anales de psicología / annals of psychology
2026, vol. 42, nº 1 (january-april), e09, 1-15
https://doi.org/10.6018/analesps.651811

Published by Editum, Servicio de Publicaciones, Universidad de Murcia (Spain), in https://revistas.um.es/analesps
ISSN online: 1695-2294. License Creative Commos 4.0 BY

Methodological quality of football studies based on observational methodology: A systematic review

Daniel López-Arenas[1], Susana Sanduvete-Chaves[1,*], M. Teresa Anguera[2], and Salvador Chacón-Moscoso[1,3]

[1] *Universidad de Sevilla, Spain*
[2] *Universidad de Barcelona, Spain*
[3] *Universidad Autónoma de Chile, Chile*

Título: Calidad metodológica de estudios de fútbol basados en metodología observacional: una revisión sistemática.

Resumen: La metodología observacional (MO) se ha empleado ampliamente en fútbol. Este estudio presenta una revisión sistemática que aplica la Escala de Calidad Metodológica para estudios basados en Metodología Observacional (MQSOM) a 236 estudios sobre fútbol basados en MO para desvelar tendencias metodológicas. Las propiedades psicométricas de la MQSOM fueron adecuadas. Se llevó a cabo un análisis descriptivo, un análisis de proporciones para identificar las diferencias procedimentales y un análisis de clúster bietápico para identificar perfiles de calidad metodológica (CM). Se encontraron diferencias en el tipo de observación (frecuentemente directa), el manual de codificación (frecuentemente especificado), el tipo de datos (raramente especificado), el instrumento de observación (frecuentemente especificado), el registro (frecuentemente especificado), control (ocasionalmente especificado) y software de análisis (frecuentemente especificado), tipo de parámetro (generalmente frecuencia), control de calidad (frecuentemente empleado) y análisis (generalmente no descriptivo). Se obtuvieron cinco perfiles de CM en orden decreciente de puntuación en MQSOM (P1: Calidad de diseño -CD-, y medición y análisis -CMA- altas; P2: CD moderada, CMA alta; P3: CD baja, CMA alta; P4: CD baja, CMA moderada; y P5: CD y CMA bajas). Este trabajo presenta una instantánea actualizada de la MO aplicada en el fútbol, esbozando futuras mejoras en la CM.

Palabras clave: Escala de Calidad Metodológica para estudios basados en Metodología Observacional (MQSOM). Perfiles de calidad metodológica. Revisión sistemática. Fútbol.

Abstract: Observational methodology (OM) has been extensively employed in football research. This study presents a systematic review that applies the Methodological Quality Scale for studies based on Observational Methodology (MQSOM) to 236 primary papers on football based on OM to unveil methodological trends. The psychometric properties of the MQSOM yielded adequate results. Descriptive statistics were provided; a proportion analysis was conducted to identify differences in procedures; and a two-step cluster analysis was conducted to identify methodological quality (MQ) profiles. The proportion analysis revealed significant differences in observation type (frequently direct), coding manual (often specified), data type (seldom specified), observational instrument (frequently specified), recording (often specified), control (occasionally specified) and analysis software (frequently specified), parameter type (usually frequency), quality control (frequently employed), and analysis (usually non-descriptive). Two-step cluster analysis yielded five MQ profiles ranked in decreasing order based on the MQSOM score (P1: high Design Quality -DQ-, and Measurement & Analysis Quality -MAQ-; P2: moderate DQ, high MAQ; P3: low DQ, high MAQ; P4: low DQ, moderate MAQ; and P5: low DQ and MAQ). This work presents an updated snapshot of OM applied in football research, outlining future improvements in MQ.

Keywords: Methodological Quality Scale for studies based on Observational Methodology (MQSOM). Methodological quality profiles. Systematic review. Football.

Introduction

Observational methodology is a research design that enables the rigorous recording, quantification and analysis of spontaneous behaviors displayed in natural contexts (Anguera et al., 2020). This methodology is widely employed due to its low level of intervention and the utilization of ad hoc instruments to observe the research object (Anguera, Blanco-Villaseñor et al., 2018). It is crucial to distinguish between observational methodology, which is based on systematic observation in everyday contexts, and an observational study in health research. An observational study is a type of quantitative research that uses empirical group comparisons to identify potential causal relationships when randomization and strict experimental control are not feasible. This includes cohort, case-control and cross-sectional studies (Cochran & Chambers, 1965).

*** Correspondence address [Dirección para correspondencia]:**
Susana Sanduvete-Chaves. Universidad de Sevilla, Facultad de Psicología, Dpto. de Psicología Experimental. Campus Ramón y Cajal. C/ Camilo José Cela s/n. 41018, Sevilla (Spain).
E-mail: sussancha@us.es
(Article received: 26-02-2025; revised: 03-06-2025; accepted: 4-06-2025)

Observational methodology is considered a mixed methods approach (Anguera et al., 2017), as it integrates both qualitative and quantitative elements through a conversion or transformation of one data type into the other (Cresswell & Plano-Clark, 2017). Observational designs can be framed in a QUAL-QUANT-QUAL sequence. In the initial QUAL (qualitative) phase, the design and observation instrument are selected, and the recording is done. In the subsequent QUANT (quantitative) phase, the recorded data is transformed into a code matrix and subjected to data quality control procedures prior to the data analysis. In the final QUAL phase, the results are interpreted considering both the initial topic and the literature (Anguera et al., 2020).

The application of observational methodology to sports research can be traced back to the latter part of the previous century. Since then, observational methodology has been progressively structured on solid methodological foundations in terms of observational design (Anguera et al., 2011) and record systematization and coding (Anguera & Blanco-Villaseñor, 2005). In recent years, efforts have been made to consolidate both direct (Anguera, 2003; Anguera et al., 2017; Sánchez-Algarra & Anguera, 2013) and indirect observation (Anguera, Portell et al., 2018).

The past decade has seen notable innovations in the assessment of methodological quality in observational methodology studies. One noteworthy contribution is the Guidelines for Reporting Evaluations Based on Observational Methodology (GREOM) (Portell et al., 2015), which provides 14 straightforward standards for conducting and reporting observational methodology studies. Based on the GREOM, a checklist of the methodological features to be considered in observational studies comprised of 16 items was developed: the Methodological Quality Checklist for Studies Based on Observational Methodology (MQCOM) (Chacón-Moscoso et al., 2019). The MQCOM's content validity and intercoder reliability has been demonstrated with empirical evidence. Recently, an instrument drawing on the MQCOM, the Methodological Quality Scale for Observational Methodology (MQSOM) (Sanduvete-Chaves et al., 2025) has also been validated. The MQSOM represents a significant milestone in the development of a methodological culture based on systematic observation. It is currently the only tool available with adequate psychometric properties to assess the methodological quality of studies based on observational methodology. Specifically, adequate validity based on the scale's internal structure and reliability has been demonstrated for the 11-item scale with a second-order factor structure.

