our four proposed dimensions. Moreover, we found that the final criteria clusters were reliably assessed. Finally, there was external validity for the TAIS in that it discriminated between participants from the different institutional contexts. As a result of these four phases of psychometric validation, we obtained a revised version of the TAIS (Figure 2) with stronger psychometric support than any of the other existing instruments for tactical assessment. This revised version of the TAIS assesses the tactical appropriateness of player behavior in each game phase, through four dimensions that integrate criteria from the four different learner's roles and the three organizational tactical levels of play, in real-game situations, and in three institutional contexts. According to the literature, what the TAIS evaluates is crucial to an aligned

Table 5. Provision of Evidence for External Validity of the Four Dimensions of Tactical Assessment Instrument in Football.

	School context				Community-based sport context				Sports club context							
Dimensions	*M*	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*	*F*	*p*	ηp^2	*SP*
Appropriate attack	.29	.58	0	3.67	.19	.44	0	3.67	.20	.44	0	4.67	7.78	<.001	.01	.95
Inappropriate attack	.11	.19	0	.67	.17	.26	0	1.67	.16	.25	0	1.67	7.92	<.001	.05	.95
Appropriate defense	.20	.31	0	2.67	.10	.24	0	2	.06	.17	0	1.67	71.41	.001	.05	1
Inappropriate defense	.13	.26	0	1.67	.20	.33	0	1.67	.21	.35	0	4.0	12.05	<.001	.01	.99

and ecological tactical assessment in football, which can enhance the teaching-learning process in PE (Godbout & Gréhaigne, 2022; Kirk, 2017).

Relating to the statistical analysis of the evaluation criteria (comprehension validity, Carretero-Dios & Perez, 2007Carretero-Dios & Pérez, 2007), but in contrast to current instruments that are difficult to use in PE lessons (e.g., FUTSAT includes up 76 criteria; Costa et al., 2011), only 12 criteria were selected for the TAIS with corrected criterion total correlation indexes between .51 and .79 (Table 2). These appropriate values are equal to or above .40, and this supported the theoretical justification of the criteria selected based on the aligned and ecological assessment demands of tactical play in the PE context (Godbout & Gréhaigne, 2020; Kirk, 2017). Moreover, all TAIS dimensions included criteria from the four learner's roles and the three organizational tactical levels in football (Barquero-Ruiz et al., 2022; Godbout & Gréhaigne, 2022; Kirk, 2017; Laakso et al., 2022; Ribeiro et al., 2019). On the one hand, the inclusion of criteria from the four learner's roles is fundamental because game behaviors as defenders and as off-the-ball attackers are essential for assessment of tactical behaviors since players spend more time off-the-ball than on (Barquero-Ruiz et al., 2020; Laakso et al., 2022). On the other hand, evaluating the interrelated game behaviors from the three organizational tactical levels is necessary to the tactical assessment, as tactical play is built through the interactions among the whole team, small groups of players, and individual players (Godbout & Gréhaigne, 2022). Aiming for an ecological and aligned assessment of tactics in the PE context, the inclusion of information from the four learner roles and the three organizational tactical levels in football could assist players in considering those criteria in deciding when, where, and how to act to solve problems in games (Casey & Kirk, 2021; Kirk, 2017).

Regarding the study of the instrument's dimensionality (construct validity), all the existing instruments, except for TSAP, showed a construct validity analysis to check the instruments' internal structure. For this purpose, authors of these instruments only assessed participants from theoretically different skills levels, considering that differences between participants would indicate the adequacy of the instruments' internal structure (Costa et al., 2011; García-López et al., 2013; Oslin et al., 1998). Nevertheless, the internal structure cannot be based solely on theoretical assumptions or on the apparent coherence of the criteria. It is necessary to analyze statistically the unity of the criteria that are clustered in each dimension (Irwing et al., 2018). In the present work (Table 3), the 12 criteria were grouped in four theoretical dimensions. The results showed the suitability of the sampling, sphericity (Irwing et al., 2018; Mettert et al., 2020), and multivariate normality, since following the Bollen (1989) recommendations, Mardia's coefficient (157.10) was less than $p\,(p+2)$, p being the number of criteria of the instrument. Following Hair et al. (2018), factor loadings were excellent given the values were greater than .70. Standardized residual covariances were as high as 2.38, indicating the absence of significant discrepancies (Hair et al., 2018). Also, standardized regression weights were appropriate, with values equal to .41 or higher, as well as AVE scores equal to .50 or higher (Hair et al., 2018). In addition, the indexes of the confirmatory factorial analysis verified the adequate fit of the model designed

(Figure 2). The χ^2/df value was less than 5, the *RMSEA* and *SRMR* indexes were lower than or equal to .05, and the goodness of fit parameters were equal or higher than .90 (Marsh et al., 2005). These satisfactory results from exploratory and confirmatory analysis supported the theoretically justified dimensions and criteria selected (Carretero-Dios & Pérez, 2007; Mettert et al., 2020).

