Muñoz, J.; Gamonales, J.M.; León, K. e Ibáñez, S.J. (2018) Formación de codificadores y fiabilidad de los registros. Una aplicación al goalball / Training of Coders and Reliability. An Application to the Goalball. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 18 (72) pp. 669-691
Http://cdeporte.rediris.es/revista/revista72/artformacion976.htm
DOI: http://doi.org/10.15366/rimcafd2018.72.005

CODER TRAINING AND DATA RELIABILITY. AN APPLICATION TO GOALBALL

FORMACIÓN DE CODIFICADORES Y FIABILIDAD DE LOS REGISTROS. UNA APLICACIÓN AL GOALBALL

Muñoz, J.[1,2]; Gamonales, J.M.[1]; León, K.[1,2] e Ibáñez, S.J.[1]

[1] Facultad de Ciencias de la Actividad Física y el Deporte. Universidad de Extremadura (Spain) suliwan@unex.es, martingamonales@unex.es, fleon@unex.es, sibanez@unex.es
[2] Universidad Autónoma de Chile (Chile) jesus.munoz@uautonoma.cl, francisco.leon@uautonoma.cl

Spanish-English translator: Rocío Domínguez Castells, rocio@sport-science.net

ACKNOWLEDGEMENTS AND/OR FUNDING

This work was developed by the Research Group for the Optimization of Training and Sport Performance (Grupo de Optimización del Entrenamiento y Rendimiento Deportivo, G.O.E.R.D.) of the Faculty of Sport Sciences of the University of Extremadura.

This work has been supported by the Funding for Research Groups (GR15122) of the Government of Extremadura (Employment and Infrastructure Office–Consejería de Empleo e Infraestructuras), with the contribution of the European Union through the European Regional Development Fund (ERDF).

UNIÓN EUROPEA
Fondo Europeo de Desarrollo Regional
Una manera de hacer Europa
JUNTA DE EXTREMADURA

Código UNESCO / UNESCO code: 5899 Educación Física y Deportiva / Physical Education and Sports
Clasificación Consejo de Europa / Council of Europe classification: 5. Didáctica y metodología / Didactics and methodology 17. Otras: Procesos de percepción / Others: Perception processes

Recibido 13 de enero de 2017 **Received** January 13, 2017

Aceptado 11 de septiembre de 2017 **Accepted** September 11, 2017

ABSTRACT

The purpose of this paper is to describe the training process and inter-rather reliability evaluation of five coders involved in a research about performance analysis of sport in goalball. The coders participated in a training process, consisted of six face-to-face sessions and four individual exercises, distributed in three phases; Theoretical, Practical and training. After the first two phases, reliability was measured during the development of training period using Multirather Kappa Free. Finally, an almost perfect level of inter-rather reliability (>.86) was obtaining in all variables defined for the research, allowing codificate all goalball match actions that formed the sample.

KEY WORDS: *goalball, observational methodology, reliability, coder training, team sports.*

RESUMEN

El propósito de este artículo es describir el proceso de formación y evaluación de la fiabilidad inter-observador de los cinco codificadores participantes en un estudio sobre el análisis del rendimiento competitivo en goalball de alta competición. Los codificadores participaron en un proceso de formación consistente en seis sesiones presenciales en grupo y cuatro ejercicios prácticos individuales y no presenciales, distribuidos en tres fases; Teórica, Práctica y de Entrenamiento. Tras las dos primeras se procedió al cálculo de la fiabilidad durante el desarrollo de la fase de entrenamiento a través de la prueba *Multirather Kappa Free*, obteniendo finalmente un nivel de fiabilidad inter-observador casi perfecto (>0,86) en todas las variables definidas para el estudio, lo que permitió la codificación de las acciones de los partidos registrados de Goalball que conformaron la muestra del mismo.

PALABAS CLAVES: goalball, metodología observacional, fiabilidad, formación de observadores, deportes de equipo.

INTRODUCTION

Sport Performance Analysis aims to gain knowledge on match context in order to improve future results (McGarry, 2009) during training or competition. All research studies that analyse sport performance are included in Sport Performance Analysis (Hughes & Bartlett, 2002). More and more researchers use this analysis because it makes decision making easier for coaches by providing an objective view of the game in children (Ureña, Morales-Rojas, León & González, 2014), amateur (García, Ibáñez, Parejo, Feu & Cañada, 2011; Ortega, Fernández, Ubal, Lorenzo & Sampaio, 2010) or professional competition (Blanco, Ibáñez, Antúnez & Hernández-Mendo, 2014; Ohnjec, Vuleta, Milanović, & Gruić, 2008; Volossovitch, Dumangane, & Rosati, 2012). This kind of studies allows for the observation of behaviours that affect sport performance with good quality, reliability, validity and accuracy (Salas & Hernández-Mendo, 2016), or for the synchronic analysis during a particular study phase (Hileno & Buscá, 2012). Therefore, they share one of the most relevant aspects of sport performance analysis: to show the study's ability to obtain significant results regarding the research context through accurate recording of previously defined variables.

The observers-coders participate in this process, being responsible for identifying all the actions happening in each context analysed. This procedure has been applied to different research fields, not only sport. González-Díaz and Iglesias-García (2015) described the coders' training and working process applied to a research project on food advertising. Within the early intervention field, Trenado, Pons-Salvador and Cerezo (2014) studied the reliability of the CITMI-R in its English version (early mother-child interaction coding system), including observer training.

The *Notational Analysis* that applies to the study of *sport performance indicators* is based on observation. In studies that use observational methodology and, therefore, depend on subjective interpretations, it is essential to ensure the highest validity and reliability possible of the recordings, with the aim to guarantee the objectivity of the collected data. By doing so, research achieves its purposes effectively and accurately, as shown by numerous studies on match analysis in sport with confirmed validity (Anguera, 1991). This is also proved by the constant and significant growth of the use of observational methodology in the last decade in the sport field (Anguera & Hernández-Mendo, 2013, 2014). Therefore, this methodology seems appropriate to solve the problems that arise in studies on physical education and sport (Blanco et al., 2014). Within this context, the role of the observer becomes essential (Arias, Argudo & Alonso, 2009a, 2009b).