In the design, measurement and analysis of a program based on observational methodology, the decisions are made within a methodological quality continuum. Recently, systematic reviews have been used to try to synthesize evidence in football literature based on observational methodology (Anguera, 2023). In these innovative methodological studies, the GREOM (Portell et al., 2015) was applied to primary football studies based on observational methodology in order to reveal latent quality profiles in the literature (Preciado et al., 2019, 2021). Preciado et al. (2019, 2021) examine the relationship between how research is reported and its underlying methods, yielding valuable insights. For example, critical elements such as the observational design or qualitative data recorded are often overlooked entirely in such studies.

Building on these remarkable endeavors, observational methodology appears to present distinctive challenges, since the viability of this mixed methods approach is inherently linked to the quantitizing of a specific type of qualitative data from a particular type of observational design into a certain type of quantitative parameter (Anguera et al., 2020). Consequently, the frequent disregard of observational design and qualitative data type in reports suggests that reporting quality does not necessarily correspond to methodological quality in the observational methodology just as, inversely, poor reporting does not seem to be inherently linked to poor methods (Huwiler-Müntener et al., 2002).

Therefore, the main objective of the systematic review conducted in this study was to use the quality criteria established in the MQSOM to further explore the methodological characteristics of observational methodology studies in football reported in the literature. The specific aims were: a) to apply the MQSOM to obtain descriptive statistics and assess the main characteristics of these studies, b) to obtain empirical evidence of significant differences between the procedures reported, and c) to detect different methodological quality profiles among studies.

Methodology

Selection criteria of the studies

In order to select the papers, a title, abstract, keyword and full text search was done for the terms "observational methodology" AND [football OR soccer OR futsal] in PsycInfo, SCOPUS, Web of Science, SPORTDiscus, PSICODOC and Google Scholar, with December 31, 2022, as the cut-off date. The reference lists from the articles collected in this step, along with that of Preciado et al. (2019, 2021), were then examined to identify additional studies.

Following PRISMA recommendations (Page et al., 2021), the inclusion criteria for the units of analysis (papers) in this study were as follow: all publications a) relied on observational methodology; b) were applied in football; c) were empirical; d) presented the standard structure of introduction, method, results and discussion; and e) were written in English or Spanish.

Search procedure

Mendeley Reference Manager was used to organize and handle the information obtained in the literature search. During the first screening, the inclusion criteria were applied to the title, keywords and abstract. The resulting papers were assessed in a second stage, in which the inclusion criteria were applied to the full texts. Two coders (DLA and IFM) applied the criteria independently. In case of disagreements, a third coder (SCM) mediated.

Coding of variables

For the data extraction, two coders (DLA and IFM) were trained to apply the MQSOM. Each item and its response options were explained. A third researcher (MTA) mediated when differences arose. Both coders then independently applied the scale to two observational methodology studies to compare the coding and, in case of disagreements, a third coder (SSC) mediated until a consensus was reached. After the training, the coders applied the scale independently to a randomly selected 25% of the sample. Finally, once a high level of agreement was achieved (an Intraclass Correlation Coefficient (ICC) > .7), DLA applied the scale to the full sample.

The primary studies included were assessed using MQSOM (Sanduvete-Chaves et al., 2025), a scale with adequate psychometric properties (RMSEA = .000, NNFI = 1, GFI = .98, AGFI = .97). The MQSOM comprises a second-order factor of Methodological Quality (ω = .87; D = .55)

containing two first-order factors: Quality of Design (6 items; ω = .90; D = .46; ICC = .933 - .967) and Quality of Measurement and Analysis (5 items; ω = .68; D = .67; ICC = .797 - .988). Each item receives a score for methodological quality levels from 0 (lowest) to 1 (highest). The MQSOM codification manual is available at https://osf.io/ca2v8/files/osfstorage. The primary studies were categorized based on the following quality criteria (Sanduvete-Chaves et al., 2025):

Design Quality I: Direct/Indirect Observation. Direct observation involves systematically recording and coding observation sessions (Sánchez-Algarra & Anguera, 2013), while indirect observation involves obtaining and analyzing data from documents, tables and databases (Anguera, 2021).

Design Quality II: Observation Unit. Observational designs can be classified as idiographic or nomothetic. Idiographic designs focus on a single participant (a player) or a unit of individuals (a football team), while nomothetic designs focused on two or more participants (players) or units (football teams) (Anguera et al., 2011).

Design Quality III: Temporal Unit. Observational designs can be cross-sectional (punctual), where data are collected in a single session, or longitudinal (with follow-up), where several sessions are held over a period of time (Anguera et al., 2011).

Design Quality IV: Dimensional Criteria. Observational designs can be classified as unidimensional, where only one level of response is considered, or multidimensional, where several levels of response are considered (Anguera et al., 2011).

Design Quality V: Coding Manual. This indicates whether the coding manual is specified and, if so, its type: rating scale, category system, field format or a combination of these (Anguera et al., 2007).

Design Quality VI: Data Type. The nature of the registered data is classified as one of four types in ascending order of complexity (Bakeman, 1978): Type I for sequential events, Type II for concurrent events, Type III for sequential states, or Type IV for concurrent states.

Measurement and Analysis Quality I: Observation Instrument Adequacy. Indicates, when specified, whether the coding manual was built ad hoc, adapted from other studies, and/or previously validated (Sanduvete-Chaves et al., 2025).

Measurement and Analysis Quality II: Recording, Control and Analysis Software. Indicates whether recording, data quality control, and/or analysis software is used. If so, it specifies what software is used (Sanduvete-Chaves et al., 2025).

Measurement and Analysis Quality III: Type of Parameter. The parameter can be classified in ascending complexity order as frequency, order or duration (Anguera et al., 2017).

Measurement and Analysis Quality IV: Data Quality Control. Indicates whether data quality is analyzed, specifying the type of analysis (qualitative or quantitative), and the specific procedure used (Blanco-Villaseñor et al., 2001).

Measurement and Analysis Quality V: Type of Data Analysis. Indicates whether data has been analyzed and distinguishes between qualitative analysis, descriptive analysis, inferential analysis or regularity detection (Sanduvete-Chaves et al., 2025).

