Design and validation of a tool for the formative assessment of invasion games

FERNANDO MANUEL OTERO-SABORIDO¹; JOSÉ ANTONIO GONZÁLEZ-JURADO²

^{1,2} Department of Sport, Pablo Olavide University, Seville, SPAIN

Published online: June 26, 2015

(Accepted for publication May 10, 2015)

DOI:10.7752/jpes.2015.02039;

Abstract:

The present work aims to design and validate a tool for formative assessment of invasion games in the context of school for future application in situations of peer evaluation. 4 experts in the field participated in the research. For this, a validation of the content was made with the Tool for Assessment and Learning of an Invasion Situation (TALIS). The purpose of the observation tool was to evaluate the ability to occupy an empty space standing on a passing line, a common situation in any invasion game. Afterwards, the tool was applied, observations were coded and the intraobserver and interobserver agreements were determined. Finally, without receiving prior training, 22 students of fourth level of primary education coded TALIS and the agreement between their observations was found. Data were analyzed from two points of view: interobserver (two expert observers) and intraobserver (in two observations with a gap a fortnight). From the point of view of the experts' analysis, the results reveal that the degree of agreement on the three TALIS variables is close to 90%. Similarly, high levels of agreement are obtained in both interobserver index (DTH: 0.908; DS: 0.840; AFD: 0.954) and intraobserver. In the case of observations of students, the degree of agreement is remarkable (DTH: 0.795; DS: 0.747; AFD: 0.783) especially when you consider that there was no specific training in TALIS. We have validated a tool for formative assessment of invasion games at school with a remarkable statistic delivery.

Key words: assessment; team sports; pupils; school; learning; physical education.

Introduction

The general concepts of evaluation, have their basis on the division of knowledge of Habermas (1997). That distinction can be summarized in the technical rationalism and practical rationalism. Each model has specific characteristics that they carry to their evaluative approaches (López 1999). In the case of technical rationalism, evaluative intention is the technical verification of the achievement of objectives and efficiency through measurement. Teachers are passive executors of measurement designed by experts. On the other hand, the evaluation under practical rationalism is not intended to measure the results but to understand them. The verification of efficacy disappears in favor of the formative value. That is, the end of the evaluation is to serve the learning. Scriven in 1967 coined this conception as "formative assessment" (Jorba and Sanmartí 2013). Although the first antecedent associated with formative assessment is in the pedagogy of Dewey (Dewey 1916) that had the precept of learning centered on experience and its subsequent thought. In short, this is the procedure that follows a formative evaluation: acting and then reflecting on the action and generate learning in a new situation. That is the sequence that develop specific strategies for formative assessment, self-assessment, and peer assessment.

In the case of Physical Education (PE), the assessment has been rated as one of the most controversial aspects of the teaching-learning process in this area. (James, Griffin, and France 2005; Otero, Calvo, and González-Jurado 2014, Tejada 2007). The export of previous models in PE in Spain has a uniform description. A first model geared towards physical performance that aims to measure the efficacy of physical student achievement through tests and standardized tests. This view does not consider the educational value of the assessment. Its intention sums up the question of López (2004) "Why do we say assessment when we really mean rating?" In contrast, the model oriented towards student participation does have a high educational value. It does not assess the driving aspect, but also considers the emotional, social and cognitive dimensions through the involvement of students in the assessment (López, 1999). Different studies have shown the benefits of involving the students to situations using self-assessment as well as peer or co-assessed assessments (Lamb, Lane, and Aldous 2013; Lopez et al 2007; Perez, Heras, and Hernán 2008; Vera Gonzalez, and Moreno 2009; NiChroinin and Cosgrave 2013).

The reviews of different papers on assessment practices in different contexts and times reveal that while the theoretical paradigm indicates the prevalence of a formative assessment, assessment practices of teachers in PE still dominate scoring models with little involvement of the students (Lorente-Catalan and Kirk 2014; Otero 2013).

Although there has been an increase in research on formative assessment in PE, there are still few studies that investigate tools such as peer assessment and self-assessment. In the same line of moderation run the formative assessment practices in PE classes despite contributions such as the Network of Formative and Assessment in Higher Education (López et al. 2007).

Furthermore, the invasion games which are the object of this paper, focus on PE preferences. Both physical education teachers as well as learners from different countries agree to designate them as, in the case of teachers, the most taught, and in the case of students the preferred choice (Shropshire and Carroll 1998; Hill and Cleven 2005; Gutiérrez, Pilsa, and Torres 2007).

In the case of invasion games there exist tools validated that involve the students in their assessment process: TSAP (Gréhaigne, Boutiher, and Godbout 1997) GPAI (Oslin, Mitchell, and Griffin 1998); BOGPI (Chen, Hendricks, and Zhu 2013) and HERJ (Garcia et al. 2013). However, the fact that there are validated tools does not ensure that a formative assessment is being performed. Some are excessively quantitative, emphasizing the amount of motor actions, not in the quality thereof. Others are too complex, they possess too many dimensions that have to be observed live by the students whose discrimination capabilities are limited (Otero 2005) and therefore its use in formative assessment processes is made difficult.