The observer may not be considered as a tool, but as an individual who has been trained to evaluate perceivable behaviours and who has an active role in the observation process (Arias et al., 2009a; Anoz, García & García, 2004; Piñar, 2005). Anguera, Blanco-Villaseñor, Losada and Hernández-Mendo (2000) exposed the need of observer training, for which several proposals exist (Medina & Delgado, 1999; Losada & Manolov, 2014). The observer training must be understood as a process through which to acquire conceptual,

empirical and technological maturity, which enables them to conduct the observation (Anguera, 2003). After theoretical training, the observers undergo a practical evaluation using agreement or association coefficients. According to O´Donoghue (2007), a large number of performance analysis methods use non-automatic techniques for data collection, what may yield errors that limit the reliability of the results.

It is important to develop a rigorous coder-training program, with the aim to determine the reliability and objectivity of the process (Brewer & Jones, 2002). This validity and reliability should guarantee that the collected data constitute a true reflection of performance in the analysed context. To do so, it is essential to choose the necessary statistic tools to conduct the analysis, in order to assess reliability (Hughes, Cooper & Nevil, 2002). Furthermore, reliability should reflect the way in which notational data are analysed, so that assessments can be made regarding the variables and the coded results can be presented accurately (James, Taylor & Stanley, 2007). In this regard, several studies have adapted Medina and Delgado's (1999) proposal of coder training to specific sport contexts, such as waterpolo (García, Argudo & Alonso, 2007), basketball (Arias et al., 2009a) and volleyball (Moreno et al., 2002); to teaching environments (Viciana, 1999); or they have studied observer reliability according to their experience and training (Denis, Lortie & Bruxelles, 2002). Likewise, some studies mention the coder training process but they do not provide details about the procedure used (Chillón & Delgado, 2012; Ortega et al., 2010; Salas & Hernández-Mendo, 2016; Ureña et al., 2014).

Notational Analysis within Sport Sciences has proven to be a useful tool to improve performance, as long as it is conducted correctly, for it allows for recording and analysis of complex dynamic situations (Hughes & Franks, 2005). Its applications include, but are not limited to studying general movements or movement patterns in team sports, mainly related to tactics and strategy (Bartlett, 2001). Due to this crucial aspect, research must follow certain phases, starting with the correct definition of the variables of the sport context to be observed. Subsequently, according to Anguera (2003), the process must follow two phases. The first one is the coder training, which enables them to learn about the process and its fundamentals. The specific training during the second phase focuses on gaining knowledge on the fundamentals of the observation process. After these two phases, it is necessary to measure objectively whether high inter-coder agreement has been achieved. Statistical tests must be used to calculate the validity and intraobserver and interobserver reliability in regard to the observed behaviours, defined by the variables and their respective categories.

Observational methodology has been used to gain knowledge on certain sport areas, such as goalball: a team sport for the visually impaired, which is based on hearing and touch (Gulick & Malone, 2011). After a thorough review of scientific literature on the use of observational methodology in sport performance analysis in goalball, it is concluded that, compared with other modalities, the studies are scarce but recent. This indicates that this is an emerging field of knowledge that is beginning to generate scientific interest. Abdolmaleki, Mirzazadeh, Allahyari and Ramezani (2015) used their own

questionnaire to describe the factors that affect performance, which are related to technique, the team, the players' experience and the coaches' experience. Other studies focused on the analysis of defensive and offensive systems in goalball (Amorím, Botelho, Sampaio, Molina & Corredeira, 2010; Tosim, Junior, Leitão & Simões, 2008a; Tosim, Massolli & Beltrao, 2008b). On the other hand, Owen (2014) identified performance indicators applying a regression analysis based on the one proposed by O'Donoghue and Cullinane (2011), but specifically adapted to goalball. Molik et al. (2015) described performance in goalball related to the anthropometric characteristics of elite players by means of a Game Efficiency Sheet, while Morato, Da Cunha, Gamero, Magalhães and Almeida (2016) developed and tested an observation system for goalball analysis using the systematic observational methodology with expert observers. Weber and Link (2016) developed software for performance analysis in goalball (GoalScout, GoalView y GoalTracK).

Nonetheless, the coder training process and data reliability assessment are not described with detail in most of these studies. Therefore, the aim of the present research is to describe the coder training process used in a study on performance indicators in goalball, improving the procedures existing in the literature. This general aim may be divided into two specific aims: i) to design a coder training process; and ii) to analyse the observer reliability.

METHODS

Design

The research design belongs to the *instrumental studies* (Montero & León, 2007), where the coders are involved in the variable and category definition during the training process. Therefore, this work includes both the design and the analysis of the data properties (Servera & Cardo, 2006). It aims for an observer training methodology that ensures data collection reliability and, therefore, whose application may be useful in different contexts and research areas related to Physical Activity and Sport Science and, more specifically, to goalball.

Participants

Five expert observers were non-randomly selected for the present study (Anguera, 2003; Rodríguez, Gil & García, 1996). They were able to transmit knowledge about the subject under study and to provide feedback that may lead to reflection in order to help the researcher (Escobar & Cuervo, 2008). They were also accessible (Valles, 2003). The selected observers should meet the five following inclusion criteria:

First criterion. To show interest in the study.

Second criterion. To hold a university degree related to Physical Activity and Sport (Master of Science/Grade in Physical Activity and Sport Sciences).

Third criterion. To have received specific training and to have passed the courses on physical activity for the disabled within the university degree.

Fourth criterion. To have practised goalball in an inclusive or educational environment.

Fifth criterion. To commit to observing, recording and coding the assigned matches.

These criteria tried to ensure that the coders possessed previous knowledge on adapted sport and physical activity, as well as on goalball and its characteristics. The five selected participants met the established criteria.

Variables

The variables to be observed by the coders, as well as their categories (also called *Categorical Nuclei*) and their respective *Degree of Plasticity* (Anguera, 1991), followed the procedure proposed by Anguera and Hernández-Mendo (2013). This means they were chosen after a previous study, in agreement with a committee of experts (national and international level coaches and university professors, experts in sport and disability), who tried to define precisely the determining motor actions to be observed.

The variables observed by the coders were divided into groups. The first one comprised six variables related to offensive technical-tactical actions: throwing player, starting area, ending area, throwing technique, type of throw and throw outcome. The second group contained six variables related to defensive technical-tactical actions: defending system, defending player, defending technique, type of defence, area where the action occurs and outcome of the defensive action. The variables were used to analyse the performance indicators (O´Donoghue, 2010). They were all numerically coded, with the aim to make their recording and statistical analysis easier. Moreover, this made it easier for the coders to observe and record them during the training process.

