Article

Development and Validation of an Observational Game Analysis Tool with Artificial Intelligence for Handball: Handball.ai

Moises Marquina [1], Demetrio Lozano [2], Carlos García-Sánchez [1,*], Sergio Sánchez-López [3] and Alfonso de la Rubia [1]

1   Deporte y Entrenamiento Research Group, Departamento de Deportes, Facultad de Ciencias de la Actividad Física y del Deporte (INEF), Universidad Politécnica de Madrid, C/Martín Fierro 7, 28040 Madrid, Spain; moises.mnieto@upm.es (M.M.); alfonso.delarubia@upm.es (A.d.l.R.)
2   Health Sciences Faculty, Universidad San Jorge, Autov A23 km 299, Villanueva de Gállego, 50830 Zaragoza, Spain; dlozano@usj.es
3   SPORT Research Group (CTS-1024), Centro de Evaluación y Rehabilitación Neuropsicológica, University of Almería, Ctra. Sacramento, s/n, 04120 Almería, Spain; ssl832@inlumine.ual.es
*   Correspondence: c.gsanchez@upm.es; Tel.: +34-91-067-8012

Abstract: Performance analysis based on artificial intelligence together with game-related statistical models aims to provide relevant information before, during and after a competition. Due to the evaluation of handball performance focusing mainly on the result and not on the analysis of the dynamics of the game pace through artificial intelligence, the aim of this study was to design and validate a specific handball instrument based on real-time observational methodology capable of identifying, quantifying, classifying and relating individual and collective tactical behaviours during the game. First, an instrument validation by an expert panel was performed. Ten experts answered a questionnaire regarding the relevance and appropriateness of each variable presented. Subsequently, data were validated by two observers (1.5 and 2 years of handball observational analysis experience) recruited to analyse a Champions League match. Instrument validity showed a high accordance degree among experts (Cohen's kappa index (k) = 0.889). For both automatic and manual variables, a very good intra- ((automatic: Cronbach's alpha (α) = 0.984; intra-class correlation coefficient (ICC) = 0.970; k = 0.917) (manual: α = 0.959; ICC = 0.923; k = 0.858)) and inter-observer ((automatic: α = 0.976; ICC = 0.961; k = 0.874) (manual: α = 0.959; ICC = 0.923; k = 0.831) consistency and reliability was found. These results show a high degree of instrument validity, reliability and accuracy providing handball coaches, analysts, and researchers a novel tool to improve handball performance.

Keywords: handball; performance indicators; artificial intelligence

1. Introduction

The analysis and evaluation of sports performance is one of the research areas that has become increasingly relevant in the study of team sports, especially in collaborative—opposition sports [1–3]. To this effect, ad hoc tools have been developed with the operational objective of providing greater information to stakeholders and practitioners to optimise sports performance [4]. Many of these analysis instruments are based on the concepts of 'observational methodology' [5] and 'notational analysis' [6] which, through the collection, analysis and interpretation of data, aim to draw empirical conclusions by assessing the context and dynamics of the sport [7]. This methodology, characterised by being a flexible and rigorous procedure [8], aims to contribute to a deeper analysis of the dynamic reality of the game by understanding the determining performance factors, as well as the detection of evolutionary trends [9].

The efficient use of observational instruments for game analysis requires two essential conditions: validity and reliability. While the former evaluates the adequacy and appropriateness of an item or aspect [8], the latter refers to the reproducibility of the values of a test or other measurement in repeated trials on the same individuals [10]. Reliability tests in sports performance analysis are conducted to ensure objectivity in the data collection process through inter-observer agreement [11]. Nevertheless, concordance within the same observer is also necessary to guarantee data quality. Thus, an observational instrument should allow high-reliability values, both between observers (inter-observer) as well as in the different observations of the same subject (intra-observer), and it is necessary to report these data quality processes [9].

Studies in collaborative/opposition team sports based on the use of observational methodology have multiplied exponentially. Consequently, the design and validation of ad hoc observational instruments have become common and frequent in elite sports, aimed at identifying performance factors [12]. One of the sports where the production has been most prolific is football. Thus, in the systematic review by Clemente et al. [13], 21 studies were identified, of which 8 focused on tactical tests (e.g., FUT-SAT), 10 on technical tests (e.g., Loughborough Soccer Passing Test) and 3 on tests concurrent with physical aspects (e.g., MT5M). In other sports (e.g., basketball) there is an amalgam of instruments with different purposes, such as the influence of contextual variables on tactical thinking [14], the player discrimination by position according to physical–technical execution patterns [15] or the relevance of physiological parameters [16]. In other sports such as rugby [17], water polo [18] and ice hockey [19], tools based on observational methodology have also been developed.

Regarding handball, the scientific evidence shows a clear orientation towards the design of observational instruments for technical–tactical game analysis, mainly at the offensive level [20]. Some examples focused on the effectiveness of offensive game systems [21] and offensive tactical behaviour in critical phases [22], on the reliability of the register in the counterattack phase [23], on attacking effectiveness according to game and contextual variables [24], on the effectiveness of offensive actions according to gender [25], on the evaluation of individual technical–tactical performance in competition [26] or on offensive performance in elite handball [27,28]. Specifically, numerous investigations based on observational methodology at the offensive level focused on the analysis of the throwing action with the aim of evaluating the specific weight of each playing position in the performance [29–35]. Moreover, due to the regulatory changes introduced with regard to the 'empty goal' rule [36], more and more authors are concerned with designing valid and reliable tools to analyse 7 × 6 situations [37]. Thus, these kind of records are in full expansion phase in modalities such as beach handball [38,39] or associated with the combined analysis of the physical and physiological demands of players [40] and referees [41] using polar coordinates.

One of the fastest-growing research and development areas in handball is performance analysis based on artificial intelligence—AI [42]. The adoption of AI and statistical models, based on game-related statistics obtained through notational analysis and observational methodology [43], has arisen with the clear objective of providing information before, during and after the competition [42]. One of the most pursued study objects has been team discrimination—winners and losers. Research has been carried out in different sport contexts, such as the Olympic Games [44–46], World Championships [47–51], European Championships [52], Pan-American Championships [53], European Champions League [54,55], national leagues [56,57], semi-professional or amateur contexts [58,59] and youth competitions [60]. Other investigations based on game-related statistics focused on the role played by the goalkeeper [61], the impact of the 'empty goal' rule of substitution of the goalkeeper by a court player [62], the influence of variables associated with sport experience [63], the establishment of an individual performance index rating—'Play Score' [64]—the defence-attack relationship [65], the relevance of contextual variables [66] or coaching decisions (e.g., time-outs) [67].

To the best of our knowledge, no handball performance analysis software has been able to determine in depth the impact of factors affecting the dynamics of game pace in handball (e.g., number of 1×1 actions that are successfully resolved by the offense team) beyond the outcome of the actions (i.e., goal or no goal), homogenising the tracking, automatically generating statistical models in situ and classifying players according to specific performance parameters (e.g., number of times that a player/team was involved in the possession of the ball, in general and in the different phases of the game). Therefore, and because of the frequent use of observational methodology for technical–tactical game analysis in handball according to performance factors and the present relevance of AI based on game-related statistics in sport, the objective of this study was to design and validate a recording instrument based on observational methodology (hereafter, Handball.ai) capable of identifying, quantifying, classifying and relating individual and collective tactical behaviours in handball, ensuring compliance with the principles of validity, reliability, accuracy and generalisability. The authors hypothesise that Handball.ai will further analyse competitive handball performance according to automatic time and event control (i.e., number of possessions per player and per team), instantaneously generate statistics based on mathematical models independently of the tracker's intervention (i.e., player score) and complete simplified event tracking based only on the team in possession of the ball.

2. Materials and Methods

A nomothetic observational, follow-up and multidimensional design was used for the present study [8]: nomothetic because there is a sample of several teams; follow-up because several championships are analysed over time and compared with each other; and multidimensional because different dimensions are taken into account. This N/S/M design leads to a series of decisions about the sample, the observation-recording instruments and the analysis procedure.