Occasionally, having established a high reliability, interobservers are used for summative assessment (non-formative) due to its implementation difficulties in the process of student participation. In this sense, we have already pointed out the real difficulties (lack of these tools by teachers, diligence process: tool design, recording and editing situations, training of observers, etc...) to implement an assessment of invasion games in PE (Otero and Silva 2015). Likewise, no evidence that these tools have been incorporated into PE classes has been found (Alarcón et al 2010; Lorente-Catalan and Kirk 2014; Sicilia et al., 2006).

Nonetheless, despite the complexity of performing formative assessment in the teaching of invasion games, the few studies on the subject have shown the benefits to students over declarative knowledge, decision-making and execution technique on invasion games such as handball, basketball or football in the school context (Santos 2010; Pérez, Heras, and Hernán 2008; Figueiredo, Lake, and Fernández 2008).

Therefore, it is the objective of this work to design and validate a tool that serves formative assessment of invasion games in the school context. To do this, we contemplated the difficulties teachers of PE have to evaluate invasion games such as heterogeneity of the student or the multiplicity of solutions and / or behaviors to be evaluated (Otero and Silva 2014). Moreover, considering that the discrimination capacity of the human being is limited (Esnaola et al., 2003) a tool has been designed warning about the viewpoint of the student (Sanchez-Mora et al 2011; Sicilia 2003) and without imposing only the teacher's point of view (Rivadeneira and Sicily 2004). At the same time, the quality of motor actions on invasion games has been prioritized on the amount thereof. Therefore this research has the overall aim to design and validate a tool for formative assessment of invasion games within the school context. Its operational objectives aim to:

- to obtain a validation of the content of the tool of observation.
- to determine the interobserver agreement using two experienced observers.
- to determine the intraobserver agreement of two expert observers at different times.
- to determine the interobserver agreement among students of Physical Education.

Method

This is a quantitative observational study, using a structured tool for collecting data. As regards to observational studies in sports, we have a nomothetic design (as there is a plurality of units to assess), punctual (there is no intrasessional monitoring) and multidimensional (Anguera and Hernandez-Mendo 2014).

Participants

A panel of experts participated in the research. They were all PE teachers. Two were PE teachers in primary school and the other two had specific training in invasion games and were university teachers of PE. The four participated in the validation of content and only the university teachers participated in the interobserver and intraobserver tests accordingly. Moreover, 22 students of fourth level of Primary Education in Sevilla (Spain) participated in the observation sequences to determine the degree of agreement between observations. The mean age was 8.36 ± 0.49 years of which 49.9% were girls and 50.1% were boys.

Materials

Game sequences were recorded with two Sony DSC-HXU cameras. The coding of the data recorded by the subjects was performed in Microsoft Office Excel 2007 and for their analysis SPSS V.18 for Windows was used. The issue of different sequences of play to be observed by the participants was conducted with the Dartfish Team Pro version 5.5 program adapted for invasion games.

Procedure

For the design of the tool we followed several steps (Figure 1). First, an analysis on the specialized literature under study, and the subsequent extraction of conclusions. Secondly, the design of game situations in students and their observation by the panel. Third, the design of the tool with a content validation by experts who

drew upon the findings of the literature review and the observed movements in game situations. Finally, we proceeded to the coding sequences by expert observers and subsequently by the chosen students.

Fig. 1. Phases in the design of the tool of formative assessment.

Bibliographical analysis and conclusions (Phase 1)

In the first place, there was a comprehensive review of the literature specializing in structuring and detail as shown in Table 1. Firstly, the general part of our object of study (formative assessment in education) was revised up to the most specific (design and validation of tools of observation in PE).

Table 1. First step in designing the tool. Analysis of the specialized literature.

Knowledge domain	Sources
Formative assessment	Jorbá and Sanmartí (2013) Santos (2002) Perrenoud(1991)
Formative assessment in PE	López (1999) López et al. (2007) Lorente-Catalán and Kirk (2014)
Teaching and assessment of invasion games	Ardá and Casal(2003) Bengué (2005) Fradua and Moreno (2001) Hernández (2000) Memmert and Harvey (2010) Méndez (2010) Otero (2004) Otero, Calvo, González-Jurado(2014) Otero et al. (2014) Sánchez-Mora et al. (2010) Forrest (2013)
Formative assessment in invasion games	Oslin, Mitchells, and Griffins (1998) Chen, Hendricks, and Zhu, (2013) Gréhaigne, Bouthier, and Godbout (1997) Figueiredo, Lago, and Fernández (2008) Otero-Saborido, Calvo, González-Jurado (2015) Pérez, Heras, and Hernán (2008) Nadeau et al. (2008) Santos (2010)
Design and validation of tools for observation	Anguera and Hernández-Mendo (2014) Etxeazarra., Castellano, J., Usabiaga, (2013) Pope et al. (2002) Pradas et al. (2012)