Data analysis

Psychometric properties were studied using JASP version 0.16. The ICC was calculated to assess both inter- and intra-coder reliability. Values over .7 were considered indicators of adequate concordance (Portney & Walkins, 2000). In order to confirm the MQSOM's reliability in relation to the sample, McDonald's Omega (*ω*) was calculated, with results higher than .80 considered strong reliability and .65 – .80, acceptable (Kalkbrenner, 2023). For item discrimination, the mean item discrimination indexes were computed. Results were interpreted as excellent for values higher than .40, good for values between .30 – .40, adequate for .20 – .30, and inadequate for < .20 (Holgado-Tello, 2015). Additionally, descriptive statistics were provided for the MQSOM items, publication year, author affiliation country, journal, event observed, object of research and study settings.

A proportion analysis was then conducted using the binomial non-parametric test in SPSS 29.0 to obtain empirical evidence of significant differences between research procedures derived from MQSOM quality criteria. Ratios over the null hypothesis of .45-.55 (95% CI) were considered indicators of statistically different proportions (Hollander et al., 2013).

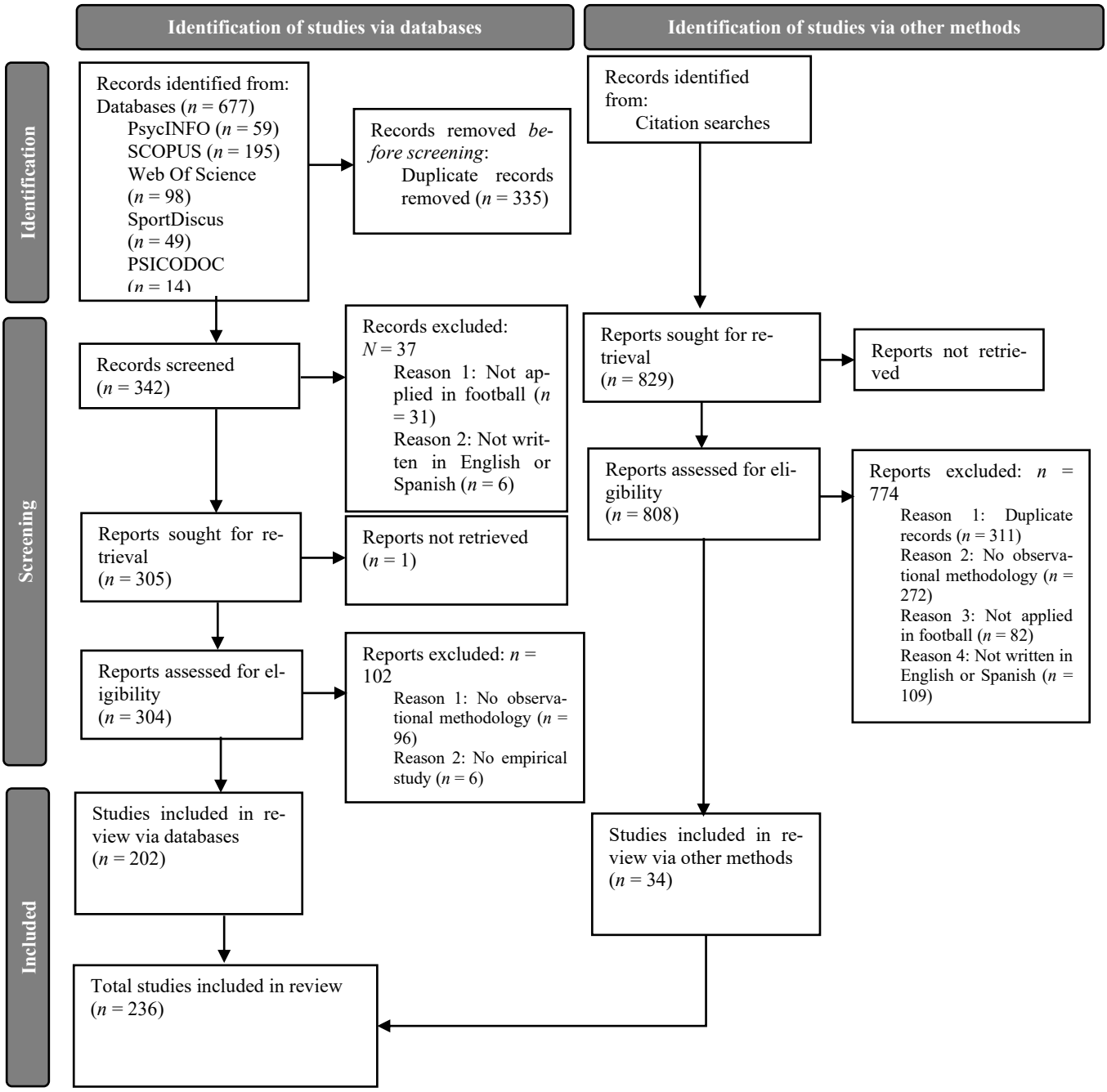
Finally, a two-step cluster analysis was conducted in SPSS 29.0 to group studies by methodological quality profiles based on their research procedures. This grouping allowed both categorical and continuous data to be analyzed simultaneously through a log-likelihood algorithm, thereby facilitating the identification of natural groupings within a dataset that might otherwise remain unidentified. Although the algorithm assumes a normal distribution for continuous variables, it is robust when the assumption is not met (IBM, 2025). The two-step clustering procedure is comprised of a pre-clustering stage in which the original cases are clustered in order to reduce the dimensions of the matrix containing distances between all potential pairs of cases. The cases represent the objects to be clustered, whereas the variables represent the attributes the cluster analysis is based on. In this procedure, the log-likelihood algorithm randomly assigns an observed case to a given cluster. As the case is read, the algorithm determines whether it should be merged with one of the existing clusters or if a new cluster should be created. Once pre-clustering is completed, all cases within the same cluster are treated as a single entity (Norusis, 2012). For dimension reduction, principal component analysis (PCA) was used and clusters were determined with Schwarz's Bayesian Information Criterion (BIC) (Chiu et al., 2001). Cluster models with both low levels of BIC and high ratios of distance measures were deemed adequate (Schwarz, 1978).

Results

Selection of studies

Figure 1 summarizes the selection process of the papers for this study. The final sample used for the review consisted of 236 primary papers. The list of articles included is available at https://osf.io/ca2v8/files/osfstorage.

Figure 1
PRISMA 2020 flow diagram for systematic review (Page et al., 2021).