The process begins with data capture, which can be accomplished through different means, such as video cameras, sensors or tracking devices. These data are collected during sporting events and provide valuable information on athlete performance, tactics used, game dynamics and other relevant aspects. Once the data are collected, AI comes into play. Machine learning algorithms and natural language processing are used to analyse and extract meaningful information from observational data. These algorithms are pre-trained using sample data sets containing labelled examples, allowing the AI to learn patterns and trends from the data [68].

Therefore, the AI in the Handball.ai tool can perform a wide variety of tasks in sports observational analysis. For example, it can automatically identify and follow players during a match, detect movement patterns and analyse the tactics and strategies used by teams. It can also perform advanced statistical analysis, such as calculating the average performance of a player in different situations or comparing the performance of different teams. In addition, the AI can provide visualisations and graphical representations of the data, making it easier for coaches, players and fans to understand and analyse. These visualisations can include motion graphs, heat maps, tactical diagrams and other visual elements that help reveal hidden patterns and trends in observational data.

2.1. Procedures

The design and validation of the instrument followed a four-stage process (Figure 1). The first phase of the research was a literature review that allowed the definition of the theoretical framework and the procedures to be carried out and the determination of the study design [69].

The second phase was to define the variables in a criteria system and category analysis. Descriptors were established, characterising the behaviours and classifying them, based on the scientific literature, into different categories and variables (Table 1).

Table 1. Definitions of the game-related statistics.

| Dimension	Variable	Definition
Period	First Half	Division of the time duration of a handball match. Comprises the first and second half.
	Second Half	
Team	Name	Name of each of the contending clubs or teams in the match.
Event	Goal	Moment that one team with the possession of the ball gets the ball into the opponent's goal.
	Save	Moment that the goalkeeper stops the ball from a shot by the opposing team.
	Post Out	Moment that one team with the possession of the ball realizes one shot, with the finality to score a goal, but the shot goes out or against the post.
	Attack Fault	Infraction committed by a player as a result of a foul on a defender indicated by the referee.
	Lost Ball	The attacking team loses ball possession, and the defending team takes ball possession.
	Technical Mistake	Loss of ball possession through a rule infringement, e.g., steps, double dribble.
	Passive	Infraction indicated by the referee if the team in ball possession does not change its attacking behaviour or does not take a throw-in after the referee has indicated the passive play warning.
	Fault	Stopping of an opposing player by the player in ball possession without generating any suspension (yellow card, red card, blue card and 2 min).
	Shot Blocked	Technical action of a defending player who intercepts the trajectory of the ball in a throwing attempt on goal.
	2 min provoked	Player in ball possession attempts to score a goal and is illegally prevented from doing so by another player on defence, incurring a 2 min suspension.
	7 m provoked	Player in ball possession tries to score a goal and another player on defence stops him, illegally interrupting a clear goal-scoring opportunity.
	7 m provoked + 2 min suspension	Player in ball possession tries to score a goal and another player in defence stops him illegally, interrupting a clear goal-scoring opportunity and incurring a 2 min suspension.
	1 × 1 Won	Player in ball possession overtakes the player on defence, generating a clear goal-scoring opportunity or a 7 m provoked, 2 min suspension and 7 m provoked + 2 min suspension.
	2 min	Exclusion of a player for 2 min indicated by the referee due to repeated fouls, unsporting conduct, an incorrect change or as a result of a disqualification.
	Yellow Card	Disciplinary sanction awarded by the referee to a player and/or coaching staff member as a result of a fault or unsporting behaviour.
	Red Card	Disciplinary sanction awarded by the referee to a player and/or coaching staff member for severe misconduct or for receiving a 2 min suspension three times.
	Blue Card	Disciplinary sanction imposed by the referee on a player and/or coaching staff member as a consequence of serious misconduct or misconduct involving a written report.
Phase	Positional Offense	Chance to score a goal by the attacking team against a structured defence. Action of attacking teams playing with structured attack in the opponent's field.
	Fastbreak	Recovery of ball possession by the defence due to an event (e.g., turnover) and this team searches for a clear action to score a goal against an unstructured defence.
	Countergoal	Recovery of ball possession by the defence due to a goal, and this team searches for a clear action to score a goal against an unstructured defence.
Defence System		Structure that both teams have in any moment of defending time during handball game.
Offense System		Structure that both teams have in any moment of attacking time during handball game.
Offense Player Position		Playing position occupied by an offensive player on the court
Defence Player Position		Playing position occupied by a defence player on the court
Shooter Position		On the field (left wing, 6 m left, 6 m centre, 6 m right, right wing, 9 m left, 9 m centre, 9 m right, opposite field and 7 meters).
Shot Location		(In the goal) with a difference in number 1, 2, 3 (from left to right up); 4, 5, 6 (from left to right in the middle); and 7,8,9 (from left to right down).
Possession		Number of times that a player/team was involved in the possession of the ball, in general and in the different phases of the game (positional offense, countergoal and fastbreak).
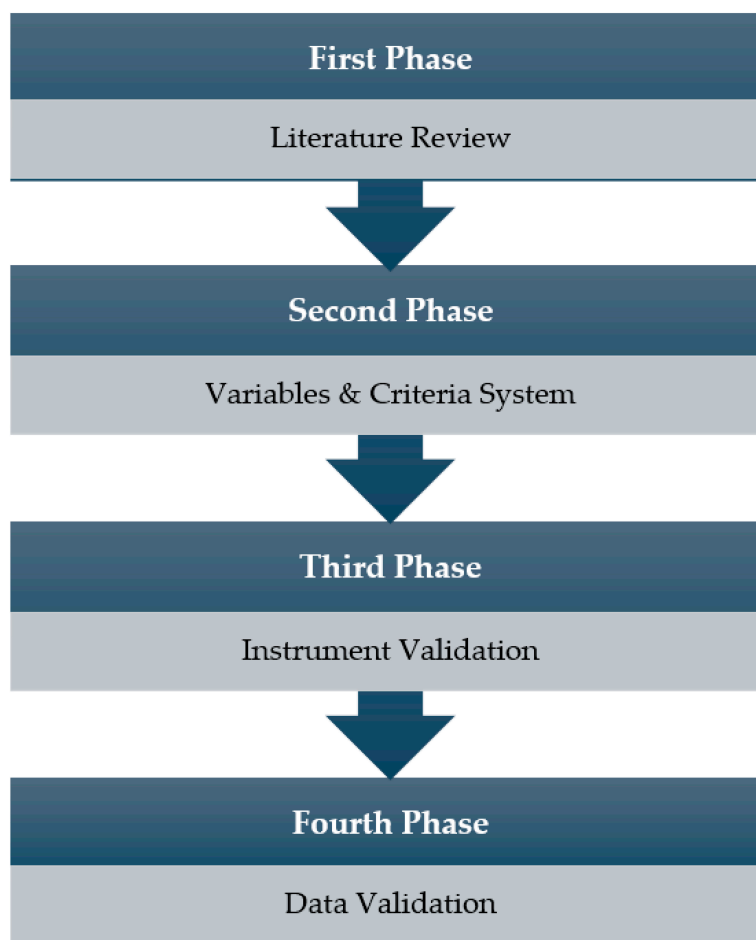
Figure 1. Structure of different phases of the validation process.

The third phase was the validation of the instrument. The content validity was determined by expert judgement. First, a questionnaire was created. Ten experts of different nationalities participated in the validation process.

An invitation explaining the context, description and purpose of the study, together with a guide to assess the following aspects of the components and their levels, was sent via email: The experts assessed the "wording" and "suitability" sections of each item, using a quantitative Lickert-type scale from 1 to 10. In addition, they made an overall qualitative assessment of each element regarding whether they considered it appropriate, where they expressed their alternatives to certain aspects that they would personally improve. To do so, they recorded: (a) the degree to which each item belongs to the object of study ("suitability"). A quantitative score was given by each expert to each unit of the instrument to determine the relevance with which each item should form part of the observation instrument. (b) Degree of precision and correctness (wording): by means of using quantitative scores given by the experts, the degree to which each item was correctly worded and defined was observed. (c) Degree of qualitative comprehension: the reflections provided by the experts on the different questions of the instrument were collected.

Subsequently, the data were transformed to a 0.1–1 scale, averaging the experts' responses for each criterion and category, according to factor. All the categories of the instrument presented agreement and acceptance values above 0.8 and were accepted [70].