They came to the following specific conclusions about the design of the tool after the reading and analysis of the aforementioned documents:

First, cognitive and motor skills are more heterogeneous in school students than in the students within a sporting context. Therefore, the observation tool must be comprehensible to all students regardless of their level of motor intelligence and / or technical execution in invasion games. In this regard, and taking into account that a design principle of an object of observation is to narrow down that which is observed, and that the discrimination ability is limited, the design of a tool is proposed with the required dimensions and with a very precise definition of the categories that does not support errors in its codification. In the present study, once the tool was coded by the experts, the degree of agreement with pupils of Primary fourth level will be determined.

Secondly, according to the literature review in the teaching of invasion games in the school context, it is recommended to start from situations of numerical superiority favoring the point of view of the student when designing the game situation that has to be assessed.

Thirdly, the design of the tool aims to be used in future training purposes for peer review. Therefore, the variables to observe should be directed towards improving the quality of student movements (within a global context of play) above quantitative aspects.

Secondly, according to the literature review, in the teaching of invasion games in the school context it is recommended to start from situations of numerical superiority favoring the point of view of the student in the design of the game situation to assess.

Thirdly the design of the tool aims to be used in future training purposes for peer review. Therefore, the variables to observe should be directed towards improving the quality of student movements (within a global context of play) above quantitative aspects.

Design and observation of game situations (Phase 2)

For the observation of game situations, a teaching unit on basic principles of invasion games was implemented in 77 primary school students. The situation observed in this study was a passing game that pitted two teams. The teams always ran in a context of numerical superiority thanks to the existence of a neutral player. They passed the ball with their hands, prioritizing the mechanism of decision taking and perception and to the technical difficulty of children playing with their foot within a designated area. There were no physical objectives, the goal was to reach a number of passes that increased depending on the evolution of the pupils. Using videos, the students were taught about important factors when clearing, as well as on the occupation of open spaces and the speed of displacement. The passes that saved an opponent by throwing the ball above them were not counted. We stimulated the clearings which had passing lanes occurred in areas with little adversary density. Four observers watched all the images recorded to make a list of possible behaviors to be included in the designing of the tool.

Design and validation of the tool (Phase 3)

After viewing the games of passes, the panel got the bibliographical conclusions of the tool to design. It was reported that the purpose of the observation tool was to assess the ability to fill a gap standing in line of pass. After specifying the objective of the tool, the purpose of it, the game situation proposal and the application context, experts were urged to make a proposal for a tool. In the first movement each expert made his proposal tool. In the second movement the four proposals were merged into one. The observation tool had first proposed four variables with two categories each. After a pilot with a group of students in Primary, the complexity that students had to observe four variables was verified in vivo. In the third movement of the panel two variables were integrated into one leaving the final tool in three variables: 'Trajectory displacement with respect to holder' (TDH), 'Scroll Speed' receptor (DS) and 'density opponents and colleagues' (AFD) in the clear. The notes taken following the observations of students forced to redefine some categories in a more accurate coding to avoid confusion. The tool for the assessment and learning situations invasion (TALIS) was finally as follows:

Table 2. Variables of the tool for the assessment and learning of situations invasion (TALIS).

Variables	Categories	Definition	Coding
(DTH) Displacement trajectory in regards to the holder	AC	Approaches holder. Offensive space of action is reduced	0
	AL	It moves away from the holder. Increases the offensive action space.	1
(DS) Displacement speed	L	Walks or jumps advancing little space	0
	R	Jogs runs or sprints	1
(AFD) Adversary and fellows' density	EOCA	The receiver receives in a space shared with adversary and fellow. Increases the probability of interception and action space for	0

		attacker is reduced.
	EOA	The receiver receives in a space shared with the adversary. 1 Increases the probability to intercept
	EOC	The receiver receives in a space shared with a fellow potential receiver. The proximity of both 2 reduces the space of offensive action
	EL	The receiver receives in a free space. Not occupied by fellows 3 or adversaries.

Coding (Phase 4)

A total of 123 sequences of passing lanes were shot and edited. Those whose shots were inadequate for observation and/or coding were eliminated. Finally, we made a random selection of 44 sequences. The two expert observers coded the 44 sequences of passing lanes using the tool for assessment and learning of situations of invasion (TALIS). Since both had participated in developing the tool it was not necessary to give a specific training. However, there did exist a unification of criteria on categories that could present doubts. After 15 days they returned to begin the coding in order to assess the reliability of the intraobservers.