Psychometric properties and descriptive statistics

The data extraction database is available at https://osf.io/ca2v8/files/osfstorage. Table 1 presents the descriptive statistics of the items from MQSOM. ICC coefficients were adequate, ranging from .71 to 1. Reliability yielded strong to acceptable evidence, ω = .89; ω_1 = .91; ω_2 = .70, and discrimination was excellent, D = .61; D_1 = .77; D_2 = .41. In 69.07% of the sample (163 studies), the benchmark for a high methodological quality overall was not met (≥ 0.75); 74.15% (175 studies) scored below the benchmark for design quality, and 41.95% (99 studies) fell below the benchmark for measurement and analysis dimension.

The means were between .13 and .96, with standard deviations ranging between .15 and .45, and no normal distribution of the items. The median was 0 for four out of six (66.7%) of the design quality items (items 1 to 6), and 1 for four out of five (80%) of the measurement and analysis items (items 7 to 11).

Table 1
Descriptive statistics and inter-intracoder reliability of the items from MQSOM.

Item	Intercoder Reliability			Intracoder Reliability			Descriptive Statistics					
	ICC	*LL*	*UL*	*ICC*	*LL*	*UL*	*M*	*Mdn*	*SD*	*S*	*K*	*KS*
1	.771	.643	.857	.741	.381	.908	.72	1	.42	-0.97	-0.87	.41
2	.902	.841	.941	1	1	1	.40	0	.44	0.42	-1.60	.33
3	.902	.841	.961	1	1	1	.40	0	.44	0.42	-1.59	.33
4	.865	.784	.918	1	1	1	.40	0	.45	0.41	-1.64	.34
5	.851	.763	.909	.865	.641	.954	.13	0	.32	2.28	3.35	.52
6	.625	.442	.758	1	1	1	.65	.5	.31	-0.29	-0.63	.30
7	.851	.763	.909	1	1	1	.71	1	.43	-0.92	-0.98	.41
8	.609	.421	.747	.772	.441	.920	.71	1	.37	-0.89	-0.67	.36
9	.712	.561	.818	.860	.631	.952	.61	.5	.25	0.32	-0.01	.41
10	.816	.709	.886	1	1	1	.79	1	.40	-1.47	0.18	.48
11	.897	.833	.938	.772	.441	.920	.96	1	.15	-4.74	24.88	.52

Note. ICC = intraclass correlation; *S* = skewness; *K* = kurtosis; *KS* = Kolmogorov-Smirnov normality test. *ICC* and *KS* obtained $p < .05$.

Table 2 presents the descriptive statistics by author affiliation country, with 64.8% of the sample (153 papers) originating in Spain.

Table 2
Paper distribution by author affiliation country.

Country	*f*	%
Spain	153	64.8
Portugal	16	8.5
UK	11	4.7
Italy	8	3.4
Brazil	8	3.4
Colombia	4	1.7
Germany	4	1.7
Greece	4	1.7
Uruguay	3	1.3
Norway	3	1.3
Poland	3	1.3
Chile	2	0.9
France	2	0.9
Turkey	2	0.9
Austria	1	0.4
Cuba	1	0.4
Iran	1	0.4
Australia	1	0.4
Romania	1	0.4
Slovakia	1	0.4
Sweden	1	0.4
USA	1	0.4
Japan	1	0.4
Total	196	100

Table 3 presents the descriptive statistics by publication year. More than one-third (78) of the primary works included in this study were published in or after 2019, and the oldest one dates to 1996.

Table 3
Paper distribution by publication year.

Year	*f*	%	c%
2022	21	8.9	8.9
2021	19	8.1	17
2020	19	8.1	25.1
2019	19	8.1	33.2
2018	20	8.5	41.7
2017	11	4.7	46.4
2016	12	5.1	51.5
2015	15	6.4	57.9
2014	17	7.2	65.1
2013	14	5.9	71
2012	10	4.2	75.2
2011	8	3.4	78.6
2010	11	4.7	83.3
2009	9	3.8	87.1
2008	4	1.7	88.8
2007	5	2	90.8
2006	4	1.7	92.5
2005	4	1.7	94.2
2004	2	0.8	95
2003	2	0.8	95.8
2002	3	1.3	97.1
2001	3	1.3	98.4
2000	2	0.8	99.2
1998	1	0.4	99.6
1996	1	0.4	100
Total	236	100	

Note. c% = cumulative percentage.

Table 4 summarizes the distribution of the primary studies by journal. The primary studies included in this review were published in 69 journals, 40 of which had an impact factor -IF- in Web of Science (163 papers, 69.1 %); 20% were in quartile (Q) 1, 35% in Q2, 22.5% in Q3, and 22.5% in Q4; the most common category was Sport Science (45%). The journals where these primary studies appeared with the highest frequency were *Frontiers in Psychology* (20; IF = 2.6), *Apunts* (19; IF = 1.6), and *Cuadernos de Psicología del Deporte* (18; no IF).

Table 4
Paper distribution by journal.