Once the instrument had been validated, a fourth and final phase was carried out to validate the data. For this purpose, the observers were trained and made several observations of the same match to achieve data reliability using Cohen's kappa concordance coefficient [71], as indicated by several studies [58,72].

2.2. Participants

To validate the instrument (third phase), 10 experts of different nationalities participated in the validation process. The sample who participated to validate the instrument had to meet the following inclusion criteria established in the research: (I) qualification: coach with Master Coach level certified by the EHF (European Handball Federation) or Level III EHF Rinck Convention; (II) years of experience: have been training for more than 10 years in national and international categories; and (III) current season: be training in the national category in the current season [73].

For data validation (fourth phase), two observers were recruited. These expert observers had Level III EHF Rinck Convention certification with 1.5 and 2 years of experience, respectively, in analysing handball using observational software. Both observers were of different nationalities (mean age: 36 years) and had experience as coaches for 13 ± 4 years. A training process of the observers lasted 21 days.

The match selected for validation corresponded to the Champions League round of 16 match between Telekom Veszprém HC and OTP Bank PICK Szeged, played on 30 March 2023. To analyse the match, data were accessed using the match recording made by the broadcasting rights company according to the data collection criteria [74].

To define the study sample, the following inclusion criteria were used: (a) the match to be analysed had to belong to an international competition. Because this type of competition has a high level of competitiveness, to access this championship, teams must qualify in their respective national leagues and then pass a preliminary phase that only allows the teams classified in the first positions to continue. For this reason, a match of the 2022/2023 Champions League was used.

2.3. Observation Tool

An ad hoc software was designed to record the individual actions of players during a handball match in real time. This software allows working with the multidimensional observation system and was named Handball.ai™ (Version: 1.2.4 (Heidelberg, Germany)). This software allows the collection of match data through a graphical interface containing two windows, one showing the video of the match and the other a graphical representation of the players and their numbers and the different team systems. The software is presented on a screen where different buttons appear to record each variable. A mobile device with a touch screen (tablet) was used to record all the events in real time as quickly and easily as possible (Figure 2).
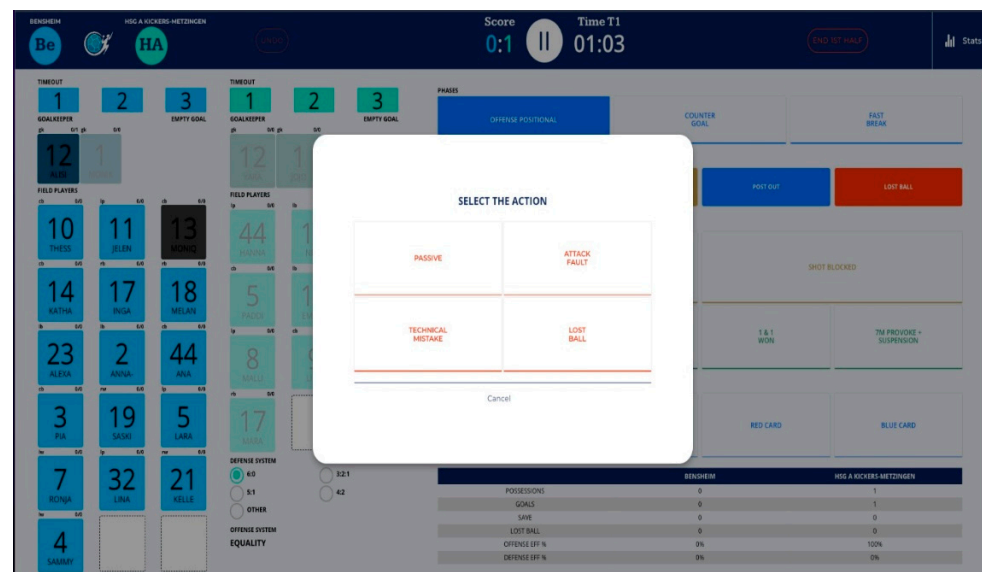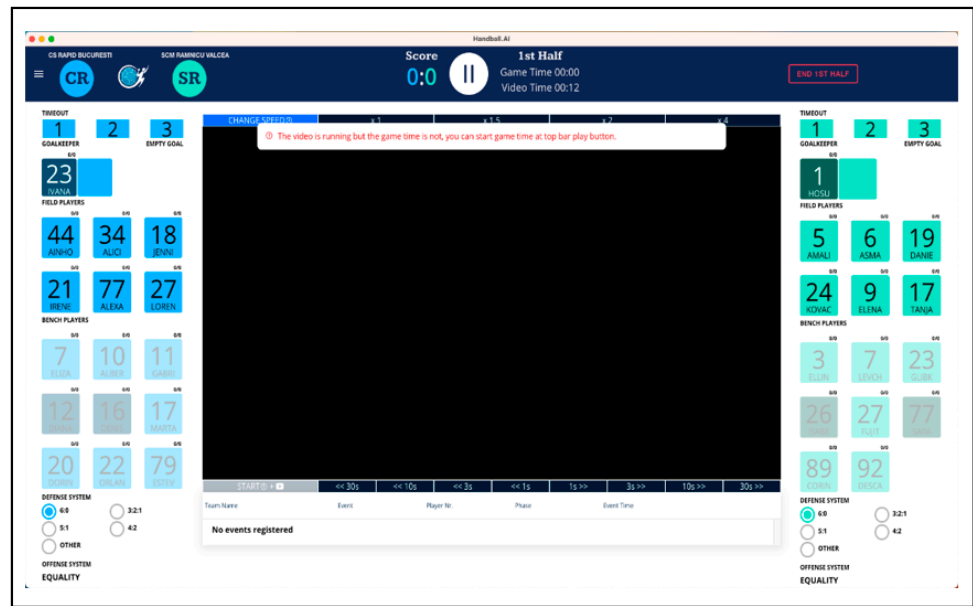

Figure 2. *Cont.*

Figure 2. Buttons of the different actions of Handball.ai.