After the encoding done by experts, the TALIS to 22 students in the fourth level of primary education was applied. Students were not subjected to a specific training tool. They received only explanation in a session where they watched examples of the three categories with examples of each of them. Subsequently, encodings experts and students were subjected to statistical analysis.

Statistical analysis

For the construction of TALIS, different qualitative and quantitative strategies were used. For the qualitative aspect, the opinion of experts to validate the contents of the tool was taken into account. Quantitatively analyzed from two points of view: interobserver (two expert observers) and intraobserver (in two observations with a gap of fifteen days). In the first case, it was analyzed from the descriptive perspective and from the concordance. Descriptively by analyzing the percentages of agreement between the two observers in all categories of the tool. For agreement of the experts, we used Cohen's Kappa index (k), which determines the proportion of cases in which there is agreement between two observers after excluding the proportion of cases in which the agreement between them is due to chance. To measure the degree of agreement between the observations of the student's agreement between them, we used the Kendall test as indicated for 3 or more observers.

Results

Following the steps described in the method section, the resulting tool TALIS contained 3 dimensions with a total of 8 categories. Table 3 shows frequency and percentage of agreements and disagreements in the three categories. An initial exploration of the results from a descriptive point of view reveals that the agreement between the expert observers is very high in all categories (Table 3). In the case of DS and environment AFD it reaches 95% of agreements in the 44 observed sequences. The DTH category obtains 93.2% of agreements.

Table 3. Frequencies and percentages of agreement and disagreement between the two expert observers of TALIS

Categories	N	Agreements		Disagreements	
		Frequency	%	Frequency	%
DTH	2	41	93.2	3	6.8
DS	2	41	93.2	3	6.8
AFD	2	42	95.5	2	4.5

DTH: displacement trajectory in regards to the holder; DS: Displacement speed; AFD: Adversary and fellows' density

Subsequently, Table 4 shows the interobserver agreement between the two experts. The results show values around 0.9 (almost perfect) in the three categories of TALIS according to Cohen's Kappa index (k) determining the proportion of cases in which there was no agreement among observers after excluding the proportion of cases where the agreement between them is the result of chance (DTH: 0.908; DS: 0.840; AFD: 0.954). The three dimensions get high values of significance (Table 4).

Table 4. Interobserver agreement among experts that codified TALIS

Level of agreement Kappa index	Agreement strength	Categories	N	Value	Mean error	Approx. T	Sig.
0,00	Poor						
0,01 - 0,20	Slight	DTH	2	,908	,076	5,776	,000
0,21 - 0,40	Fair						
0,41-0,60	Moderate	DS	2	,840	,089	5,579	,000
0,61-0,80	Substantial						
0,91-1,00	Almost perfect	AFD	2	,954	,085	7,041	,000
,00	Poor						

DTH: displacement trajectory in regards to the holder; DS: Displacement speed; AFD: Adversary and fellows' density

Finally, Table 5 shows the degree of agreement between two measurements at two different times in each of the observers. It took 15 days between the two measurements. The results indicate that both observers can make an appropriate and accurate tool use and there is a high stability in the results. Both observers obtain almost perfect values in the AFD and DTH categories and considerable on DS. The three categories obtain high levels of significance (Table 5).

Table 5. Interobserver agreement in coding IASI. Observer A and B

Categories	Agreement value		Mean error		Approx. T		Sig.	
	Observer		Observed		Observer		Observed	
	A	B	A	B	A	B	A	B
DTH	,908	,954	,064	,045	6,046	6,337	,000	,000
DS	,840	,891	,089	,075	5,579	5,909	,000	,000
AFD	,936	,944	,063	,055	7,509	7,399	,000	,000

DTH: displacement trajectory in regards to the holder; DS: Displacement speed; AFD: Adversary and fellows' density

Since we have a formative assessment tool to use with students, Table 6 shows the results of matching observations from TALIS by students. Very similar coefficients are obtained in all three categories (DTH: 0.795; DS: 0.747; AFD: 0.783) and slightly lower when compared with the experts. The three results are highly significant in all three categories (Table 6).

Table 6. Interobserver agreement among students encoded with IASI

Categories	N	Kendall's W	Chi-squared	Sig. asin.	gl.
DTH	22	,795	751,634	.000	43
DS	22	,747	706,778	.000	43
AFD	22	,783	740,459	.000	43

DTH: displacement trajectory in regards to the holder; DS: Displacement speed; AFD: Adversary and fellows' density

According to Landis and Koch approaches (1977), who proposed the margins to assess the degree of agreement based on the index k (Table 3), TALIS's excellent statistical results obtained on the observations of experts, as the strength of the agreement is 'significant' and 'almost perfect' in most measurements. For its part, the degree of agreement on the observations of students is "considerable".