Journal	Indexing	IF	Quartile	Category	*f*	%
Frontiers in Psychology	PubMed, PMC, Scopus, Google Scholar, DOAJ, CrossRef, PsycInfo, Semantic Scholar, Ulrich's Periodicals Directory, CLOCKSS, SSCI, EBSCO, OpenAIRE, & Zetoc	2.6	Q2	Psychology, Multidisciplinary	20	8.47
Apunts Educación Física y Deportes	ABI/INFORM, Academic Research Library, Academic Search Complete, CBUC, CNKI Scholar, Compludoc, DICE, EBSCO, Heracles, IRESIE, ISOC, OpenAlex, Periodical Index Online, ProQuest, Research Library, RESH, & SPORTDiscus	1.6	Q2	Education & Educational Research	19	8.05
Cuadernos de Psicología del Deporte	Scopus, SPORTDiscus, EBSCO, In-Recs, Latindex, DICE, ISOC, Dialnet, Resh Oceano Psicodoc, IBECS, PsycInfo, & Redalyc	No	-	-	18	7.63
Journal of Sports Sciences	Adis International, British Library Board, CABI, EBSCO, Elsevier, Focus On, Sports Science and Medicine, Gale, Ovid, ProQuest, Taylor & Francis Clarivate Analytics, & MEDLINE	2.3	Q2	Sport Science	16	6.78
Retos	Scopus, SPORTDiscus, e-revistas, Google Scholar, IN-RECS, IRESIE, Latindex, LivRe!, MIAR, REBIUN, Recolecta, BVS, Capes, Psicodoc, Redalyc, RESH, Sherpa, Sicapes, Socol@r, USIC, ARIADNA, Sport´docs, SUDOC, COPAC, HINARI, & Publindex	1.2	Q3	Hospitality, Leisure, Sport & Tourism	14	5.93
International Journal of Performance Analysis in Sport	SCIE, SSCI, Scopus, SPORTDiscus, DOAJ, ProQuest, EBSCO, ABSEES, & CABI	1.9	Q2	Sport Science	11	4.07
Revista de Psicología del Deporte	Clarivate analytics, Scimago Journal Country Rank, & RECYT	No	-	-	9	3.81
Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte	SCIE, Scopus, Dialnet, DOAJ, & SPORTDiscus	No	-	-	8	3.39
Journal of Human Kinetics	Scopus, SPORTDiscus, ABI/INFORM, ABSEES, CABI, CEEAS, & DOAJ	1.9	Q2	Sport Science	8	3.39
European Journal of Human Movement	ESCI, DOAJ, Latindex, & Google Scholar	0.5	Q4	Sport Science	7	2.97
International Journal of Sports Medicine	MEDLINE, SCI, EMBASE, SCOPUS, BIOBASE, EBSCO, & ProQuest	2	Q2	Sport Science	5	2.12
Journal of Physical Education and Sport	SPORTDiscus, ProQuest, DOAJ, & Open J-GATE	No	-	-	5	2.12
Journal of Sports Science and Medicine	SCIE, SciSearch, MEDLINE, PubMed, Physical Education Index, Nursing & Allied Health Database, SCOPUS, EBSCO, SPORTDiscus, J-Gate, GALE, Google Scholar, ERA, & SPONET	2.4	Q2	Sport Science	5	2.12
European Journal of Sport Science	CABI, EBSCO, CINAHL, Embase, IFIS, Medline, Pubmed, SCIE, Scopus, SPORTDiscus, STM source, & SSCI	2.4	Q2	Sport Science	4	1.69
Quality & Quantity	UGC CARE, & Scopus	1.07	Q3	Social Sciences, Interdisciplinary	4	1.69

Journal	Indexing	IF	Quartile	Category	*f*	%
Revista Iberoamericana de Psicología del Ejercicio y el Deporte	Scopus, ESCI, ERIHPLUS, Latindex, Redalyc, WebQualis, ISOC, DICE, RESH, IN-RECS, Psicodoc, EBSCO, OpenDOAR, Ulrichs Web, Pub Psych, Copac, Recolecta, Dulcinea, Redib, & Sherpa Romeo	No	-	-	4	1.69
Anales de Psicología	SSCI, SCI, Redib, Scielo, PsycInfo, & Latindex	1.4	Q3	Psychology, Multidisciplinary	3	1.27
Cultura, Ciencia y Deporte	SSCI, Scopus, EBSCO, MIAR, Latindex, REDIB, Redalyc, Dialnet, RECOLECTA, ERIH PLUS, Redinet, SPORTDiscus, Psicodoc, DOAJ, ISOC, IN-RECS, Dulcinea, Scirus, WorldCat, LILACS, Gtbib, Research Gate, SafetyLit, REBIUN, E-Revistas, Cabell's Directory, Fuente Académica Plus, ERA BVS, PRESCOPUS Russia, JournalTOCs, & Genamics	0.22	Q3	Hospitality, Leisure, Sport & Tourism	3	1.27
Journal of Human Sport & Exercise	ESCI, Scopus, ROAD, Latindex, Dialnet, Dulcinea, REBIUN, Google Scholar, EBSCOhost, ERIH Plus, MIAR, & Sherpa Romeo	0.5	Q4	Sport Science	3	1.27
Journal of Sport & Exercise Psychology	SSCI, Scopus, PubMed, MEDLINE, ProQuest, PsycInfo, EBSCOhost, & Google Scholar	2	Q2	Hospitality, Leisure, Sport & Tourism	3	1.27
RICYDE: Revista Internacional de Ciencias del Deporte	SSCI, ESCI, Google Scholar, & REDIB	0.7	Q4	Sport Science	3	1.27
Sports	BibCnrs, CNKI, CNPIEC, Dimensions, DOAJ, EBSCO, Scopus, Gale, J-Gate, Pubmed, PMC, OpenAIRE, ProQuest, SafetyLit, ESCI	2.2	Q2	Sport Science	3	1.27
Sustainability	ADS, AGRIS, CABI, CAPlus, CNKI, CNPIEC, Dimensions, EBSCO, Scopus, GEOBASE, FRIDOC, FSTA, Gale, GeoRef, Inspec, J-Gate, OpenAIRE, OSTI, ProQuest, PSYNDEX, EconPapers, IDEAS, SafetyLit, SCIE, & SSCI	3.3	Q3	Green & Sustainable Science & Technology	3	1.27
Perceptual and Motor Skills	SSCI, Index Medicus, PsycInfo, & PubMed	1.4	Q4	Psychology, Experimental	3	1.27
International Journal of Environmental Research and Public Health	AGRIS, BibCnrs, CABI, CAPlus, CNKI, CNPIEC, Dimensions, EBSCO, Scopus, GEOBASE, Embase, FSTA, Gale, J-Gate, PubMed, MEDLINE, PMC, OpenAIRE, OSTI, PATENTSCOPE, ProQuest, PSYNDEX, RePEc, & SafetyLit	4.61	Q1	Public, Environmental & Occupational Health	2	0.85
Journal of Sport and Health Research	DOAJ, Scopus, Dialnet, Latindex, DICE, REDIB, ISOC, DULCINEA, EBSCO, Google Scholar, MIAR, CIRC, RESH, Worldcat, CAPES, Sherpa Romeo, ULRICHS, ERIH PLUS, & ESCI	0.7	Q4	Hospitality, Leisure, Sport & Tourism	2	0.85
Journal of Sports Engineering and Technology	SCIE, ProQuest, & Scopus	1.1	Q3	Sport Sciences	2	0.85
Metodología de las ciencias del Comportamiento	Discontinued as of 2004	No	-	-	2	0.85
Motricidade	Scopus, PsycInfo, IndexCopernicus, Scielo, CABI, SPORTDiscus, EBSCO, ProQues, DOAJ, J-Gate, Latindex, Gale, Google Scholar, SIIC, BVS, Sherpa Romeo, e-Revistas, OCLC, ScienceCentral, Genamics JournalSeek, SafetyLit, NLM, SCIRUS, & BASE	No	-	-	2	0.85
PLoS ONE	Crossref, Dimensions, DOAJ, Google Scholar, PubMed, Scopus, SCIE, AGRICOLA, CABI, CAPES, CAS, CNKI, EconBiz, Embase, MEDLINE, PsycInfo, RePEc, & Zoological Record	2.9	Q1	Multidisciplinary Sciences	2	0.85
Psicothema	SSCI, Fuente Académica Plus, Scopus, IBZ, Google Scholar, Periodicals Index Online, SciELO, MEDLINE, Dialnet, EMBASE, EBSCO, ERIH	3.2	Q1	Psychology, Multidisciplinary	2	0.85