The following variables were defined and described to clarify data collection. The variables that define the individual actions of the players were divided into groups (Table 1): period, team, event, phase, offense system, defence system (Figure 3), offense player position, defence player position, shooter position (Figure 4), shot location (Figure 5) and possession.
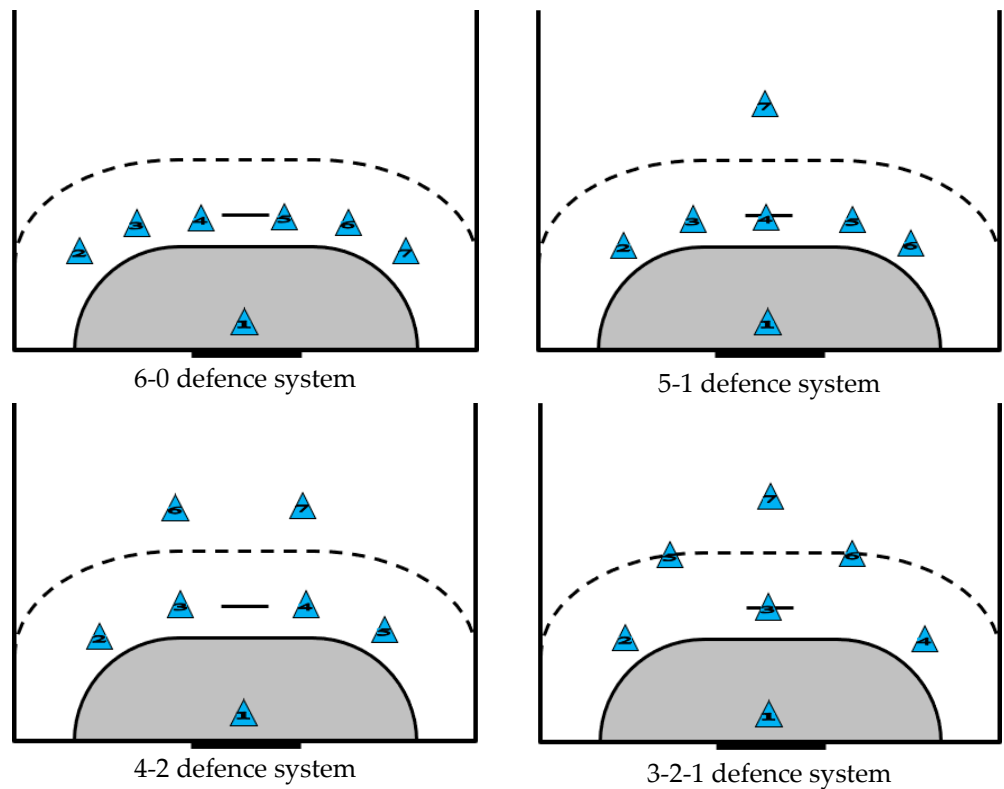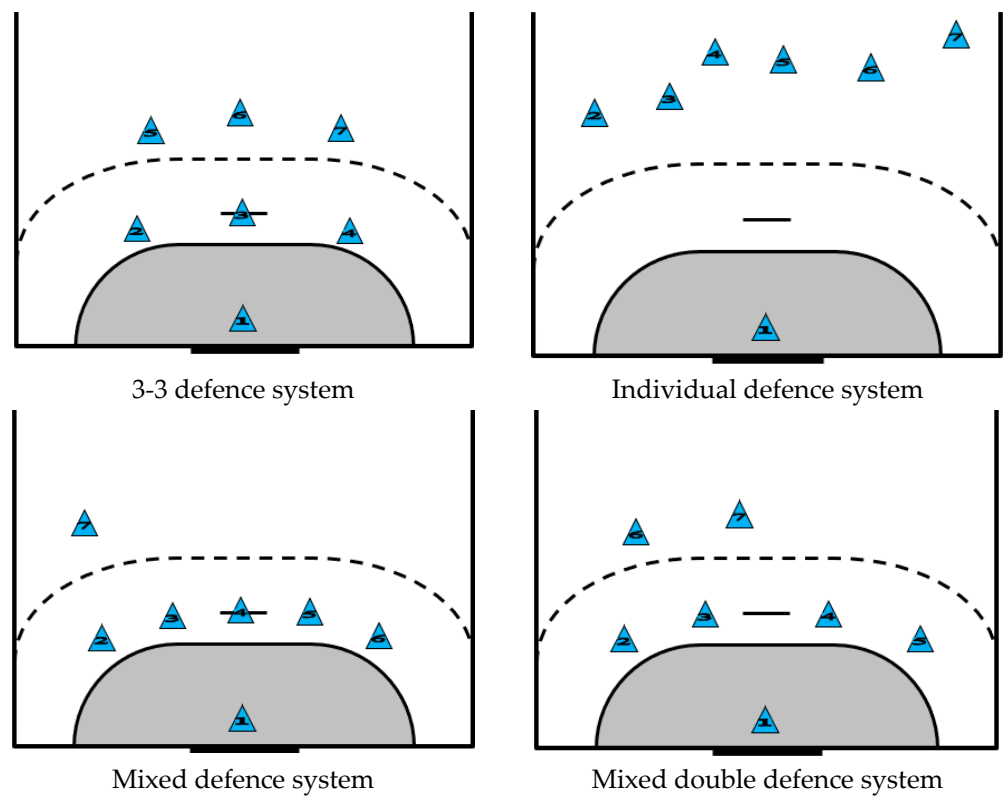

6-0 defence system

5-1 defence system

4-2 defence system

3-2-1 defence system

Figure 3. *Cont.*

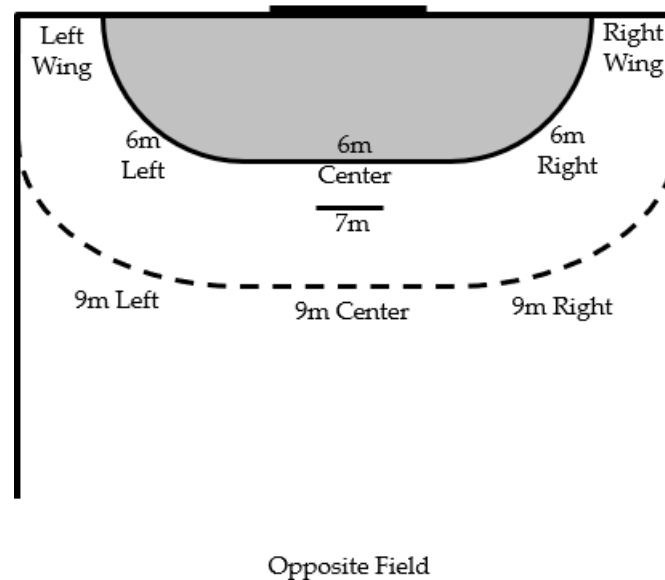**Figure 3.** Graphic representation of the different defence systems.

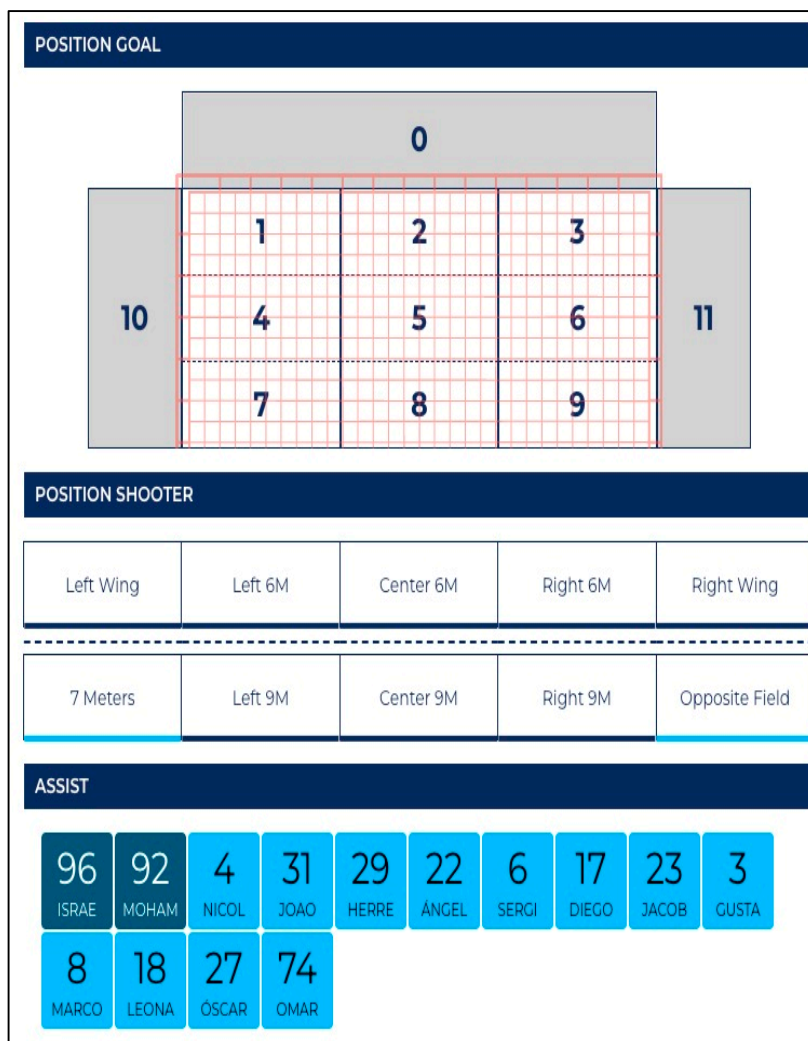**Figure 4.** Graphic representation of shooter position.

Figure 5. Graphic representation of shot location.

Within the different dimensions, some of the variables are calculated manually, i.e., by the observer clicking. On the other hand, there are other variables calculated automatically by the software. This calculation is performed by means of automatic calculation parameters, i.e., an automatic calculation based on manually marked events (Table 2).

Table 2. Dimensions analysed and their calculation method.

Dimension	Tracking
Period	Automatic
Team	Automatic
Offense System	Automatic
Offense Player Position	Automatic
Defence Player Position	Automatic
Possession	Automatic
Duration	Automatic
Players	Automatic
Event	Manual
Phase	Manual
Defence System	Manual
Shooter Position	Manual
Shot Location	Manual

2.4. Data Validation

The inter-observer reliability assessment process was similar to methods used in previous research [58,75]. The onboarding of the observers consisted of the following phases: (1) meeting with the platform developer to get to know the platform and become familiar with it; (2) from 1 week to 10 days for the observers to test it and ask questions about the process; (3) a 2-h session, normally separated by 1 h, where they observed a match together with the platform developer and recorded the observed data simultaneously; and (4) a session made up of the analysis of 2 complete live matches, comparing the results. In each training session the observers were subjected to the same conditions: (i) the observer was isolated in a room to maintain the intra-sessional connection; and (ii), observations were made at the same time and in the same place under stable conditions and without the presence of any person who might interfere directly or indirectly in the process, except the researcher.