Discussion

The design of observation tools for student participation in invasion games has multiple antecedents (Chen, Hendricks, and Zhu 2013, Gréhaigne, Godbout, and Bouthier 1997, Oslin, Mitchell, and Griffin 1998). Some of these tools establish general variables to specify in any invasion game (decision making, technical execution ...). However, it is left to the discretion of the teacher or applicator tool to define and specify variables for each case where no one can know whether the formulation of a category is relevant (eg definition of when a decision is taken or technical execution considered right or wrong).

In this sense TALIS defines precisely in three dimensions (with their respective categories) the team in possession of the ball thereof for any of the invasion games. There is therefore room for doubt or ambiguity in its application. The teacher needs the exact parameters for coding. Other inputs from TALIS is having a clear intention of formative assessment, ie a tool of reflection and improvement of the student. While other validated tools offer the students only practical quantitative data (number of shots on goal, number of turnovers ...) that do not result in an improvement process, the dimensions of TALIS allow the students to reflect on the quality of the actions (DTH, DS and AFD) without prioritizing the quantitative aspect.

By gambling on ball possession, it brings the greatest amount of time and distance of the total volume of the game in invasion games. Between 80% and 95% of players playing any of these modalities are not in possession of the mobile one, hence its importance and relevance in the tool of this object of study (Avila 1999; Dellal et al 2011 ; Cárdenas 1996; Collet 2013; Di Salvo et al 2009; Perez, Heras, and Hernán 2008; Alamo 1996; Wein 1995).

Within the attack phase, the game without the ball is performed by the 'team player without the ball' (Hernández 2000) or 'attacker without ball' (Wein 1995). And within that role, the displacement (either toward or away from the player with the ball) is the technical and tactical ability most representative (Vegas 2007; King 2012; Serra, Gonzalez and Garcia 2011). Therefore, as a major contribution, the TALIS tool has focused on the ability of the role in which the attacking players spend more time of the game and travel greater distances.

Moreover, the game situation used in TALIS, a passing game with their hands, including tactical behaviors common to all invasion games. These patterns can be transferred between sports if you opt for a model of horizontal education that brings different modes of invasion (Bengué 2005; Hernandez and others 2000). In this sense, although the literature had theorized about the desirability of a joint teaching in school of invasion games, there are experimental studies which have reinforced in different ways the horizontal model based on the results obtained (Memmert and Harvey 2010; Ramirez 2009; Yanez and Castejón 2011). These results empirically confirm the possibility of transferring learning between sports disciplines included in this group of sports. Therefore, TALIS and situation of assessment given between themselves is very versatile in the school curriculum as it is feasible to implement in many invasion games (basketball, handball, football, lacrosse, rugby, ultimate, hockey ...).

In that vein, work on education invasion games emphasize the importance of teaching to attack rather than defend in the early formative stages (and therefore in the school context object of this paper) and, above all, the importance of technical-tactical contents the team in possession of the ball (displacements, shimmies) (Vegas 2007; Painter and Cardenas 2001; Fradua and Moreno 2001) In this sense, TALIS is not only a tool for very important phase of the game, but it does so on the ability not very relevant training ball ages in school context.

Moreover, levels of agreement obtained in TALIS both from the perspective of interobserver and intraobserver resemble those achieved by different researches with certain equivalences to this (Pope et al 2002; Pradas et al 2012; Reina-Gómez, Hernandez-Mendo, and Fernández-García 2010). In either case, as already emphasized in the results, levels of agreement obtained in this investigation for Cohen Kappa index can be considered remarkable in some cases and excellent in others, if we stick to the prescription by Landis and Koch (1977) to the K index.

For the degree of agreement on observations of students, other studies provide moderately higher coefficients. However, these works refer to a quantitative tool as the TSAP not using the Kendall test but the intraclass correlation (Richard, Godbout, and Gréhaigne 2000; Gréhaigne, Boutiher, and Godbout 1997; Nadeau, Godbout, and Richard 2008). Also, students were older and, unlike the present study, they did receive the proper training of observers for validating a tool (Medina and Delgado 1999). Presumably, a specific training tool results raise the coefficient Kendall.

Therefore, after checking the shortage of formative assessment practices in teaching invasion games in school (López 1999; Lorente-Catalan and Kirk 2014; Otero 2013), TALIS appears as a contribution that can be incorporated in the context of school to be a simple tool in its application and with a remarkable reliability. This statistic makes it useful, in addition to its educational value, because of the possibility of coding performed between pairs of students can be used as assessment.

In this sense, there is evidence that peer assessment outperforms learning of invasion games over traditional teacher-student assessments. Different research credited better learning outcomes in a formative assessment using peer assessment tools between similar pairs to TALIS (Santos 2010; Pérez, Heras, and Hernán 2008; Figueiredo, Lake, and Fernández 2008). Therefore, both the reality of the classroom (greater learning and

helping the rating), for its statistical benefit, TALIS is posited as a possible tool in the teaching and learning of invasion games in PE.