Concerning the instrument's reliability, the four other tactical assessment instruments presented adequate values when the instruments were under construction and based on the inter- or intra-observers' agreements (Costa et al., 2011; García-López et al., 2013; Gréhaigne et al., 1997; Oslin et al., 1998). In this respect, the reliability analysis of the TAIS showed Cronbach alpha and Omega McDonald values greater than .70 in the four dimensions, which indicates a high association between criteria in those four dimensions according to reference values (Viladrich et al., 2017). These positive results, based on the final version of the TAIS confirmed the high degree of interrelationship, homogeneity, and consistency among the criteria of each dimension (Carretero-Dios & Pérez, 2005; Irwing et al., 2018).

In relation to external validity, the TSAP study was the only one that analyzed this through consultation with two football experts (Gréhaigne et al., 1997). In contrast, in the present version of the TAIS, we decided, first, to analyze the assessment of the same construct across the three institutional contexts through three confirmatory factorial analyses, one for each context (school, community-based sports, and sports clubs), reporting suitable psychometric properties (Table 4) and, second, to compare participants of the three institutional contexts, showing differences between them in all the dimensions. These results suggest that the TAIS is useful for tactical assessment in the three institutional contexts and allows discrimination between participants who are taught from different approaches. We found that participants from the school context obtained better values than those from community-based sports and sports clubs (Table 5). These better values could be due to the higher current interest in teaching games from a tactical perspective in school PE in comparison to youth sports, given the difficulties in developing children's tactical learning that coaches perceive (Harvey & Jarrett, 2014; Kinnerk et al., 2018; Morales-Belando et al., 2022), and because the improvement of children's tactics is a learning outcome in the PE curriculum in the country where this study was conducted (Royal Decree 157/2022). Moreover, the lack of statistical differences between participants from community-based sports and sports clubs in appropriate attack, inappropriate attack, and inappropriate defense dimensions could be explained on the basis that coaches traditionally teach football in these contexts from a technical perspective (O'Connor et al., 2018; Stonebridge & Cushion, 2018). In this regard, the main contextual difference between participants from community-based sports and sports clubs resides in the fact that the second group usually train more hours per week. According to Arias-Estero et al. (2020), for tactical learning, practice time should be considered along with the teaching-learning approach used because more practice time does not guarantee tactical learning if the teaching-learning approach is not appropriate.

Limitations and Directions for Further Assessment

The instrument's temporal stability and predictive validity were not measured, and our sampling method does not allow us to generalize these results to a broad population. Also, the sequence of game presentation to observers was not counterbalanced to mitigate potential effects of order regarding fatigue or demotivation in data collection. In addition, as participants were mainly boys, studies with more girls are needed.

Conclusion

This study confirmed that the TAIS is a valid and reliable instrument with appropriate psychometric properties to assess football tactics in PE and youth sport teams (with 8–12 year old players). In practical terms, this instrument can enhance the teaching-learning process as it enables an aligned and ecological assessment through four player dimensions that correspond to two game phases of attack and defense, differentiate appropriate and inappropriate behavior for the 12 criteria from the four different learners' roles and the three organizational tactical levels of play (Figure 2).

The excellent factor loadings values, the goodness of fit parameters of the exploratory and confirmatory factorial analysis, the adequate standardized regression weights, standardized residual covariances, AVE, internal consistency, and the reported differences between the three institutional contexts as provision of evidence for external validity, theoretically and statistically, are an important milestone in support of this revised version of the TAIS. Researchers and teachers can now use the TAIS in interventions studies and instruction because it was created specifically with a tactical perspective. Of note, however, this instrument was designed specifically for football, and it would need to be validated for use with other invasion games.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Ministerio de Educación, Cultura y Deporte (FPU17/00606).