Once the training period was over, observation was carried out to validate the data. The same match was recorded by the two observers, and the results were compared between them (inter-observer). One week later, the same match was observed again by the same observers, and the results were compared with the previous analysis (intra-observer).

The reliability and internal consistency of the data collected were analysed. To determine the case proportion, Cohen's kappa index (k) was used, which analyses the agreement degree between observers after excluding the proportion of cases in which the agreement between them is the result of chance [76]. Kappa values can range from −1.0 to 1.0. Agreement in the interpretation of the kappa value was assessed as follows: <0 no agreement; 0.01–0.20 poor agreement; 0.21–0.40 fair agreement; 0.41–0.60 moderate agreement; 0.61–0.80 good agreement; and 0.81–0.99 very good agreement [77]. Cronbach's alpha (α) and intraclass correlation coefficients (ICC) were used to determine internal consistency. Internal consistency indices and reliability thresholds ranged from 0 to 1 [78] and were set at: for α (internal consistency), < 0.50 unacceptable, 0.51–0.60 poor, 0.61–0.70 questionable, 0.71–0.80 acceptable, 0.81–0.90 good and ≥ 0.91 excellent [79]; for ICC (reliability), ≤ 0.50 poor, 0.51–0.75 moderate, 0.76–0.90 good and ≥ 0.91 excellent [80].

3. Results

Cohen's kappa index (k) was used to determine the degree of agreement among the experts. A Cohen's kappa concordance was achieved of very good agreement at 0.889 (Table 3). The analysis of variance (one-factor ANOVA) regarding variables of ranges with a normal distribution (Table 3) did not show significant differences between the averages of the different experts (0.101 f; 1.000 sig).

Table 3. Validation data experts panel.

Variables	Expert 1 κ	Expert 2 κ	Expert 3 κ	Expert 4 κ	Expert 5 κ	Expert 6 κ	Expert 7 κ	Expert 8 κ	Expert 9 κ	Expert 10 κ
Period	1	1	1	1	1	1	1	1	1	1
Team	1	1	1	1	1	1	1	1	1	1
Offense System	1	0.78	1	0.78	1	0.8	1	0.78	1	0.82
Offense Player Position	0.79	0.64	0.82	0.64	0.78	0.78	0.64	0.64	0.78	0.78
Defence Player Position	0.81	0.78	0.83	0.78	0.82	0.82	0.78	0.78	0.82	0.82
Possession	1	1	1	1	1	1	1	1	1	1
Duration	0.8	0.82	0.79	0.82	0.78	0.82	0.82	0.82	0.82	0.82
Players	1	1	1	1	1	1	1	1	1	1
Event	1	1	1	1	0.82	1	0.82	1	0.82	1
Phase	0.8	0.78	0.61	0.78	0.78	0.82	0.78	0.78	0.78	0.82
Defence System	1	1	1	1	1	1	1	1	1	1
Shooter Position	0.79	0.82	0.78	0.82	0.82	0.78	0.82	0.82	0.78	0.78
Shot Location	0.81	0.78	0.77	0.78	0.8	0.8	0.8	0.78	0.8	0.8
Mean	0.908	0.877	0.892	0.877	0.892	0.894	0.882	0.877	0.892	0.895
Total						0.889				
ANOVA (Sig) 1.000 F (0.101)										

Note: κ: Cohen's kappa. Sig: Significance. F: Variance.

Table 4 shows the mean intra-observer and inter-observer reliability and internal consistency tests of the automatic variable data collected by the Handball.ai software. Both can be considered very good at 0.957 and 0.937, respectively.

Table 4. Validity coefficients for intra- and inter-observer internal consistency (Cronbach's alpha) and reliability (intra-class correlation coefficients and Cohen's kappa) for automatic variables.

Variable Group	Intra-Observer			Inter-Observer		
	α	ICC (95%)	κ	α	ICC (95%)	κ
Period	1	1	1	0.995	0.989	0.929
Team	0.951	0.907	0.908	0.989	0.978	0.919
Duration	-	0.977	0.955	0.895	0.845	-
Offense System	0.986	0.972	0.98	0.992	0.984	0.91
Offense Player Position	0.982	0.965	0.958	0.984	0.969	0.915
Possession	1	1	0.976	1	1	0.958
Team Players		-	0.645	-	-	0.61
Mean	0.984	0.970	0.917	0.976	0.961	0.874
		0.957			0.937	

Note: α: Cronbach's alpha; ICC: Intra-class correlation coefficients; κ: Cohen's kappa.

Table 5 shows the means of the intra-observer and inter-observer reliability and internal consistency tests for the data of the manual variables collected by the observers. Both can be considered very good at 0.913 and 0.904, respectively.

Table 5. Validity coefficients for intra- and inter-observer internal consistency (Cronbach's alpha) and reliability (intra-class correlation coefficients and Cohen's kappa) for manual variables.

Variable Group	Intra-Observer			Inter-Observer		
	α	ICC (95%)	κ	α	ICC (95%)	κ
Event	0.996	0.993	0.942	1	0.999	0.955
Phase	0.891	0.803	0.83	0.909	0.833	0.814
Shooter Position	0.945	0.895	0.821	0.981	0.963	0.601
Defence Player Position	0.984	0.968	0.797	0.992	0.984	0.873
Shot Location	0.977	0.955	0.901	0.912	0.838	0.913
Mean	0.959	0.923	0.858	0.959	0.923	0.831
		0.913			0.904	

Note: α: Cronbach's alpha; ICC: Intra-class correlation coefficients; κ: Cohen's kappa.

4. Discussion

In this study we analysed the validity and the inter- and intra-observer reliability of Handball.ai in identifying, quantifying, classifying and relating individual and collective tactical behaviours in handball. The main results obtained in the reliability tests (ICC and Cohen's kappa) and internal consistency tests (Cronbach's alpha and ICC), intra-observer and inter-observer, demonstrate that Handball.ai is a valid, reliable and precise tool to register individual and collective tactical behaviours in handball. In particular, the data of the automatic variables collected by Handball.ai (ICC = 0.957 and 0.937) and the data of the manual variables collected by the observers (ICC = 0.913 and 0.904) can be considered very good. Therefore, this reveals that the system is a very reliable software capable of registering automatic and manual variables, serving as an observational instrument to analyse elite handball.

The expert panel has played an important role in improving and clarifying the definition, relevance and appropriateness of each variable used in Handball.ai, as in similar studies developed in other sports [17,81,82]. Specifically, the Cohen's kappa index (κ = 0.889) showed a very good agreement between experts in order to validate each variable of the instrument. In addition, one of the main strengths of the expert panel was that the

10 coaches have experience at international competitions and are aware of the observational methodology.

Handball.ai contains some dimensions and variables that have been previously validated in other instruments and that serve to contextualise the situation of the game in which the event takes place, such as the period of the match, the phase of the game, the offense system or the defence system [21–24,26,28]. Some of these variables (e.g., period or offense system) are recorded automatically by Handball.ai, which is the first handball observation instrument that uses the advantages of AI [41,42,69]. This innovation is a great advance with respect to the software developed and validated until now [83,84], as it will reduce the time required to record and code the events during the match.

In addition, the instrument includes other variables to expand the information related to a shot on goal, such as shooter position and shot location. These variables have also been widely used in other observation instruments that have been validated [22–24,26,27,29]. Likewise, Handball.ai is a valid and reliable tool to register game events that have already been considered decisive for achieving maximum performance in handball, such as goals, saves, turnovers, technical mistakes, assists, etc. [43–59].