Material

The following material was used during the coder training process, the match observation, and the data recording, codification and analysis: a virtual platform within Moodle environment, an introductory manual (containing the game violations and the variables to be observed in the study), videos of goalball matches, reliability analysis software (Randolph, 2005), and a cloud storage service.

Procedure

In order to design a coder training process to assess goalball offensive and defensive technical-tactical actions, it was necessary to follow a methodical procedure, taking many factors into account, such as the literature review, procedure design, sample selection, documentation delivery, data collection

and analysis, and final instrument preparation (Gamonales, León, Muñoz, González-Espinosa & Ibáñez, 2017).

Therefore, bearing the above and Medina and Delgado's (1999) approach in mind, the coders were involved during their training process in recoding the categories to be observed, with the aim to improve constructively some of the proposals regarding coder training existing in the scientific literature (Arias et al., 2009a; Moreno et al., 2002). Likewise, the training phases suggested by Anguera (2003) were taken into account to design the coder training process. This process was divided into two moments: the *coder selection* and the *coder training process*, including the inter-coder reliability assessment. Figure 1 shows the phases followed by the researchers and coders during the training process.

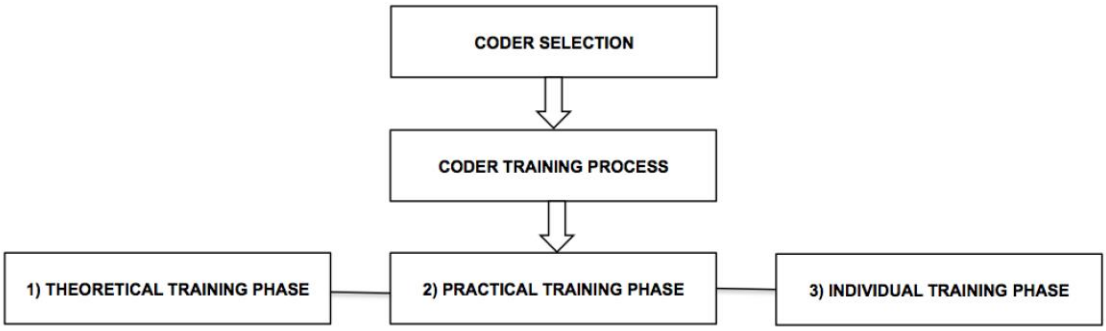

Figure 1. Diagram of the research phases.

The variables and categories were previously defined and revised by the experts, what simplified the process by suppressing the division into two phases proposed by these authors (preparation–with two subphases–and training). Therefore, the coder training process applied in the present study consisted of three phases, implemented in Moodle virtual platform:

1) Theoretical training phase, to gain knowledge on the sport context in general and on the study variables in particular. Instruments that bring the coders closer to the context to be observed were expressly designed for this phase. Different kinds of documents were produced in several formats: video, screencast, match fragments and presentations. Table 1 lists the instruments produced and used during the Theoretical training phase.

Table 1. Instruments produced and used during the Theoretical training phase.

Instrument	Description
Presentations	Document that provides, in a brief, schematic and visual manner, all the information regarding the analysed sport context, in this case goalball.
Video tutorial - Screencast	Video file containing the image and sound recording of what the user does and says on the computer. Here it was used to explain the main characteristics of goalball while a match fragment was viewed.
Video fragments	Goalball match fragments from the championship under study, randomly selected, that were used for the training process and the reliability assessment.
PRES01 – 1st video tutorial	General explanation of the game: video tutorial - screencast to explain the main characteristics of goalball.
PRES02 - Violations	Presentation in ".pdf" format about the main game violations contained in the oficial rules, with images that help with their comprehension.
PRE03 – The study	Presentation in ".pdf" format regarding the main aims and characteristics of the study, with the purpose to engage the coders.
PRE04 - Context variables	Presentation in ".pdf" format containing the definition of the study context variables.
PRE05 – Offensive variables	Presentation in ".pdf" format regarding the offensive variables.
PRE06 - Defensive variables	Presentation in ".pdf" format regarding the defensive variables.
PRE07 - Outcome	Presentation in ".pdf" format regarding the action final outcome.

This phase was completely face-to-face, with training days divided into six sessions of 40 to 60 minutes, depending on the content. Table 2 contains the content and material distribution of the face-to-face training sessions.

Table 2. Timing of the practical face-to-face sessions.

Session	Name	Duration	Activities	Remarks / material
1	Initial requirements	10´	0.1- Installing SPSS 0.2- Installing and using DROPBOX	Each coder with his/her PC
	Description of the sport characteristics	30´	1.1- Video tutorial	PRE01
			1.2- Presentation on violations	PRE02
2	Introduction of the study, the variables and the data collection instrument	40´	2.1- Presentation about the study	PRE03
			2.2- SPSS: file for data collection	goalball_def.sav
3	Context variables	40´	3.1- Presentation on *context variables* 3.2- Filling in the context variables: - Indexer's data - File name - Match schedule and number - How to fill in the CV	PRE04 goalball_def.sav
4	Offensive variables	60´	4.1- Presentation on offensive variables 4.2- Practical exercise: 10 actions of each offensive variable 4.3- 20´ of comments and remarks related to the exercise	PRE05 Simplified ".pdf" diagram Paper and pen
5	Defensive variables	60´	5.1- Presentation on defen*sive variables*. 5.2- Practical exercise: 10 actions of each defensive variable 5.3- 15´ of comments and remarks about the exercise	PRE06 Simplified ".pdf" diagram Spreadsheet
6	Outcome variable	60´	6.1- Presentation on the *variable outcome*. 6.2- Practical exercise: 20 actions of each outcome variable 6.3- 20´ of comments and remarks about the exercise	PRE06 Simplified ".pdf" diagram

A computer with a large format screen was used during the training day and the coders were distributed in front of it. Every session followed the same structure: the main researcher explained the content with the help of the presentations and images from the championship. During the explanation, the coders asked questions to clarify the concepts. Afterwards, practical coding exercises of real situations related to the four variable groups (context, offensive, defensive and outcome) were done. These groups were introduced in successive sessions.