ORCID iD

Jose L. Arias-Estero https://orcid.org/0000-0002-6201-4710

References

Arias-Estero, J. L., Jaquero, P., Martínez-López, A. N., & Morales-Belando, M. T. (2020). Effects of two TGfU lessons period on game performance, knowledge and psychosocial variables in

elementary physical education. *International Journal of Environmental Research and Public Health*, *17*(10), 3378. https://doi.org/10.3390/ijerph17103378

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), 3330. 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

Barquero-Ruiz, C., Kirk, D., & Arias-Estero, J. L. (2022). Design and validation of the tactical assessment instrument in football (TAIS). *Research Quarterly for Exercise & Sport*, *93*(3), 615–632. https://doi.org/10.1080/02701367.2021.1889457

Biggs, J. (2014). Constructive alignment in university teaching. *HERDSA Review of Higher Education*, *1*, 5–22.

Bollen, K. A. (1989). *Structural equations with latent variables*, John Wiley and Sons.

Care, E., Kim, H., Vista, A., & Anderson, K. (2018). *Education system alignment for 21st century skills: Focus on assessment. Center for Universal Education*, The Brookings Institution.

Carretero-Dios, H., & Pérez, C. (2005). Normas para el desarrollo y revisión de estudios instrumentales [Standards for the development and review of instrumental studies]. *International Journal of Clinical and Health Psychology*, *5*(3), 521–551.

Carretero-Dios, H., & Pérez, C. (2007). Standards for the development and review of instrumental studies: Considerations about test selection in psychological research. *International Journal of Clinical and Health Psychology*, *7*(3), 863–882.

Casey, A., & Kirk, D. (2021). *Models-based practice in physical education*, Routledge.

Costa, I. T., Garganta, J., Greco, P. J., Mesquita, I., & Maia, J. (2011). System of tactical assessment in soccer (FUT-SAT): Development and preliminary validation. *Motricidade*, *7*(1), 69–83. https://doi.org/10.6063/motricidade.121

Doyle, W. (2006). Ecological approaches to classroom management. In C. M. Evertson & C. S. Weinstein (eds), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 97–125), Laurence Erlbaum and Associates.

FIFA. (2018). Report of the FIFA. *Global club football report 2018*. FIFA. Available at: https://digitalhub.fifa.com/m/3f37419da40770b7/original/plstmtbjzbps7c2vkfir-pdf.pdf (accessed 16 March 2022).

García-López, L. M., González-Víllora, S. G., Gutiérrez-Díaz, D. C., & Serra, J. (2013). Development and validation of the game performance evaluation tool (GPET) in soccer. *Sport TK*, *2*(1), 89–99. https://doi.org/10.6018/185791

Godbout, P., & Gréhaigne, J. F. (2020). Revisiting the tactical-decision learning model. *Quest*, *72*(4), 430–447. https://doi.org/10.1080/17408989.2020.1861232

Godbout, P., & Gréhaigne, J. F. (2022). Regulation of tactical learning in team sports. The case of the tactical-decision learning model. *Physical Education and Sport Pedagogy*, *27*(3), 215–230. https://doi.org/10.1080/17408989.2020.1861232

Gréhaigne, J. F., Godbout, P., & Bouthier, D. (1997). Performance assessment in team sports. *Journal of Teaching in Physical Education*, *16*(4), 500–516. https://doi.org/10.1123/jtpe.16.4.500

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). *Multivariate data analysis* (8th ed.), Cengage Learning EMEA.

Harvey, S., & Jarrett, K. (2014). A review of the game-centred approaches to teaching and coaching literature since 2006. *Physical Education and Sport Pedagogy*, *19*(3), 278–300. https://doi.org/10.1080/17408989.2012.754005

Irwing, P., Booth, T., & Hughes, D. J. (2018). *The wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development*, John Wiley and Sons.

Kinnerk, P., Harvey, S., MacDonncha, C., & Lyons, M. (2018). A review of the game-based approaches to coaching literature in competitive team sport settings. *Quest*, *70*(4), 401–418. https://doi.org/10.1080/00336297.2018.1439390

Kirk, D. (2017). Teaching games in physical education: Towards a pedagogical model. *Revista Portuguesa de Ciências do Desporto*, *17*(S1.A), 17–26. https://doi.org/10.5628/rpcd.17.S1A.17

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.), Guilford Press.