Therefore, this instrument incorporates new variables that had not been included to date in an observation tool with similar characteristics [21–24,26,28]. One of these variables is the number of possessions, an aspect that will allow us to analyse in depth the dynamics of game pace in competition. It also includes another new variable that allows us to quantify the number of 1×1 actions that are successfully resolved by the offensive team. This variable represents a new advance in handball game analysis, as it will allow us to know exactly which player has generated the offensive superiority, regardless of whether the action ends in a goal or 7 m. This information was not considered in other observation instruments [21–24,26,28] and will be especially useful for handball coaches and performance analysts.

It would be appropriate to note that although Handball.ai has shown good validity and reliability both intra- and inter-observer in elite handball (0.913–0.957 and 0.904–0.937), we should be careful with the use of this software in other handball modalities, such as beach handball and wheelchair handball.

Future studies could evaluate the generalisation of the system with a larger sample of matches in similar or different competitions, such as national teams' championships or national league tournaments.

5. Conclusions

The findings reported in this study demonstrate that Handball.ai is a valid, reliable, and precise tool for identifying, quantifying, classifying and relating individual and collective tactical behaviours in elite handball competitions. In particular, the main results obtained in the reliability tests and internal consistency, both intra-observer and inter-observer, demonstrate that the data of the automatic and manual variables (collected by the Handball.ai and by the observers, respectively) can be considered very good.

Furthermore, this instrument presents two main advantages compared to other software validated to date. Firstly, the use of AI to collect automatic variables will reduce the time required to record and code the events during the match. Secondly, this instrument incorporates new variables, such as number of possessions and number of 1×1 actions that are successfully resolved by the offensive team. These variables will allow coaches to perform a deeper analysis of handball.

Therefore, handball coaches and performance analysts should consider the high reliability and internal consistency of the automatic and manual variables to conduct analysis during elite competitions. Specifically, this tool would help technical staff to: (i) design better training programs focused on improving the decisive technical actions of the game; (ii) make comprehensive analyses of the technical–tactical performance (individual and collective) of their opponents; (iii) develop a more effective game plan; and (iv) improve talent identification and player recruitment. In addition, data collected by Handball.ai

might be used by researchers for the development of new research in performance analysis in handball.

There is no doubt that AI will continue to transform analysis and evaluation of elite sports. Consequently, further research is needed to analyse other AI applications, such as predicting match outcomes, assisting coaches on line-ups, tactics and player rotations, supporting referee decisions during the competition and understanding the value of players from an economic perspective.

References

1. de la Rubia, A.; Ugalde-Ramírez, A.; Gutiérrez-Vargas, R.; Pino-Ortega, J. Does the New Resin-Free Molten D60 Ball Have an Impact on the Velocity and Accuracy of Handball Throws? *Appl. Sci.* **2022**, *13*, 425. [CrossRef]
2. García-Aliaga, A.; Martín-Castellanos, A.; Marquina Nieto, M.; Muriarte Solana, D.; Resta, R.; López del Campo, R.; Mon-López, D.; Refoyo, I. Effect of Increasing the Number of Substitutions on Physical Performance during Periods of Congested Fixtures in Football. *Sports* **2023**, *11*, 25. [CrossRef] [PubMed]
3. Ibáñez, S.J.; Piñar, M.I.; García, D.; Mancha-Triguero, D. Physical Fitness as a Predictor of Performance during Competition in Professional Women's Basketball Players. *Int. J. Environ. Res. Public Health* **2023**, *20*, 988. [CrossRef] [PubMed]
4. Kelly, A.L.; Williams, C.A.; Cook, R.; Sáiz, S.L.J.; Wilson, M.R. A Multidisciplinary Investigation into the Talent Development Processes at an English Football Academy: A Machine Learning Approach. *Sports* **2022**, *10*, 159. [CrossRef]
5. Hughes, M. Notational Analysis. In *Science and Soccer*; Routledge: Oxford, UK, 2003; pp. 253–272.
6. Anguera, M.T. Metodología Observacional. In *Metodología de la Investigación en Ciencias del Comportamiento*; Universidad de Murcia: Murcia, Spain, 1990; Volume 1, pp. 125–236.
7. Anguera, M.T.; Hernández-Mendo, A. La Metodología Observacional En El Ámbito Del Deporte. Observational Methodology in Sport Sciences. *Rev. Cienc. Deporte* **2013**, *9*, 135–160.
8. Anguera, M.T.; Blanco, A.; Losada, J.L. Diseños Observacionales, Cuestión Clave En El Proceso de La Metodología Observacional. *Metodol. Cienc. Comport.* **2001**, *3*, 135–161.
9. Santos, S.; Sarmento, H.; Alves, J.; Campaniço, J. Construcción de Un Instrumento Para La Observación y El Análisis de Las Interacciones En El Waterpolo. *Rev. Psicol. Deporte* **2014**, *23*, 191–200.
10. Atkinson, G.; Nevill, A. Measures of Reliability in Sports Medicine and Science. *Sports Med.* **2000**, *30*, 375–381. [CrossRef]
11. O'Donoghue, P. Reliability Issues in Performance Analysis. *Int. J. Perform. Anal. Sport* **2007**, *7*, 35–48. [CrossRef]
12. Ávila-Moreno, F.M.; Chirosa-Ríos, L.J.; Ureña-Espá, A.; Lozano-Jarque, D.; Ulloa-Díaz, D. Evaluation of Tactical Performance in Invasion Team Sports: A Systematic Review. *Int. J. Perform. Anal. Sport* **2018**, *18*, 195–216. [CrossRef]
13. Clemente, F.M.; Praça, G.; Oliveira, R.; Aquino, R.; Araújo, R.; Silva, R.; Sarmento, H.; Afonso, J. A Systematic Review of the Criterion Validity and Reliability of Technical and Tactical Field-Based Tests in Soccer. *Int. J. Sports Sci. Coach.* **2022**, *17*, 1462–1487. [CrossRef]
14. Reis, C.P.; Morales, J.C.P.; Gomes, C.M.A.; Pereira, F.D.A.A.; Ibáñez, S.J. Construct Validation of a New Instrument to Measure Declarative Tactical Knowledge in Basketball. *Percept. Mot. Skills* **2021**, *128*, 1712–1729. [CrossRef] [PubMed]
15. Vučković, I.; Gadžić, A.; Sekulić, Ž.; Marković, S. Validnost I Pouzdanost Testa Za Procenu Vremena Reakcije I Brzine Kretanja U Odbrambenom Kosarkaskom Stavu. *Facta Univ. Ser. Phys. Educ. Sport* **2022**, *20*, 61–71. [CrossRef]

16. Gottlieb, R.; Shalom, A.; Alcaraz, P.E.; Calleja-González, J. Validity and Reliability of a Unique Aerobic Field Test for Estimating VO2max among Basketball Players. *Sci. J. Sport Perform.* **2022**, *1*, 112–123. [CrossRef]

17. Villarejo, D.; Ortega, E.; Gómez, M.-Á.; Palao, J.-M. Design, Validation, and Reliability of an Observational Instrument for Ball Possessions in Rugby Union. *Int. J. Perform. Anal. Sport* **2014**, *14*, 955–967. [CrossRef]

18. Sabio Lago, Y.S.; Balic, M.G.; Sanromà, J.C.; Solá-Santesmases, J.; Argudo Iturriaga, F. Design, Validation and Reliability of an Instrument to Analyze Technical-Tactical Actions in Water Polo. *Retos Nuevas Perspect. Educ. Física Deporte Recreación* **2018**, *34*, 57–65.

19. Nadeau, L.; Richard, J.-F.; Godbout, P. The Validity and Reliability of a Performance Assessment Procedure in Ice Hockey. *Phys. Educ. Sport Pedagog.* **2008**, *13*, 65–83. [CrossRef]

20. Prieto, J.; Gómez-Ruano, M.A.; Sampaio, J. From a Static to a Dynamic Perspective in Handball Match Anal-Ysis: A Systematic Review. *Open Sports Med. J.* **2015**, *8*, 25–34.

21. Lozano Jarque, D.; Camerino Foguet, O. Eficacia de Los Sistemas Ofensivos En Balonmano. *Apunt. Educ. Física Deportes* **2012**, *2*, 70–81. [CrossRef]

22. Lozano, D.; Camerino, O.; Hileno, R. Análisis Del Comportamiento Táctico Ofensivo En Momentos Críticos de Juego En El Alto Rendimiento En Balonmano: Un Estudio Mixed Methods. *Cuad. Psicol. Deporte* **2016**, *16*, 151–160.