Conclusions

Once the current state of affairs on the formative evaluation of invasion games in the school context has been reviewed, this study has designed a tool of formative assessment in this area, through validation of content. The results subsequently in both interobserver agreement of experts and students as intraobserver reveal a high statistical reliability.

After this first stage of TALIS, the conclusions invite us to continue, in future work, to validate the tool, contrasting it with the students of different ages and who have received specific training in TALIS. Also, its school application should aim to establish that use of this tool generates more learning at both declarative and procedural level students. To do this they must use other experimental designs with both qualitative and quantitative techniques. Finally, in order to facilitate the incorporation of the tool to physical education classes, automation of the tool in digital format would help promote its use. More specifically, the students in codifying software by clicking on the categories while watching the game would increase the benefits of the tool.

References

- Álamo, J. M. (1996). Las conductas motrices en balonmano: Estudio comparativo entre puestos específicos. *Red: Revista De Entrenamiento Deportivo*, 10(1), 31-36.
- Alarcón, F., Cárdenas, D., Miranda, M. T., Ureña, N., & Piñar, M. I. (2010). La metodología de enseñanza de los deportes de equipo. *Revista De Investigación En Educación*, (7), 91-103.
- Anguera, M. T., & Hernández-Mendo, A. (2014). Observational methodology and sports psychology: State of affairs. *Revista De Psicología Del Deporte*, 23(1), 103-109.
- Ardá, A., & Casal, C. (2003). *Fútbol: Metodología de la enseñanza del fútbol* (1st ed.) Barcelona : Paidotribo, 2003.
- Ávila, F. M. (1999). Los desplazamientos sin balón. factores perceptivos y decisionales en los deportes colectivos. *Habilidad Motriz*, (14), 22-32.
- Bengué, L. (2005). *Fundamentos transversales para los deportes de equipo* (1st ed.) INDE Publicaciones.
- Cárdenas, D. (1996). *Análisis de las demandas físicas del juego del baloncesto*. Malaga: I.A.D.
- Chen, W., Hendricks, K., & Zhu, W. (2013). Development and validation of the basketball offensive game performance instrument. *Journal of Teaching in Physical Education*, 32(1), 100-109.
- Collet, C. (2013). The possession game? A comparative analysis of ball retention and team success in european and international football, 2007-2010. *Journal of Sports Sciences*, 31(2), 123-136.
- Dellal, A., Chamari, K., Wong, D. P., Ahmaidi, S., Keller, D., Barros, R., . . . Carling, C. (2011). Comparison of physical and technical performance in european soccer match-play: Fa premier league and la liga. *European Journal of Sport Science*, 11(1), 51-59.
- Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. London: MacMillan.
- Di Salvo, V., Gregson, W., Atkinson, G., Tordoff, P., & Drust, B. (2009). Analysis of high intensity activity in premier league soccer. *International Journal of Sports Medicine*, 30(3), 205-212.
- Esnaola, I., Ruiz de Auza, S., Zulaika, I., Rodríguez, A., & Goñi, A. (2003). Autoconcepto físico y desarrollo personal: Perspectivas de investigación. *Revista De Psicodidáctica*, (15), 7-64.
- Etxezarra, I., Castellano, J., & Usabiaga, O. 2013. Aplicación de diferentes estrategias para el control de calidad del dato de una herramienta observacional en fútbol formación. *Revista Iberoamericana de Psicología del Ejercicio y el Deporte*, 8 (2),301-316.
- Figueiredo, L. M., Lago, C., & Fernández, M. (2008). Análisis del efecto de un modelo de evaluación recíproca sobre el aprendizaje de los deportes de equipo en el contexto escolar. *Motricidad: Revista De Ciencias De La Actividad Física Y Del Deporte*, (21), 99-117.
- Forrest, G. (2013). Systematic assessment of game-centred approach practices – the game-centred approach assessment scaffold. *Physical Education and Sport Pedagogy*, , 1-15. doi:10.1080/17408989.2013.803526
- Fradua, J. L., & Moreno, R. (2001). La iniciación al fútbol en el medio escolar. In F. Ruiz, A. García & A. J. Casimiro (Eds.), (pp. 145-178). Madrid: Gymnos.
- García, L. M., González, S., Gutiérrez, D., & Serra, J. (2013). Development and validation of the game performance evaluation tool (GPET) in soccer. *Revista Euroamericana De Ciencias Del Deporte*, 2(1), 89-99.
- Gréhaigne, J., Boutin, D., & Godbout, P. (1997). Performance assessment in team sports (evaluation de la performance en sports collectifs). *Journal of Teaching in Physical Education*, 16(4), 500-516.
- Gutiérrez, M., Pilsa, C., & Torres, E. (2007). Perfil de la educación física y sus profesores desde el el punto de vista de los alumnos. *Revista Internacional De Ciencias Del Deporte = International Journal of Sport Science*, 3(8), 39-52.
- Habermas, J. (1997). *Teoría de la acción comunicativa: Complementos y estudios previos*. Madrid: Cátedra.