The exercises consisted in viewing and coding individually blocks of ten actions displayed by the main researcher. Lastly, every session ended with the collective correction and comments on each exercise, emphasizing confusing actions.

2) Practical training phase. Its aim was to analyse the images and to achieve high inter-coder reliability in the recorded observations. The exercise structure in both the practical and the individual training phases was similar and included the same resources. Table 3 displays the permanent resources available in Moodle virtual platform to exchange information and save data.

Table 3. Permanent resources available in Moodle virtual platform

Resource	Description
Virtual forum	To share and solve doubts or questions the exercises.
Handing-in virtual space	To upload every exercise once finished.
Report with the results and recommendations regarding every variable	To continue with the training process after the coders' exercises pass the reliability tests

3) Individual training phase, in which reliability was assessed and the necessary corrections and adjustments were progressively made with three aims: to provide immediate access to the training material, to share through virtual forums the questions that may arise and to let the coders do and save the exercises during both phases. Lastly, there were two types of training sessions: face-to-face group sessions and remote individual sessions (Table 4).

Table 4. Types of training sessions

Type	Description
Face-to-face group sessions	These sessions were imparted by the study's main researcher and were mainly oriented to theoretical training.
Remote individual sessions	Individual working sessions oriented to observation practice. This type of session constituted an advantage, since the training could be adapted to each coder's personal needs.

Statistical analysis

The reliability assessment of the data collected by the coders was conducted using *Multirater Kappa free* (MKF) statistic, a variation of *Kappa* coefficient (Cohen, 1960). This tool was especially designed to test reliability (Randolph, 2005) in cases where there are more than two coders who must observe variables with more than two categories and where it is not mandatory for the observers to assign a number of cases to each category (Brennan & Prediger, 1981).

Kappa values may vary between -1.00 and 1.00 and the acceptable inter-coder agreement *Kappa* coefficient varies depending on the authors (James & et al., 2007). Landis and Koch (1977) established a nomenclature to describe the relative strength of agreement with *Kappa* Statistic (Table 5).

Table 5. Lables for *Kappa* statistic associated to the strength of inter-observer agreement (Landis & Koch, 1977)

Kappa statistic	Strength of agreement	Key
<0.00	Poor	P
0.00 – 0.20	Slight	SL
0.21 – 0.40	Fair	F
0.41 – 0.60	Moderate	M
0.61 – 0.80	Substantial	S
0.81 – 1.00	Almost perfect	AP

The labels proposed by Landis and Koch (1977) were used to describe the strength of inter-observer agreement during the training period. On the other hand, Randolph (2005) stated that an MKF value of 0.70 or higher indicates high strength of inter-coder agreement. This value of inter-coder reliability was established as the minimum necessary for the present study. Lastly, the average reliability of the data collected by the coders regarding the defensive and offensive variables was calculated, as well as for the whole set of variables. This enabled us to verify whether the coder training process under evaluation was appropriate.

RESULTS

Table 6 shows the evolution of the strength of agreement obtained in the exercises of the individual training phase for the variables regarding offensive and defensive technical-tactical actions. The number of actions to be analysed increased progressively along this phase until reaching a complete match, similar to the observation units.

Table 6. Evolution of the strength of agreement in the individual training exercises.

Variable	Vid04 20 Thr.	SA	Vid05 30 Thr.	SA	Vid06 40 Thr.	SA	Vid07 Match half	SA
V1O. Throwing player.	0.97	AP	0.90	AP	0.94	AP	0.92	AP
V2O. Starting area.	0.84	AP	0.83	AP	0.87	AP	0.87	AP
V3O. Ending area.	0.68	S	0.72	S	0.73	S	0.85	AP
V4O. Throwing technique.	0.93	AP	0.96	AP	0.92	AP	0.96	AP
V5O. Type of throw.	0.73	S	0.87	AP	0.73	S	0.76	S
V6O. Throw outcome.	0.92	AP	1.00	AP	0.92	AP	1.00	AP
V1D. Defending system.	0.53	M	0.85	AP	0.85	AP	0.92	AP
V2D. Defending player.	0.85	AP	0.92	AP	0.93	AP	0.93	AP
V3D. Defending technique.	0.72	S	0.61	S	0.65	S	0.61	S
V3D*. Defending technique - recoded.	0.91	AP	0.90	AP	0.90	AP	0.90	AP
V4D. Type of defence.	0.74	S	0.74	S	0.83	AP	0.74	S
V5D. Area where the action occurs.	0.74	S	0.85	AP	0.82	AP	0.85	AP
V6D. Outcome	0.83	AP	0.88	AP	0.90	AP	0.90	AP

SA: Strength of agreement labels according to Landis and Koch (1977)
Thr.: Throws

Most of the variables showed high strength of agreement from the beginning. In others, the results improved noticeably along the individual training phase and at the end they largely exceeded 0.70, the value established by Randolph (2005) as significant. Only for the variable *Defending technique* the value stayed below 0.70. This indicates low strength of inter-observer agreement, according to the reference value established in this study, despite being "substantial" according to Landis and Koch (1977). With the purpose to achieve higher strength of agreement, and taking the source of inter-observer errors described by James et al. (2007) and Anguera (1988) into account, a recodification and/or new definition process was applied to the affected variables, as shown in Figure 2.

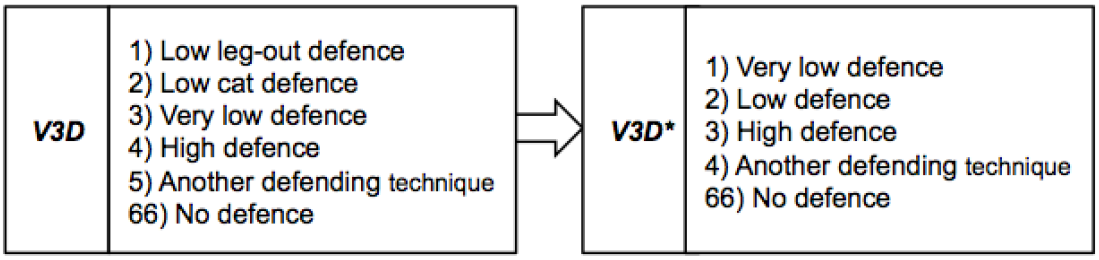

Figure 2. Variable recodification diagram

Table 7 contains the evolution of the reliability of the data collected by the coders during the individual training exercises.