23. Jiménez-Salas, J.; Hernández-Mendo, A. Analysis of Data Quality and Generalizability Observing System Counter in Elite Handball. *Rev. Cienc. Deporte* **2016**, *12*, 35–44.

24. González Ramírez, A. Fiabilidad de Un Instrumento de Observación Para Analizar El Rendimiento En Handball. *Cienc. Psicológicas* **2016**, *10*, 199–207. [CrossRef]

25. Pic, M. The Observation of Gender Differences in Handball. *Rev. Cienc. Deporte* **2017**, *13*, 191–198.

26. Tulha Duarte Ferreira, A.L.; Santos Graça, A.B.; Dias Estriga, M.L. Development and Validation of an Individual Technical-Tactical Assessment System at Handball: The Handball Tactical Performance Evaluation (HTPE). *J. Sports Sci.* **2018**, *14*, 141–154.

27. Quiñones Rodríguez, Y.; Morillo-Baro, J.P.; Reigal, R.E.; Morales-Sánchez, V.; Vázquez-Diz, J.A.; Hernández-Mendo, A. El Ataque Posicional En Balonmano: Validación de Un Sistema de Observación. *Cuad. Psicol. Deporte* **2019**, *19*, 113–124. [CrossRef]

28. Ferrari, W.R.; Vaz, V.; Sousa, T.; Sarmento, H.; Dias, G. Development and Validation of a Notational Instrument to Study the Offensive Process in Handball. *J. Sport Pedagog. Res.* **2018**, *4*, 27–34.

29. Montoya Fernández, M.; Moras i Feliu, G.; Anguera i Argilaga, M.T. Análisis de Las Finalizaciones de Los Extremos En Balonmano. *Apunt. Educ. Física Deportes* **2013**, *113*, 52–59. [CrossRef]

30. Bilge, M. Game Analysis of Olympic, World and European Championships in Men's Handball. *J. Hum. Kinet.* **2012**, *35*, 109–118. [CrossRef] [PubMed]

31. Gutiérrez, Ó.; Ruiz, J.L. Data Envelopment Analysis and Cross-Efficiency Evaluation in the Management of Sports Teams: The Assessment of Game Performance of Players in the Spanish Handball League. *J. Sport Manag.* **2013**, *27*, 217–229. [CrossRef]

32. Hassan, A. Team Handball World Cup Championship 2013—Analysis Study. *J. Hum. Sport Exerc.* **2014**, *9*, S409–S416. [CrossRef]

33. Santos, F.M.; Fernandez, J.; Oliveira, M.C.; Leitão, C.; Anguera, T.; Campaniço, J. The Pivot Player in Handball and Patterns Detection—Instrument. *Motricidade* **2009**, *5*, 29–36. [CrossRef]

34. Meletakos, P.; Vagenas, G.; Bayios, I. A Multivariate Assessment of Offensive Performance Indicators in Men's Handball: Trends and Differences in the World Championships. *Int. J. Perform. Anal. Sport* **2011**, *11*, 284–294. [CrossRef]

35. Ávila-Moreno, F.; Lozano-Jarque, D.; Chirosa-Ríos, L.J.; Ortega-Becerra, M. Design, Validation, and Reliability of an Observational Instrument Assessing Pivot Influence on Tactical Effectiveness in Team-Handball. *Kinesiology* **2022**, *54*, 166–174. [CrossRef]

36. International Handball Federation (IHF). *Rules of the Game*; Indoor Handball: Copenhagen, Denmark, 2016.

37. Prudente, J.N.; Cardoso, A.R.; Rodrigues, A.J.; Mendes, J.N.; Fernando, C.; Lopes, H.; Sousa, D.F. Playing 7 vs. 6 with an Empty Goal: Is It Really an Option for Coaches? A Comparative Analysis between Portugal and the Other Teams during the Men's European Handball Championship 2020. *Front. Psychol.* **2022**, *13*, 809909. [CrossRef]

38. Vázquez-Diz, J.A.; Morillo-Baro, J.P.; Reigal, R.E.; Morales-Sánchez, V.; Hernández-Mendo, A. Design and Validation of an Observational Instrument for Goalkeepers in Beach Handball Elaboração e Validação de Um Instrumento de Observação Para Guarda Redes de Andebol de Praia. *Cuad. Psicol. Deporte* **2019**, *19*, 135–146.

39. Navarro, A.; Morillo, J.P.; Reigal, R.E.; Hernández-Mendo, A. Polar Coordinate Analysis in the Study of Positional Attacks in Beach Handball. *Int. J. Perform. Anal. Sport* **2018**, *18*, 151–167. [CrossRef]

40. Font, R.; Daza, G.; Irurtia, A.; Tremps, V.; Cadens, M.; Mesas, J.A.; Iglesias, X. Analysis of the Variables Influencing Success in Elite Handball with Polar Coordinates. *Sustainability* **2022**, *14*, 15542. [CrossRef]

41. Morillo, J.P.; Reigal, R.E.; Hernández-Mendo, A.; Montaña, A.; Morales-Sánchez, V. Decision-Making by Handball Referees: Design of an Ad Hoc Observation Instrument and Polar Coordinate Analysis. *Front. Psychol.* **2017**, *8*, 1842. [CrossRef]

42. Host, K.; Pobar, M.; Ivasic-Kos, M. Analysis of Movement and Activities of Handball Players Using Deep Neural Networks. *J. Imaging* **2023**, *9*, 80. [CrossRef] [PubMed]

43. Ding, P. Analysis of Artificial Intelligence (AI) Application in Sports. *J. Phys. Conf. Ser.* **2019**, *1302*, 032044. [CrossRef]

44. Saavedra, J.M.; Þorgeirsson, S.; Chang, M.; Kristjánsdóttir, H.; García-Hermoso, A. Discriminatory Power of Women's Handball Game-Related Statistics at the Olympic Games (2004–2016). *J. Hum. Kinet.* **2018**, *62*, 221–229. [CrossRef]

45. Vuleta, D.; Sporiš, G.; Milanović, D. Indicators of Situational Efficiency of Winning and Defeated Male Handball Teams in Matches of the Olympic Tournament 2012. *Acta Kinesiol.* **2015**, *9*, 40–49.

46. Milanović, D.; Vuleta, D.; Ohnjec, K. Performance Indicators of Winning and Defeated Female Handball Teams in Matches of the 2012 Olympic Games Tournament. *J. Hum, Kinet.* **2018**, *64*, 247–253. [CrossRef]

47. Almeida, A.G.; Merlin, M.; Pinto, A.; Torres, R.D.S.; Cunha, S.A. Performance-Level Indicators of Male Elite Handball Teams. *Int. J. Perform. Anal. Sport* **2020**, *20*, 1–9. [CrossRef]

48. de Paula, L.V.; Costa, F.E.; Ferreira, R.M.; Menezes, R.P.; Werneck, F.Z.; Coelho, E.F.; Greco, P.J. Analysis of Discriminatory Game Variables between Winners and Losers in Women's Handball World Championships from 2007 to 2017. *Kinesiology* **2020**, *52*, 54–63. [CrossRef]

49. Konstantinos, N.S.; Elissavet, R.N.; Panagiotis, G.M.; Ioannis, B.A.; Konstantinos, B.D. Performance Indicators and Competition Ranking in Women's and Men's World. *J. Phys. Educ. Sport* **2018**, *18*, 1761–1766.

50. Daza, G.; Andrés, A.; Tarragó, R. Match Statistics as Predictors of Team's Performance in Elite Competitive Handball. *RICYDE Rev. Int. Cienc. Deporte* **2017**, *48*, 149–161. [CrossRef]

51. Meletakos, G.P.; Konstantinos, N.S.; Ioannis, B.A. Stable and Changing Characteristics of High-Level Handball as Evidenced from World Men's Championships. *J. Phys. Educ. Sport* **2020**, *20*, 1354–1361.

52. Skarbalius, A.; Pukenas, K.; Vidünaite, G. Sport Performance Profile in Men's European Modern Handball: Discriminant Analysis between Winners and Losers. *Balt. J. Health Phys. Act.* **2013**, *3*, 44–54.