- Hernández, J., Castro, U. S., Cruz, H., Gil, G., Guerra, G., Quiroga, M., . . . Rodríguez, J. P. (2000). *La iniciación a los deportes desde su estructura y dinámica: Aplicación a la educación física escolar y al entrenamiento* (1st ed.) Barcelona : Inde, 2000.
- Hill, G. M., & Cleven, B. (2005). A comparison of students' choices of 9th grade physical education activities by ethnicity. *The High School Journal*, 89(2), 16-23.
- James, A. R., Griffin, L. L., & France, T. (2005). Perceptions of assessment in elementary physical education: A case study. *Physical Educator*, 62(2), 85-95.
- Jorba, J., & Sanmartí, N. (2013). La función pedagógica de la evaluación. In A. Parcerisa (Ed.), *Evaluación como ayuda al aprendizaje* (pp. 21-42). Barcelona: Grao.
- Lamb, P., Lane, K., & Aldous, D. (2013). Enhancing the spaces of reflection: A buddy peer-review process within physical education initial teacher education. *European Physical Education Review*, 19(1), 21-38.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.
- López, V. M. (1999). *Prácticas de evaluación en educación física: Estudio de casos en primaria, secundaria y formación del profesorado. tesis doctoral publicada*. Valladolid: Secretariado de Publicaciones de la Universidad de Valladolid.
- López, V. M. (2004). Evaluación, calificación, credencialismo y formación inicial del profesorado: Efectos y patologías generadas en la enseñanza universitaria. *Revista Interuniversitaria De Formación Del Profesorado*, (51), 221-232.
- López, V. M., Barba, J. J., Monjas, R., Manrique, J. C., Heras, C., González, M., & Gómez, J. M. (2007). Trece años de evaluación compartida en educación física. *Revista Internacional De Ciencias De La Actividad Física Y El Deporte*, 7(26)
- Lorente-Catalán, E., & Kirk, D. (2014). Making the case for democratic assessment practices within a critical pedagogy of physical education teacher education. *European Physical Education Review*, 20(1), 104-119.
- Medina, J., & Delgado, M. A. (1999). Metodología de entrenamiento de observadores para investigaciones sobre educación física y deporte en las que se utilice como método la observación. *Motricidad: Revista De Ciencias De La Actividad Física Y Del Deporte*, 5, 69-86.
- Memmert, D., & Harvey, S. (2010). Identification of non-specific tactical tasks in invasion games. *Physical Education & Sport Pedagogy*, 15(3), 287-305.
- Méndez, A. (2010). *Modelos actuales de iniciación deportiva. unidades didácticas sobre deportes de invasión*. Sevilla: Wanceulen.
- Nadeau, L., Godbout, P., & Richard, J. (2008).
 the validity and reliability of a performance assessment procedure in ice hockey. *Physical Education and Sport Pedagogy.*, 13(1), 65-83.
- Ní Chróinín, D., & Cosgrave, C. (2013). Implementing formative assessment in primary physical education: Teacher perspectives and experiences. *Physical Education and Sport Pedagogy*, 18(2), 219-233.
- 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.
- Otero, F. M., Carmona, J. A., Calvo, A., & Díaz-Mancha, J. A. (2014). Metodología de enseñanza de los deportes de invasión en primaria. *Revista Internacional De Medicina Y Ciencias De La Actividad Física Y Del Deporte*, (53), 69-87.
- Otero, F. M. (2004). Pautas metodológicas para el diseño de un instrumento de evaluación a partir de un normotipo criterial: La lista de control. *Revista De Educación Física*, (94), 21-26.
- Otero, F. M. (2005). Análisis de los elementos determinantes en la aplicación de una tarea evaluativa en educación física: Propuesta para un deporte de cooperación-oposición y participación simultánea. *Tándem*, (18), 59-69.
- Otero, F. M. (2013). Información y cesión de la responsabilidad en la evaluación de los deportes de invasión en primaria. *EmásF: Revista Digital De Educación Física*, 22, 56-71.
- Otero, F. M., Calvo, Á., & González-Jurado, J. A. (2014). Analysis of the assessment of invasion sports in elementary school. *Cultura, Ciencia Y Deporte*, 9(26), 139-153.
- Otero-Saborido, F. M., & Silva, J. (2015). Fiabilidad en la observación en fútbol sala a través de una herramienta de evaluación de los deportes de equipo (team sport assessment procedure). *Revista Iberoamericana De Psicología Del Deporte*, 10(2), 259-266.
- Otero-Saborido, F.M., Calvo, A., and González-Jurado, J.A.(2015). Agreement and reliability in the use of the Team Sport Assessment Procedure in basketball: A case study, *South African Journal for Research in Sport, Physical Education and Recreation*. 37 (2) In press
- Pérez, Á., Heras, C., & Hernán, I. (2008). Evaluación formativa en la educación secundaria obligatoria. su aplicación a una unidad didáctica de deportes colectivos en el marco del estilo actitudinal. *Revista Española De Educación Física Y Deportes*, 9(Julio-Diciembre), 45-66.
- Perrenoud, P. (1991). Pour une approche pragmatique de l'évaluation formative. *Mesure Et Évaluation En Éducation*, 13(4), 49-81.
- Pintor, D., & Cárdenas, D. (2001). La iniciación al baloncesto en el medio escolar. In F. Ruiz, A. García & A. J. Casimiro (Eds.), (pp. 105-144). Madrid: Gymnos.