Table 7. Average of the data collected by the coders.

	Vid04 20 Thr.	SA	Vid05 30 Thr.	SA	Vid06 40 Thr.	SA	Vid07 Match half	SA
Offensive Variables	0.84	AP	0.88	AP	0.85	AP	0.89	AP
Defensive Variables	0.76	S	0.82	AP	0.84	AP	0.83	AP
Total	0.79	S	0.84	AP	0.84	AP	0.85	AP

SA: Strength of agreement labels according to Landis and Koch (1977)
Thr.: Throws

A partial and total improvement in the inter-coder agreement was observed along the training process.

DISCUSSION

Several studies have adapted Medina and Delgado's (1999) coder training proposal and have described the different process phases with detail (Arias et al., 2009a; García et al., 2007; Viciana, 1999; Denis et al., 2002; Moreno et al., 2002). However, other studies mention the coder training process but do not describe clearly the procedures followed to design the sessions, to determine the number of coders or to define frequent errors (Chillón & Delgado, 2012; Ortega et al., 2010; Salas & Hernández-Mendo, 2016; Ureña et al., 2014). On the other hand, the studies found in the literature concerning goalball do not provide detailed information about the training process (Amorím et al., 2010; Morato et al., 2016) or the inter-observer reliability assessment (Abdolmaleki et al., 2015; Molik et al., 2015; Oliveira & De Martino, 2012; Owen, 2014; Tosim et al., 2008a; Weber & Link, 2016). Some authors justify this lack of precision due to the process complexity, arguing that it is impossible to explain the training process or the reliability tests used (O´Donoghue, 2010). This entails a significant lack of information on the coder training process (Arias et al., 2009a), what, according to Anguera (1998), may constitute a *source of error* when applying the Observational Methodology.

The aim of this research was to analyse the reliability of a group of participants who underwent a coder training process. The coder training applied in the present study met Medina and Delgado's (1999) and Anguera's (2003) guidelines, and included the following phases: Theoretical training, Practical training and Individual training.

1) Theoretical training phase

This first phase concerned the coder training, assuming their occasional participation in variable and category redefinition. Determining the number of sessions included in the process becomes especially important in this phase. Some authors from the Physical Activity and Sport field suggested a face-to-face coder training process with three sessions of approximately ninety minutes each (Moreno et al., 2002), seven sessions with different durations, ranging between one and a half hour and three hours (Arias et al., 2009a), twelve sessions (Blanco et al., 2014) or fifteen sessions of unknown duration (Medina, 1996). However, Viciana (1999) stated that the number of sessions would

depend on the contents to be imparted. He included twelve sessions in his study, four in each phase. Lastly, Usuabiaga, Castellano, Blanco-Villaseñor and Casamichana (2013) implemented a four-week (thirty-hour) training programme based on a designed observation protocol. In other research fields, the training proposal comprised five sessions of four hours each (Trenado et al., 2014). This reveals the large variability in the number and duration of the sessions.

Nonetheless, the number of sessions included in this study (six) was in accordance with the study content. Six sessions seemed enough to achieve the aim established by Medina and Delgado (1999) for the coder training phase: to let the observers familiarise with the behaviour to be observed, to involve them in the category design process and to show them the kind of task they will have to perform during the observation process. Furthermore, the sessions were designed to be as productive, participatory and clarifying as possible. They were limited to a maximum duration of sixty minutes, in order to avoid fatigue that would lead to a decrease in the coders' attention and concentration.

On the other hand, the number of coders involved in the training process varies depending on the autor. The selection of the group of coders is crucial: they must be the most suitable for the study to make the collection process successful. Medina (1996) proposed eight as an appropriate number of observers to be trained and to reach high enough inter-coder reliability. Viciana (1999) involved fifteen students of the Master of Science in Physical Education (divided into three groups of five) as observers. Moreno et al. (2002) selected ten coders without specifying the requirements to participate in the study. Arias et al. (2009a) recruited four novel participants from different academic fields and one expert, while Hernández-Mendo, Montoro, Reina and Fernández (2012) selected six participants, divided into two groups, to conduct non-planned, systematic observations. Eight pairs of observers participated in the study conducted by Usuabiaga et al. (2013). Two of them were excluded because they did not complete the training plan. Lastly, Casal, Losada and Ardá (2015) involved four coders. Other studies do not specify the number of coders, although they do specify that the coders followed a training program (Jiménez & Hernández-Mendo 2016; Morillo & Hernández-Mendo, 2015).

Five coders were recruited for the present study. This number seemed enough to deal with possible withdrawals during the process and to ensure a correct data collection. All the participants were able to transmit knowledge and information about the subject under study. The selection criteria showed similarities with those from other studies, such as Medina (1996): Physical Activity and Sport professionals, with enough previous knowledge on the subject under study, were selected. The observers involved in the present study had higher educational level than the participants of Viciana's (1999) study, since they held an academic certification of a specific competence level. On the other hand, as requested in Arias et al.'s (2009a) study, they were required to show interest and commitment to complete the observation, recording and codification of the assigned matches.

Therefore, the aim of this phase was to explain the coders the observation plan, minimising questions or doubts about the process. For this reason, it was

important to select coders who were responsible and commited with the study. In doing so, errors (Anguera et al., 2000) and withdrawal along the process would be avoided. It was also important to provide the observers with the necessary videos and files after every session, so that they could be ready for the rest of the sessions of the first phase. Hence, the observation process would receive continuous feedback and would constitute a convenient method to represent what happens in the observed context as faithfully as possible (Pieron, 1986).

2) Practical training phase

This phase had to consist of practical activities that would help the observers learn how to conduct the recording during the observation process (Medina & Delgado, 1999). The practical phase of the coder training process was designed aiming for simplicity and based on individual exercises to analyse and code a specific number of throws of the championship under study. *Context* variables were not included in the analysis. The number of sessions increased progressively along the training sessions, as suggested by Medina (1996).