53. Trejo-Silva, A.; Feu, S.; Camacho-Cardenosa, A.; Camacho-Cardenosa, M.; Brazo-Sayavera, J. Relation of Offensive Performance during Exclusions and Final Ranking in Female Handball. *Appl. Sci.* **2022**, *12*, 10774. [CrossRef]

54. Ferrari, W.; Dias, G.; Sousa, T.; Sarmento, H.; Vaz, V. Comparative Analysis of the Offensive Effectiveness in Winner and Losing Handball Teams. *Front. Psychol.* **2020**, *11*, 547110. [CrossRef]

55. Vaz, V.; Ferrari, W.; Sarmento, H.; Couceiro, M.; Sousa, T.; Dias, G. Handball Offensive Analysis: Comparative Evolution of Linear Tendency Lines between Finalist Teams in the EHF CL. *Appl. Sci.* **2023**, *13*, 1366. [CrossRef]

56. Ohnjec, K.; Vuleta, D.; Milanović, D.; Gruić, I. Performance Indicators of Teams at the 2003 World Handball Championship for Women in Croatia. *Kinesiology* **2008**, *40*, 69–79.

57. Gómez, M.Á.; Lago-Peñas, C.; Viaño, J.; González-Garcia, I. Effects of Game Location, Team Quality and Final Outcome on Game-Related Statistics in Professional Handball Close Games. *Kinesiology* **2014**, *46*, 249–257.

58. Þorgeirsson, S.; Pic, M.; Lozano, D.; Sigurgeirsson, O.; Sekulic, D.; Saavedra, J.M. Gender-Based Differences in Game-Related Statistics between Winning and Losing Teams in an Amateur Handball League. *Acta Gymnica* **2022**, *52*, e2022. [CrossRef]

59. Þorgeirsson, S.; Pic, M.; Lozano, D.; Sigurgeirsson, O.; Sekulic, D.; Saavedra, J.M. The Difference Between Winners and Losers in Balanced Handball Games in the Final 10 Minutes. *Montenegrin J. Sports Sci. Med.* **2022**, *11*, 37–43. [CrossRef]

60. Antúnez-Medina, A.; Garcías-Rubia, J.; Sáez-Blázquez, F.J.; Valle-Blanco, A.; García-Martín, A. Differences in Performance Indicator between Winning and Losing Team in Formative Stages According to Sex and Final Score Differences. *Rev. Cienc. Deporte* **2013**, *9*, 5–16.

61. Hatzimanouil, D. Goalkeeper's Efficiency in Relation with Throws from Different Attacking Court Areas in Team Handball. *J. Phys. Educ. Sports Manag.* **2020**, *7*, 11–18. [CrossRef]

62. Gümüş, H.; Gençoğlu, C. The Effects of the Goalkeeper Substitution Rule as a New Strategy in Handball: Analysis of Men's European Handball Championship 2020. *Acta Gymnica* **2020**, *50*, 113–121. [CrossRef]

63. Costa, G.D.C.T.; Pedrosa, G.F.; Souza, N.P.D.; Gemente, F.R.F.; Freire, A.B.; Castro, H.D.O. Type of Game Practiced in Handball According to the Positions of the Attackers: Analysis of the Women's World Handball Championship 2015. *Int. J. Perform. Anal. Sport* **2017**, *17*, 360–373. [CrossRef]

64. Wagner, H.; Hinz, M.; Melcher, K.; Radic, V.; Uhrmeister, J. The PlayerScore: A Systematic Game Observation Tool to Determine Individual Player Performance in Team Handball Competition. *Appl. Sci.* **2023**, *13*, 2327. [CrossRef]

65. Pascual, A.; Font, R.; Pascual, X.; Lago-Peñas, C. Evolution of Match Performance Parameters in Elite Men's Handball 2012–2022. *Int. J. Sports Sci. Coach.* **2023**, 174795412211424. [CrossRef]

66. Lago-Peñas, C.; Gómez, A.M.; Viaño, J.; González-García, I.; Fernández-Villarino, M. de L. ángeles Home Advantage in Elite Handball: The Impact of the Quality of Opposition on Team Performance. *Int. J. Perform. Anal. Sport* **2013**, *13*, 724–733. [CrossRef]

67. Prieto, J.; Gómez, M.-Á.; Volossovitch, A.; Sampaio, J. Effects of Team Timeouts on the Teams' Scoring Performance in Elite Handball Close Games. *Kinesiology* **2016**, *48*, 115–123. [CrossRef]

68. Araújo, D.; Couceiro, M.; Seifert, L.; Sarmento, H.; Davids, K. *Artificial Intelligence in Sport Performance Analysis*; Routledge: New York, NY, USA, 2021; ISBN 9781003163589.

69. Hughes, M.D.; Bartlett, R.M. The Use of Performance Indicators in Performance Analysis. *J. Sports Sci.* **2002**, *20*, 739–754. [CrossRef]

70. Bulger, S.M.; Housner, L.D. Modified Delphi Investigation of Exercise Science in Physical Education Teacher Education. *J. Teach. Phys. Educ.* **2007**, *26*, 57–80. [CrossRef]

71. Cohen, J. *Statistical Power Analysis for the Behavioral Science*, 2nd ed.; Lawrence Erlbaum Associates: New York, NY, USA, 1988; ISBN 0-8058-0283-5.

72. Anguera, M.T.; Camerino, O.; Castañer, M.; Sánchez-Algarra, P.; Onwuegbuzie, A.J. The Specificity of Observational Studies in Physical Activity and Sports Sciences: Moving Forward in Mixed Methods Research and Proposals for Achieving Quantitative and Qualitative Symmetry. *Front. Psychol.* **2017**, *8*, 2196. [CrossRef]

73. Blanco Villaseñor, A.; Losada, J.L.; Anguera, T. Data Analysis Techniques in Observational Designs Applied to the Environment-Behaviour Relation1 Analysis of Tactical-Tecnhical Performance in Team Handaball View Project. *Medio Ambiente Comport. Hum.* **2003**, *4*, 111–126.

74. Heinemann, K. *Introducción a la Metodología de la Investigación Empírica en Las Ciencias Del Deporte*; Paidotribo: Barcelona, Spain, 2003.

75. Fortes, A.M.; Gómez, M.A.; Hongyou, L.; Sampedro, J. Inter-Operator Reliability of Videoobserver^{TM}. *Cuad. Psicol. Deporte* **2016**, *16*, 137–152.

76. O'Donoghue, P.; Holmes, L. *Data Analysis in Sport*; Routledge: London, UK, 2015.

77. Altman, D.G. *Practical Statistics for Medical Research*; Chapman & Hall: London, UK, 1991.

78. Peterson, R.A.; Kim, Y. On the Relationship between Coefficient Alpha and Composite Reliability. *J. Appl. Psychol.* **2013**, *98*, 194–198. [CrossRef]

79. George, D.; Mallery, P. *SPSS for Windows Step by Step: A Simple Guide and Reference (11.0 Update)*, 4th ed.; Allyn & Bacon: Boston, MA, USA, 2003.

80. Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [CrossRef]

81. García-Martín, A.; Antúnez, A.; Ibáñez, S.J. Análisis Del Proceso Formativo En Jugadores Expertos: Validación de Instrumento/Analysis of Expert Players' Training Process: Validation of Tools. *Rev. Int. Med. Cienc. Act. Física Deporte* **2016**, *61*. [CrossRef]

82. Chacón-Moscoso, S.; Sanduvete-Chaves, S.; Anguera, M.T.; Losada, J.L.; Portell, M.; Lozano-Lozano, J.A. Preliminary Checklist for Reporting Observational Studies in Sports Areas: Content Validity. *Front. Psychol.* **2018**, *9*, 291. [CrossRef] [PubMed]

83. Hernández-Mendo, A.; López, J.A.; Castellano, J.; Morales, V.; Pastrana, J.L. Program for Use in Methodology Observacional. *Cuad. Psicol. Deporte* **2012**, *12*, 55–78.

84. Gabin, B.; Camerino, O.; Anguera, M.T.; Castañer, M. Lince: Multiplatform Sport Analysis Software. *Procedia Soc. Behav. Sci.* **2012**, *46*, 4692–4694. [CrossRef]