- Pope, R. P., Coleman, K. J., Gonzalez, E. C., Barron, F., & Heath, E. M. (2002). Validity of a revised system for observing fitness instruction time (SOFIT). *Pediatric Exercise Science*, 14(2), 135-146.
- Pradas, F., Floría, P., González-Jurado, J., Carrasco, L., & V, B. (2012). Desarrollo de una herramienta de observación para el análisis de la modalidad individual del tenis de mesa. *Journal of Sport and Health Research*, 4(3), 255-268.
- Ramírez, A. A. (2009). *El efecto positivo de la transferencia en el aprendizaje de habilidades específicas del balonmano sobre las habilidades del baloncesto durante la clase de educación física*
- Reina, Á. (2012). *Evaluación de factores implicados en el rendimiento del fútbol profesional*
- Reina-Gómez, A., Hernández-Mendo, A., & Fernández-García, J. C. (2010). Multi-facet design for goal scoring in soccer-7. *Quality and Quantity*, 44(5), 1025-1035.
- Richard, J. F., Godbout, P., & Gréhaigne, J. F. (2000).
 students' precision and reliability of team sport performance. , 70, vol. 1, 85-91(2000). *Research Quarterly for Exercise & Sport*, 1(70), 85-91.
- Rivadeneira, M. L., & Sicilia, A. (2004). *La percepción espacio-temporal y la iniciación a los deportes de equipo en primaria: Unidades didácticas para primaria* (1st ed.) INDE Publicaciones.
- Robles, J. (2009). *Tratamiento del deporte dentro del área de educación física durante la etapa de Educación Secundaria Obligatoria en la provincia de Huelva*. Universidad de Huelva. Departamento de Expresión Corporal).
- Sánchez-Mora, D., García, L. M., Del Valle, S., & Solera, I. (2011). Spanish primary school students' knowledge of invasion games. *Physical Education and Sport Pedagogy*, 16(3), 251-264.
- Santos, S. (2010). La utilización de hojas de registro en la enseñanza de los deportes colectivos. *Tándem: Didáctica De La Educación Física*, (34), 91-108.
- Santos, M. A. (2002). Una flecha en la diana: La evaluación como aprendizaje. *Aula De Innovación Educativa*, 34, 7-9.
- Serra, J., González, S., & García, L. M. (2011). Comparación del rendimiento de juego de jugadores de fútbol de 8-9 años en dos juegos modificados 3 contra 3. *Cuadernos De Psicología Del Deporte*, 11(2), 77-92.
- Shropshire, J., & Carroll, B. (1998). Final year primary school children's physical activity levels and choices. *European Journal of Physical Education*, 3(2), 156-166.
- Sicilia, A., Delgado, M. A., Sáenz-López, P., Manzano, J. I., Varela, R., Cañadas, J., & Félix, M. (2006). La evaluación de aprendizajes en educación física. diferencias en función del nivel educativo. *Motricidad: Revista De Ciencias De La Actividad Física Y Del Deporte*, (17), 71-95.
- Sicilia, A. (2003). La investigación sobre el pensamiento del alumnado. una revisión desde la educación física. *Revista De Educación*, (331), 577-613.
- Tejada, J. (2007). *La evaluación en educación física en huelva y su provincia*
- Vegas, G. (2007). *Metodología de enseñanza basada en la implicación cognitiva del jugador de fútbol base*. (Unpublished Departamento de Educación Físico Deportiva. Universidad de Granada, Granada.
- Vera, J. A., González, M., & Moreno, J. A. (2009). Relaciones entre la cesión de responsabilidad en la evaluación y la percepción de igualdad en la enseñanza de la educación física escolar. / relationships between the transfer of responsibility in assessment and the perception of equality in physical education classes. *Cultura, Ciencia Y Deporte*, 4(10), 25-31.
- Wein, H. (1995). Fútbol a la medida del niño. un óptimo modelo de formación como clave de futuros éxitos. *CEDIF. RFEF. Madrid*,
- Yáñez, J., & Castejón, F. J. (2011). La utilización de la transferencia para el aprendizaje de la táctica colectiva deportiva en educación secundaria. *Infancia Y Aprendizaje*, 34(1), 95-107.