Journal	Indexing	IF	Quartile	Category	*f*	%
	PLUS, PubMed, Latindex, PsycInfo, MIAR, IBECS, CARHUS, Redinet, Rebiun, Psicodoc, DOAJ, Pubpsych, & Crossref					
Brazilian Journal of Kinanthropometry and Human Performance	DOAJ, Genamics JournalSeek, Index Corpenicus, Latindex, Lilacs, LivRe, ProQuest, SciELO, Scopus, Sibradid, & SPORTDiscus	No	-	-	2	0.85
Revista Iberoamericana de Ciencias de la Actividad Física y el Deporte	SSCI, ESCI, Latindex, REDIB, EBSCO, ERIH PLUS, SJI, Open Academic Journal Index, Index Copernicus, MIAR, CIRC, Dialnet, SPORTDiscus, RCS, Dulcinea, WorldCat, Ulrichs, Sherpa Romeo, RiUMA, Dart Europe, Driver, EasyApps, Europeana, ScientificCommons, OpenDoar, OCLC, & BASE	No	-	-	2	0.85
Journal of Education and Training Studies	BASE, CNKI, Crossref, EuroPub, Google Scholar, JournalTOCs, JUFO, Qualis Index, RePEc, ROAD, Scinapse, VOCED & WorldCat	No	-	-	2	0.85
Revista Andaluza de Medicina del Deporte	IBECS & Scopus	No	-	-	1	0.42
Serbian Journal of Sports Sciences	Index Copernicus & SCindeks	No	-	-	1	0.42
Pamukkale Journal of Sport Sciences	Scopus, TR Dizin, EBSCO, SPORTDiscus, DOAJ, Index Copernicus, CAB, & Turkiye Atif Dizini	No	-	-	1	0.42
Portuguese Journal of Sport Sciences	SciELO, SPORTDiscus, & Latindex	No	-	-	1	0.42
Artificial Intelligence in Project Management	Book	No	-	-	1	0.42
Baltic Journal of Health and Physical Activity	SSCI, EZB, Openalex, Scopus, Sherpa Romeo, Sudoc, Wikidata, ZDB, Title DOI, FatCat, Crossref, DOAJ, Road, & PubMed	0.7	Q4	Sport Sciences	1	0.42
Behavior Research Methods	SSCI, UGC, Scopus, PubMed, ProQuest, Medline, Google Scholar, & BIOSIS	4.6	Q1	Psychology, Mathematical	1	0.42
Children	BibCnrs, CNKI, CNPIEC, Dimensions, DOAJ, EBSCO, Scopus, Embase, Gale, J-Gate, PubMed, PMC, OpenAIRE, ProQuest, PSYNDEX, SafetyLit, & SCIE	2	Q2	Pediatrics	1	0.42
Lecturas: EF y Deportes	Dialnet, Latindex, CAPES, Google Scholar, MIAR, Carhus, BASE, Crossref, CIRC, LIVRE, WorldCat, REDIB, LatinREV, ERIH, BINPAR, Dimensions, Library Hub Discover, REBIUN, Malena, CEV, OAIster, AURA, Open University Library, ZDB, ResearchBib, UBL Leipzig, DRJI, ESJI, RootIndexing, EuroPub, Biblat, CLASE, ASCI, & ISI	No	-	-	1	0.42
Germina	Crossref, Scilit, LatinREV, ROAD, Latindex, & Meta Biblioteca	No	-	-	1	0.42
Human Movement Science	Scopus, Medline, SSCI, SCIE, SJR, & SNIP	1.6	Q3	Sport Sciences	1	0.42
International Journal of Computer Science in Sport	Baidu Scholar, Cabell's Journalytics, CNKI Scholar, Digital Library of Zielona Góra, Dimensions, DBLP Computer Science Bibliography, DOAJ, EBSCO, Reaxys, Engineering Village, ExLibris, Google Scholar, Inspec, JST, J-Gate, Jisc, JournalGuide, JournalTOCs, KESLI-NDSL, MyScienceWork, Naver Academic, Naviga, ProQuest, ReadCube, ScienceON, SCILIT, Scite, SCOPUS, Semantic Scholar, TDNet, Ulrichsweb, WanFang Data, WorldCat, & X-MOL	No	-	-	1	0.42
International Journal of Social Psychology	SSCI, PsycInfo, Scopus, Academic Search, SocINDEX, ERIH, Latindex, & ISOC	0.9	Q4	Psychology, Social	1	0.42