The assessment of the strength of inter-coder agreement started with this exercise and was later applied to determine reliability and validity. Most of the variables involved in the first exercise of the practical phase (Vid04), which included twenty throws, showed high strength of agreement (>0.70) according to Randolph (2005), except the variables V30 and V1D. This reveals that the coders had doubts and/or made errors. According to Anguera (1988), the *sources of errors* associated to the observers can be prevented by means of good planning and training. Therefore, after the completion of the exercise, a report with the results and some recommendations was provided.

This process was also included in the individual training phase. These reports aimed to identify and to solve errors in the variables without lack of agreement, as well as to improve category comprehension by focusing on correcting the situation and preventing future errors. According to Viciana (1999), each phase must develop depending on the training strategy applied. Hence, the coders were reminded the main aspects to be born in mind before the beginning of every exercise. By doing so, potential errors in codification were prevented. Furthermore, it is important to note that the technical means used in the coder training process and the reliability assessment must support the observers' work (Anguera, 1988). For this reason, Moodle platform was used in this phase. It enabled us to provide all the necessary information and resources for the correct development of the process, as well as to save the coders' work. This was all based on the assumption that appropriate coder training provides research with higher objectivity and credibility (Medina, 1996).

3) Individual training phase

The observation exercises were performed individually again in this phase, and individual reliability was assessed while making corrections and adjustments by means of the result and recommendation report. The number of offensive and

defensive actions to be observed increased progressively up to 30 in “Vid05”, 40 in “Vid 06” and all the actions contained in half a match in “Vid07”. As in the previous phases, the coders uploaded a file with the observation results of every exercise to Moodle, where they could also make questions in the forums available for this purpose. This kind of support is very useful to organise and manage the information provided to the coders and constitutes an efficient and accessible update of the tools used in the coder training process.

Most of the variables coded in the exercises of this individual training phase showed high strength of agreement from the beginning. The results of other variables improved considerably along the individual training phase until largely exceeding the established value, 0.70 (Randolph, 2005). Only the variable *Defending technique* did not reach this value after the whole process, which indicated insufficient strength of inter-observer agreement compared to the reference value established in this study. In this case, it was necessary to look for the source of disagreement. James et al. (2007) suggested three sources of error that may lead to lack of agreement between observers and, therefore, low inter-observer reliability: *Operational error*, *Observational error* and *Definition error*. After analysing all observers’ notes, it was concluded that the error in the case of the variable *Defending technique* was a *Definition error*. Two of this variable’s categories were very similar and, hence, difficult to distinguish in the image due to the camera filming angle. It was then decided to recode the variable, which consists in merging several categories into one. After testing several category combinations, the two which led to error were detected and merged into one category. Subsequently, the observations were again tested for reliability, yielding the results showed in Table 6. A noticeable increase in reliability was observed, reaching very high agreement values. As a result, the variable *Defending technique* was recoded from six categories to five, joining the first and second categories. The categories were also assigned a new, more logical, order. Once the inter-observer reliability is high enough for every variable, the sample collection may begin.

The process described in the present document had the purpose to ensure a high objectivity of the study sample, necessary for research. The coder training, as well as the reliability assessment, was described. The latter is essential in studies that apply the Observational Methogology (Anguera, 1988, 1991). As it can be noticed, the strength of inter-coder agreement in this study was in accordance with the minimum reliability established. The coder training process in goalball yielded a value of 0.86, higher than in other research studies (Den Hollander, Brown, Lambert, Treu & Hendricks, 2016; Cañadas, Ibáñez & Leite, 2015; Ibáñez, Santos & García, 2015) and slightly lower than the one obtained by García, Ibáñez, Gómez and Sampaio (2009), 0.88. Thus, in light of the results, it can be stated that the procedure’s reliability was “almost perfect” (Landis & Koch, 1977) and its results were similar to previous research.

CONCLUSION

The purpose of the coder training process was to achieve high inter-observer agreement for every variable to be analysed. This training process was divided

into three phases: theoretical training phase, practical training phase and individual training phase.

The session design and development is crucial to be successful and must meet the original goals and expectations. Therefore, the individual training phase must last until the inter-observer reliability is high enough to ensure data objectivity. In order to improve the inter-observer reliability, the sources of error were searched, the errors were analysed and the coders were provided with feedback.

The coder training process presented in this paper may be applied in the Physical Activity and Sport Sciences field and, particularly, in goalball.

The scarcity of studies that describe the coder training process in detail constitutes a limitation of the present study, so it is necessary to conduct specific literature reviews on the coder training process in the future.

REFERENCES

Abdolmaleki, H., Mirzazadeh, Z., Allahyari, M., & Ramezani, M. (2015). Identify and analysis of performance evaluation indicators of Iranian Goalball coaches. *Annals of Applied Sport Science, 3*(3), 43-56.

Altman, D. G. (1991). *Practical Satistics for Medical Research*. London: Chapman y Hall.

Amorím, M. T., Botelho, M., Sampaio, E., Molina, J., & Corredeira, R. (2010). Caracterización de los patrones comportamentales de los atletas con discapacidad visual practicantes de Goalball. *Revista Electrónica Interuniversitaria de Formación del Profesorado, 13*(3), 47-57.

Anguera, M. T. (1988). *Observación en la escuela*. Barcelona: Graó.

Anguera, M. T. (1991). *Metodología observacional en la investigación psicológica*. Barcelona: PPU.

Anguera, M. T. (2003). La observación. En C. Moreno Rosset (ed.), *Evaluación psicológica. Concepto, proceso y aplicación en las áreas del desarrollo y de la inteligencia* (pp. 271-308). Madrid: Sanz y Torres.

Anguera, M. T., & Hernández-Mendo, A. (2013). Metodología observacional en el ámbito del deporte. *E-balonmano.com: Revista de Ciencias del Deporte, 9*(3), 135-160.

Anguera, M. T., & Hernández-Mendo, A. (2014). Metodología observacional y psicología del deporte: Estado de la cuestión. *Revista de Psicología del Deporte, 23*(1), 103-109.

Anguera, M. T., Blanco-Villaseñor, A., Losada, J. L., & Hernández-Mendo, A. (2000). La metodología observacional en el deporte: Conceptos básicos. *Lecturas: E.F. y Deportes, Revista Digital, 24.*

Anoz, I., García, T., & García, J. A. (2004) Análisis de la estructura del ataque en equipos de alto nivel de balonmano. *Apunts, Educación Física y Deportes* (76), 53-58.