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & T omás-Marco, I. (2014). Exploratory item factor analysis: A practical guide revised and updated. *Anales de Psicología*, *30*(3), 1151–1169. https://doi.org/10.6018/analesps.30.3.199361

Laakso, T., Davids, K., Luhtanen, P., Liukkonen, J., & Travassos, B. (2022). How football team composition constrains emergent individual and collective tactical behaviours: Effects of player roles in creating different landscapes for shared affordances in small-sided and conditioned games. *International Journal of Sports Science and Coaching*, *17*(2), 346–354. https://doi.org/10.1177/17479541211030076

MacPhail, A., Kirk, D., & Griffin, L. (2008). Throwing and catching as relational skills in game 788 play: Situated learning in a modified game unit. *Journal of Teaching in Physical 789 Education*, *27*(1), 100–115. https://doi.org/10.1123/jtpe.27.1.100

Marsh, H. W., Hau, K. T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares & J. McCardle (eds), *Contemporary psychometrics: A festschrift to Roderick P. McDonald* (pp. 275–340), Laurence Erlbaum and Associates.

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

Mettert, K., Lewis, C., Dorsey, C., Halko, H., & Weiner, B. (2020). Measuring implementation outcomes: An updated systematic review of measures' psychometric properties. *Implementation Research and Practice*, *1*(2), 1–28. https://doi.org/10.1177/2633489520936644

Morales-Belando, M. T., Kirk, D., & Arias-Estero, J. L. (2022). A systematic review of Teaching Games for Understanding intervention studies from a practice-referenced perspective. *Research Quarterly for Exercise & Sport*, *93*(4), 670–681. https://doi.org/10.1080/02701367.2021.1897066

Morley, D., Van Rossum, T., Richardson, D., & Foweather, L. (2019). Expert recommendations for the design of a children's movement competence assessment tool for use by primary

school teachers. *European Physical Education Review*, *25*(2), 524–543. https://doi.org/10.1177/1356336X17751358

Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing*, *5*(2), 159–168. https://doi.org/10.1207/s15327574ijt0502_4

O'Connor, D., Larkin, P., & Williams, A. M. (2018). Observations of youth football training: How do coaches structure training sessions for player development? *Journal of Sports Sciences*, *36*(1), 39–47. https://doi.org/10.1080/02640414.2016.1277034

Oslin, J. L., Mitchell, S. A., & Griffin, L. L. (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

Parlebas, P. (2001). *Juegos, deporte y sociedad. Léxico de praxiología motriz*. Paidotribo.

Piaget, J. (1977). *The development of thought. Equilibration of cognitive structures*, Basil Blackwell.

Ribeiro, J., Davids, K., Araújo, D., Guilherme, J., Silva, P., & Garganta, J. (2019). Exploiting bi-directional self-organizing tendencies in team sports: The role of the game model and tactical principles of play. *Frontiers in Psychology*, *10*(2213), 1–8. https://doi.org/10.3389/fpsyg.2019.02213

Robinson, G., & O'Donoghue, P. (2007). A weighted kappa statistic for reliability testing in performance analysis of sport. *International Journal of Performance Analysis in Sport*, *7*(1), 801–819. https://doi.org/10.1080/24748668.2007.11868383

Royal Decree (157/2022). por el que se establecen la ordenación y las enseñanzas mínimas de la Educación Primaria. *Ministerio de Educación y Formación Profesional*.

Stonebridge, I., & Cushion, C. (2018). An exploration of the relationship between educational background and the coaching behaviours and practice activities of professional youth soccer coaches. *Physical Education and Sport Pedagogy*, *23*(6), 636–656. https://doi.org/10.1080/17408989.2018.1485143

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Anales de Psicología*, *33*(3), 755–782. https://doi.org/10.6018/analesps.33.3.268401

Vors, O., & Kirk, D. (2016). Classroom life and the importance of context in physical education: From ecological approaches to situationist theories. *Recherches & Education*, *15*, 15–28. https://doi.org/10.4000/rechercheseducations.3033

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, *44*(3), 219–246. https://doi.org/10.1177/0095798418771807

Author Biographies

Carmen Barquero-Ruiz is an Assistant Professor within the Physical Education and Sport Sciences (PESS) Department at the University of Limerick (Ireland). Her main research and teaching interests revolve around teaching games in physical education and initial teaching education between different cultural context and focusing on the benefits of an aligned assessment within invasion games.

David Kirk is Professor of Education in the School of Education at the University of Strathclyde. His main teaching and research interest revolve around educational innovation, curriculum history, and physical education and sport pedagogy.

Lourdes Meroño is Senior Lecturer and Dean in the Faculty of Sport at UCAM Catholic University of Murcia (Spain). Her main research and teaching interests revolve around (physical education) teacher education, and innovative pedagogies for teaching and learning.

José L. Arias-Estero is a Senior Lecturer in the Facultad de Educación at Universidad de Murcia (Spain). His main teaching and research interest revolve around pedagogical approaches for teaching and learning in physical education and initial teaching education.