Journal	Indexing	IF	Quartile	Category	*f*	%
International Journal of Sport and Exercise Psychology	PsycInfo, Scopus, & SSCI	2.3	Q2	Psychology, Applied	1	0.42
Journal of Strength and Conditioning Research	Academix OneFile, Academic Search, CINAHL, EBSCO, EMBASE, Ex Libris, HINARI, JournalGuide, Physical education Index, proquest, PuibMed, Rehabilitation 6 Sports medicine Source, SCIE, Scopus, SPORTDiscus, & TDNet	2.5	Q2	Sport Sciences	1	0.42
Kinesiology	UGC CARE, Scopus, SCIE, & DOAJ	0.9	Q4	Sport Sciences	1	0.42
Kronos	UGC CARE, Scopus, & DOAJ	No	-	-	1	0.42
Medicina (Kaunas)	BibCnrs, CNKI, CNPIEC, Dimensions, DOAJ, EBSCO, Scopus, Gale, J-Gate, PubMed, Medline, PMC, OpenAIRE, OSTI, PATENTSCOPE, Proquest, & SCIE	2.4	Q1	Medicine, General & Internal	1	0.42
Physiology & Behavior	Scopus, Medline, SSCI, & SCIE	2.4	Q2	Behavioral Sciences	1	0.42
Polish Journal of Sport & Tourism	Arianta, Baidu Scholar, Cabell's Journalytics, CABI, CNKI Scholar, Dimensions, EBSCO, ERIH PLUS, ExLibris, Google Scholar, JST, J-Gate, Jisc, JournalGuide, JournalTOCs, KESLI-NDSL, MIAR, MyScienceWork, Naver Academic, Naviga, OpenAlex, Polish Scientific Journals Database, ProQuest, ReadCube, RePEc, ScienceON, SCILIT, SJR, Scite, SCOPUS, Semantic Scholar, TDNet, Ulrichsweb, WanFang Data, WorldCat, & X-MOL	No	-	-	1	0.42
Procedia - Social and Behavioral Sciences	Discontinued as of 2019	No	-	-	1	0.42
Proceedings from the 9th European Congress of Methodology	RODERIC	No	-	-	1	0.42
Psychology of Sport and Exercise	Scopus, Medline, SSCI, SCIE, & PsycInfo	3.1	Q1	Psychology	1	0.42
Revista Brasileira de Ciencias do Esporte	DOAJ, Google Scholar, LILACS, SciELO, Scopus, & Latindex	No	-	-	1	0.42
Revista Brasileira de Futsal e Futebol	GALE, Advanced Sciences Index, BASE, Dialnet, DOAJ, DRJI, EBSCO, Hinari, OAJI, REDIB, Sherpa Roemo, SPORTDiscus, SCIE, SSCI, AHCI, ESCI, & WorldCat	0.1	Q3	Hospitality, Leisure, Sport & Tourism	1	0.42
Revista digital: Actividad Física y Deporte	REDIB	No	-	-	1	0.42
Revista de Psicología General y Aplicada	Scopus, EBSCO, CARHUS, Crossref, Dialnet, DICE, RESH, DOAJ, Dimensions, E-libro, Google Scholar, Grupo Océano, IBECS, ISOC, Latindex, MIAR, OpenAlex, Proquest, Psicodoc, PsycInfo, Psyke, Redalyc, Rebuin, REDIB, SciELO, ScienceDirect & SJR	No	-	-	1	0.42
Scandinavian Journal of Medicine & Science in Sports	CABI, EBSCO, AMED, Current Contents: Clinical Medicine, ProQuest, Medline, Pubmed, PsycInfo, SCI, SCIE, & RECAL	3.5	Q1	Sport Sciences	1	0.42
The UB Journal of Psychology	Racó, RCUB, Scopus, Latindex, CARHUS, CIRC, DOAJ, ERIH, ECI, JCREI, MIAR, & SJR	0.6	Q4	Psychology	1	0.42
Journal of Sports Medicine and Physical Fitness	CINAHL, Current Contents/Clinical Medicine, EMBASE, PubMed, MEDLINE, SCIE, & Scopus	1.2	Q3	Sport Sciences	1	0.42
Revista de Preparación Física en el Fútbol	Google Scholar	No	-	-	1	0.42

Journal	Indexing	IF	Quartile	Category	*f*	%
Revista Española de Educación Física y Deportes	Dialnet, CIRC, REBIUN, & Latindex	No	-	-	1	0.42
Journal of Science and Medicine in Sport	Scopus, Medline, SCIE, SJR, & SNIP	3	Q1	Sport Sciences	1	0.42
Facta universitatis	DOAJ, EBSCO, Sponet, Proquest	No	-	-	1	0.42
Advances in Physical Education	Aritrilibrary, CABI, CALIS, ChaoXing, Citefactor, Cnplinker, COPAC, CrossRef, DTU Findit, EZB, ERA, GFMER, GETIT@YALE, Infotrieve, i-Scholar, JournalSeek, NSTL, NSLC, OAJSE, Open Access Library, J-Gate, SciLit, Sherpa Romeo, WorldCat, & ZDB	No	-	-	1	0.42
Total					236	100

Note. IF = impact factor.

Table 5 shows the frequency of study objects. The primary papers included in this review were grouped based on the literature (Carling et al., 2005, 2009; Hughes & Franks, 1997; Preciado et al., 2019, 2021; Sarmento et al., 2014) into 26 broad categories (e.g., any action in which a stationary ball reenters play, such as free kicks, corner kicks, throw-ins or penalties, were grouped into the category "set pieces").

Table 5
Objects of research considered in the sample.

Study objects	*f*
Combination plays	82
Set pieces	25
Defensive blocks	23
Tool validation	21
Performance prediction	15
Scoring patterns	15
Spaces & movements	10
Goalkeeper analysis	8
Possession transitions	7
Coach influence on footballers	6
Interaction between players	6
Distance covered	5
Grassroot football	4
Passing sequences	5
Ball recovery	3
Methodological	3
Shots on goal	3
Counter attacks	2
Goalkeeper-player	2
One-on-one game situations	2
Game interruptions	1
Injury patterns	1
Offside analysis	1
Footedness	1
Aggressive play	1
Home advantage	1

The studies included focused primarily on combination plays (82 papers), followed by set pieces (25 papers) and defensive blocks (23 papers).

Table 6 shows the frequency of settings observed. The most frequently observed setting was the Spanish league (55 papers), followed by the FIFA World Cup (51 papers).

Table 6
Study settings considered in the sample.

Observational events	*f*
Spanish league	55
World Cup	51
Euro Cup	33
Grassroot football tournaments	29
UEFA Champions League	15
English league	15
Training sessions	13
Italian league	13
German league	9
Individual player	7
Paralympic/Olympic Games	4
Portuguese league	4
Colombian Cup	3
CONMEBOL Cup	3
UEFA Europa League	3
Norwegian league	3
Unspecified matches	2
Two players	2
Uruguayan league	2
French league	2
Brazilian league	2
A coach	1
Other leagues (Colombia, Cuba, Mexico, Romania, US, and Greece)	1
Other cups (Spain, Brazil, Italy, Portugal, Africa, Asia, Confederations Cup)	1

Proportion analysis

A binomial non-parametric test was conducted to identify differences in research procedures based on MQSOM in the sample of primary studies. Table 7 shows the proportion comparisons of quality indicators based on MQSOM.

Table 7
Proportion comparisons between quality indicators based on MQSOM.