Arias, J. L., Argudo, F. M., & Alonso, J. I. (2009a). El proceso de formación de observadores y la obtención de la fiabilidad en metodología observacional para analizar la dinámica de juego en minibásquet. *Apunts, Educació Física i Esports, 4*(98), 40-45.

Arias, J.L., Argudo, F.M., & Alonso, J. I. (2009b). Método objetivo para analizar dos modelos de la línea de tres puntos en minibasket. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, *9*(36): 349-365.

Barlett, R. M. (2001). Performance analysis: Can bringing together Biomechanics and notational analysis benefit coaches? *International Journal of Performance Analysis in Sport, 1*(1), 122-126.

Blanco, A., Ibáñez, S. J., Antúnez, A., & Hernández-Mendo, A. (2014). Estudio de fiabilidad de los indicadores de rendimiento en liga asobal. *Cuadernos de Psicología del Deporte, 15*(1), 255-264.

Brennan, R. L., & Prediger, D. J. (1981). Coefficient kappa – some uses, misuses, and alternatives. *Educational and Psychological Measurement, 41*(3), 687-699.

Brewer, C. J., & Jones, R. L. (2002). A five-stage process for establishing contextually valid systematic observation instruments: The case en rugby unión. *The Sports Phychologist, 16*, 138-156.

Cañadas, M., Ibáñez, S. J., & Leite, N. (2015). A novice coach´s planning of the technical and tactical content of youth basketball training: a case study. *International Journal of Performance Analysis in Sport, 15*, 572-587.

Casal, C., Losada, J. & Ardá, T. (2015). Análisis de los factores de rendimiento de las transiciones ofensivas en el fútbol de alto nivel. *Revista de Psicología del Deporte, 24*(1), 103-110

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46.

Chillón, P., & Delgado, M. A. (2012). Observación del profesor de educación física: una investigación de salud en el aula. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, *12*(47): 493-521.

Den Hollander, S., Brown, J., Lambert, M., Treu, P., & Hendricks, S. (2016). Skills Associated with Line Breaks in Elite Rugby Union. *Journal of Sport Science & Medicine, 15*(3), 501-508.

Denis, D., Lortie, M., & Bruxelles, M. (2002). Impact of observers´experience and training on reliability of observations for a manual handling task. *Ergonomics, 45*(6), 441-454.

Escobar, J., & Cuervo, A. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en Medición*, *6*, 27-36.

Gamonales, J. M., León, K., Muñoz, J., González-Espinosa, S., & Ibáñez, S. J. (2018). Validación del IOLF5C para la eficacia del lanzamiento en fútbol para ciegos. Revista Internacional de Medicina y Ciencias del Deporte, 18(70), 361-381. http://dx.doi.org/10.15366/rimcafd2018.70.010

García, J., Ibáñez, S. J., Gómez, M, A., & Sampaio, J. (2009). Basketball game-related statistics discriminating ACB league teams according to game location, game outcome and final score diferences. *International Journa of Performance Analysis in Sport, 14,* 443-452.

García, J., Ibáñez, S. J., Parejo, I., Feu, S., & Cañada, M. (2011). Diferencias entre nivel de juego y categoría de los jugadores en etapas de formación. *Revista Española de Educación Física y Deportes, 395*(21), 13-28.

García, P., Argudo, F., & Alonso, J. I. (2007). Validación de un entrenamiento de observadores para el análisis de una microsituación de juego en waterpolo*. Lecturas: Educación Física y Deportes, 12*(109), 1-18.

González-Díaz, C., & Iglesias-García, M. (2015). Coder Training aplicado a la investigación sobre la publicidad de alimentos. *Opción, 31*(4), 545-563.

Gulick, D. T, & Malone, L. A. (2011). Field test for measuring aerobic capacity in Paralympic goalball athletes. *International Journal of Athletic Therapy and Training, 16*(5), 22-25.

Hernández-Mendo, A., Montoro, J., Reina, Á., & Fernández, J. (2012). Desarrollo y optimización de una herramienta observacional para el bloqueo en voleibol. *Revista Iberoamericana de Psicología del Ejercicio y el Deporte, 7*(1), 15-31.

Hileno, R., & Buscá, B. (2012). Herramienta observacional para analizar la cobertura del ataque en voleibol. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte, 12*(47), 557-570.

Hughes, M. D., & Bartlett, R. M. (2002). The use of performance indicators in performance analysis. *Journal of Sports Sciences, 20*(10), 739-754.

Hughes, M. D., & Franks, I. (2005). Analysis of passing sequences, shots and goals in soccer. *Journal of Sports Sciences, 23*(5), 509-514.

Hughes, M. D., Cooper, S. M., & Nevil, A. (2002). *Analysis procedures for non-parametric data from performance in sport* (2ª ed.). Nueva York: Routledge.

Ibáñez, S. J., Santos, J. A., & García, J. (2015). Multifactorial analysis or free throw shooting in eliminatory basketball games. *International Journal of Performance Analysis in Sport, 15*, 897-912.

James, N., Taylor, J. B., & Stanley, S. (2007). Reliability procedures for categorical data in performance analysis. *International Journal of Performance Analysis in Sport, 7*, 1-11.

Jiménez, J., & Hernández-Mendo, A. (2016). Análisis de la calidad del dato y generalizabilidad de un sistema de observación del contraataque en el balonmano de élite. *E-balonmano.com: Revista de Ciencias del Deporte, 12*(1), 31-34.

Landís, J. R. & Koch, G. G. (1977). Aplication of hierarchical kappa-type statistics in assessment of majority agreement among multiple observers. *Biometrics, 33*(2), 363-374.

Losada, J. L., & Manolov, R. (2014). The process of basic training, applied training maintaining the performance of an observer. *Quality & Quantity, 48*(2). DOI 10.1007/s11135-014-9989-7

McGarry, T. (2009). Applied and theoretical perspectives of performance analysis in sport: Scientific issues and challenges. *International Journal of Performance Analysis in Sport, 9*(1), 128-140.

Medina, J. (1996). Proceso de entrenamiento de codificadores para el estudio de los diarios del profesorado de educación física. *Revista Motricidad, 2*, 113-127.

Medina, J., & Delgado, M.A. (1999). Metodología de entrenamiento de observadores para investigaciones sobre E.F. y Deporte en la que se utilice como método la observación. *Revista Motricidad, 5*, 69-86.

Molik, B.; Morgulec-Adamowicz, N.; Kosmol, A.; Perkowski, K.; Bednarczuk, G.; Skowroński, W.; Gómez, M.; Koc, K.; Rutkowska, I.; & Szyman, R. J. (2015). Game Performance Evaluation in Male Goalball Players. *Journal of Human Kinetics*, *48*(1), 43-51.

Montero, I., G., & León O. (2007). A guide for naming research studies in Phychology. *International Journal of Clinical and Health Psychology*, *7*(3), 847-862.

Morato, M., Da Cunha, O., Gamero, D., Magalhães, T., & Almeida, J. (2016). Development and evaluation of an observational system for goalball match analysis. *Revista Brasileira de Ciências do Esporte, xx*(xx), xx-xx. Pendiente de publicación / In press. http://dx.doi.org/10.1016/j.rbce.08.002

Moreno, M. P., Santos, J. A., Ramos, L. A., Sanz, D., Fuentes, J. P., & Del Villar, F. (2002). Aplicación de un sistema de codificación para el análisis de contenido de la conducta verbal del entrenador de voleibol. *Revista Motricidad, 9*, 119-140.

Morillo, J., & Hernández-Mendo, A. (2015). Análisis de la calidad del dato de un instrumento para la observación del ataque en balonmano playa. *Revista Iberoamericana de Psicología del Ejercicio y el Deporte, 10*(1), 15-22.

O'Donoghue, P. (2007). Reliability issues in performance analysis. *International Journal of Performance Analysis in Sport, 7*(1), 35-48.

O'Donoghue, P. (2010). Research methods for sports performance analysis. London: Routledge.

O'Donoghue, P., & Cullinane, A. (2011). A regression-based approach to interpreting Sports performance. *Internacional Journal of Performance Analysis in Sport, 11*, 295-307.

Ohnjec, K., Vuleta, D., Milanović, D., & Gruić, I. (2008). Performance indicators of teams at the 2003 World Handball Championship for women in Croatia. *Kinesiology, 40*(1), 69-79.

Oliveira, E. P. L.; & De Martino, J. M. (2012). Sistema de análise tática para a modalidade paraolímpica Goalball explorando processamento paraleo. En *Quinto Encontro dos Alunos e Docentes do Departamento de Engenharia de Computacao e Automacao Industrial*. Universidad Estatal de Campinas, Brasil.

Ortega, E., Fernández, R., Ubal, M., Lorenzo, A., & Sampaio, J. (2010). Indicadores de rendimiento defensivo en baloncesto en los ganadores y perdedores. *International Journanl of Sport Sciences, 6*(19), 1-12.

Owen, G. (2014). *Exploratory analysis of goalball: A regression based aproach*. Tesis Doctoral. Universidad de Chester.

Pieron, M. (1986). *Enseñanza de las actividades física y deportivas. Observación e investigación*. Málaga: Editorial Unisport.

Piñar, M. I. (2005). *Incidencia del cambio de un conjunto de reglas de juego sobre algunas de las variables que determinan el proceso de formación de los jugadores de minibásquet (9-11 años).* Granada: Universidad de Granada.

Randolph, J. J. (2005). Free-Marginal Multirater Kappa (multirater Kfree): An Alternative to Fleiss´ Fixed-Marginal Multirater Kappa. Joensuu Learning and Instruction Symposium 2005, University of the Joensuu, Finland. In http://www.eric.ed.gov/PDFS/ED490661.pdf

Rodríguez, G., Gil, J., & García, E. (1996). *Métodos de investigación cualitativa*. Málaga: Aljibe.

Salas, J., & Hernández-Mendo, A. (2016). Análisis de la calidad del dato y generalizabilidad de un sistema de observación del contraataque en el balonmano de élite. *E-balonmano.com: Revista de Ciencias del Deporte, 12*(1), 31-44.

Servera, M., & Cardo, E. (2006). Children Sustained Attention Task (CSAT): Normative, reliabillity, and validity data. *International Journal of Clinical and Health Psychology, 6*, 697-707.

Tosim, A., Junior, A. P., Leitão, M. T., & Simões, R. (2008a). Sistemas técnicos e táticos no Goalball. *Revista Mackenzie de Educacao Física e Esporte, 7*(2), 141-148 .

Tosim, A., Massolli , G., & Beltrao, M. A. (2008b). Seleção da percepção auditiva e tátil em atletas de Goalball em situação defensiva de jogo. Revista Mackenzie de *Educacao Física e Esporte, 7*(3), 181-187.

Trenado, M. T., Pons-Salvador, G., & Cerezo, M. A. (2014). Interacción Temprana: Evaluación de la Fiabilidad del Sistame Observacional CITMI-R, Versión Inglesa. *Revista Electrónica de Metodología Aplicada, 19*(1), 29-42.

Ureña, A., Morales-Rojas, J.S., León, J. & González, M. (2014). El voleibol español infantil. Estudio del nivel de participación en el juego. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, *14*(56): 771-786.

Usuabiaga, O., Castellano, J., Blanco-Villaseñor, A. & Casamichana, D. (2013). La Teoría de la Generalizabilidad en las primeras fases del método observacional aplicado en el ámbito de la iniciación deportiva: calidad del dato y estimación de la muestra. *Revista de Psicología del Deporte, 22*(1), 103-109.

Valles, M. (2003). *Técnicas cualitativas de investigación social: Reflexión metodológica y práctica profesional*. Madrid: Editorial Síntesis.

Viciana, J. (1999). Proceso de entrenamiento de grupos de codificadores específicos antes un sistema múltiple de categorías de análisis del discurso en un grupo de trabajo de docentes en educación física. *Revista Motricidad, 5*(1), 53-67.

Volossovitch, A., Dumangane, M., & Rosati, N. (2012). Factores que influenciam a dinâmica do rendimento colectivo no andebol. *Revista Española de Educación Física y Deportes, 396*(22), 13-33.

Weber, C., & Link, D. (2016). Performance Analysis in Goalball. In *Proceedings of the 10th International Symposium on Computer Science in Sports (ISCSS)* (pp. 157-160). Springer International Publishing.

Referencias totales / Total references: 69 (100%)
Referencias propias de la revista / Journal's own references: 5 (7,2%)