Ratios to compare[a]			Estimate[b]	*SE*	*95% CI*		$Z_{Estimate}$	*p*
					LL	*UL*		
Direct observation	Indirect observation							
201/236 = .85	35/236 = .15		.852	7.68	.803	.893	10.74	< .001
Observational design	No observational design							
113/236 = .48	123/236 = .52		.479	7.68	.416	.543	-0.59	.558
Coding manual	No coding manual							
170/236 = .72	66/236 = .28		.720	7.68	.661	.775	6.71	< .001
Data type specified	No data type specified							
32/236 = .14	204/236 = .86		.136	7.68	.096	.183	-11.13	< .001
Observational instrument	No observational instrument							
199/236 = .84	37/236 = .16		.843	7.68	.793	.886	10.48	< .001
Recording instrument	No recording instrument							
165/236 = .70	71/236 = .30		.699	7.68	.639	.755	6.05	< .001
Quality control software	No quality control software							
98/236 = .42	138/236 = .59		.415	7.68	.354	.479	-2.54	.011
Analysis software	No analysis software							
186/236 = .79	50/236 = .21		.788	7.68	.733	.837	8.79	< .001
Frequency parameter	Sequency parameter							
181/236 = .77	55/236 = .23		.996	7.68	.981	1	7.68	< .001
Data quality control	No data quality control							
187/236 = .79	49/236 = .21		.792	7.68	.738	.841	8.92	< .001
Descriptive statistics	Inferential analysis	Regularity detection						
15/236 = .06	162/236 = .69	59/236 = .25	.936[c]	7.68	.901	.963	13.34	< .001

Note. [a]1-β = .43 for each comparison. [b]Computed using the first value as reference. [c]Computed using the second and third value as reference.

Studies differed significantly in terms of the type of observation (85% direct), coding manual (72% specified), data type (86% unspecified), observation instrument (84% employed), recording software (70% employed), type of parameter (77% frequency), quality control software (59% unspecified), analysis software (79% employed), data quality control (79% employed) and type of analysis (69% inferential; 25% regularity detection and 6% descriptive).

Cluster analysis

A two-step cluster analysis was conducted to obtain methodological quality profiles. Although the MQSOM score did not indicate a normal distribution, $Z(236) = .088$, $p < .001$, the cluster analysis was conducted since this technique is robust to normality assumption (IBM, 2025). Table 8 shows Schwarz's Bayesian Information Criterion (BIC), BIC change, ratio of BIC changes and ratio of distance measures for one to fifteen-cluster solutions.

BIC decreased from the one-cluster model to the six-cluster model. However, the ratio of distance measures reached its maximum when five clusters are considered (1.69). Therefore, a five-cluster solution was identified as optimal model for grouping the sample of primary works included in this study (Schwarz, 1978). Table 9 shows the five clusters in descending order of methodological quality.

Table 8
BIC, BIC change, ratio of BIC changes and ratio of distance measures of clusters 1 to 15.

Cluster solutions	Bayesian Information Criterion (BIC)	BIC change	BIC changes ratio	Ratio of distance measures
1	3473.19			
2	3026.44	-446.75	1	1.58
3	2781.76	-244.68	.548	1.61
4	2669.45	-112.31	.251	1.03
5	2562.39	-107.06	.240	1.69
6	2541.07	-21.31	.048	1.38
7	2553.92	12.85	-.029	1.06
8	2572.05	18.14	-.041	1.02
9	2591.94	19.88	-.045	1.22
10	2626.78	34.84	-.078	1.11
11	2668.30	41.52	-.093	1.04
12	2711.97	43.67	-.098	1.16
13	2763.72	51.75	-.116	1.06
14	2818.39	54.67	-.122	1.03
15	2874.46	56.07	-.126	1.12

Table 9
Quality profiles derived from the systematic methodological review.

P	*f*	GD	D1	D2	Dimension 1: Quality of design					Dimension 2: Quality of measurement & analysis					
					Observation	Observational design	Coding manual	Data Type	Observation instrument	Software			Parameter	Data quality	Data analysis
										Recording	Quality control	Analysis			
1	49	.81	.78	.89	Direct	Mostly specified	Mostly specified	Frequently not specified/ Sometimes Type IV/ Rarely Types I & II	Specified	Specified	Employed	Mostly Employed	Mostly duration	Quantitative	Regularity detection/inferential analysis and rarely descriptive
2	54	.76	.70	.83	Mostly Direct	Specified	Mostly specified	Not specified/ Rarely Type IV	Specified	Mostly Specified	Sometimes employed	Mostly Employed	Mostly Frequency	Mostly quantitative	Mostly Inferential analysis
3	38	.56	.37	.79	Direct	Sometimes specified	Frequently specified	Not specified/ Rarely Type II	Mostly specified	Sometimes specified	Occasionally employed	Mostly Employed	Duration/ Order	Mostly quantitative	Mostly regularity detection and occasionally inferential
4	59	.46	.26	.70	Frequently direct	Rarely specified	Mostly specified	Not specified	Frequently specified	Sometimes specified	Rarely employed	Frequently employed	Frequency	Quantitative, but sometimes absent or qualitative	Mostly inferential and rarely descriptive
5	36	.24	.06	.45	Sometimes direct	Not specified	Not specified	Not specified	Rarely specified	Sometimes specified	Not employed	Sometimes employed	Frequency	Frequently absent/ rarely quantitative	Mostly inferential and rarely descriptive or regularity detection

Note. P = Profile; GD = General Dimension: Methodological Quality; D1 = Dimension 1: Design; D2 = Dimension 2: Measurement & Analysis.

The profile with the highest mean score in MQSOM (P1; GD = .81; D1 = .78; D2 = .89) was characterized by high levels of design and measurement and analysis quality, besides having the highest prevalence of type IV data (concurrent states). The profile with the second highest mean MQSOM score (P2; GD = .76; D1 = .70; D2 = .83) had a moderate level of design quality and a high level of measurement and analysis quality, with the highest prevalence of design specifications for the observational research. A low level of design quality and a high level of measurement and analysis quality characterized the profile with the third highest mean MQSOM score (P3; GD = .56; D1 = .37; D2 = .79), which also had the highest prevalence of order and duration parameters. The fourth profile in terms of highest mean MQSOM score (P4; GD = .46; D1 = .26; D2 = .70) had a low level of design quality and a moderate level of measurement and analysis quality, with the highest prevalence of qualitative data quality control. Finally, the profile with the lowest mean MQSOM score (P5; GD = .24; D1 = .06; D2 = .45) presented low levels of design and measurement and analysis quality. It was the only one that did not specify the observational design, coding manual or the quality control software.

Study design for the observational research, the observation instrument and, most notably, the data quality software are the research procedures that varied the most in the different clusters. The study design was specified or predominantly specified in P1 (high quality of design, and measurement and analysis) and P2 (moderate quality of design and high quality of measurement and analysis), occasionally in P3 (low quality of design and high quality of measurement and analysis), infrequently in P4 (low quality of design and moderate quality of measurement and analysis) and not specified in P5 (low quality of design, and measurement and analysis). The observation instrument was specified in P1 and P2, mostly specified in P3, frequently specified in P4, and rarely specified in P5. Furthermore, the use of software to obtain data quality evidence was consistent in P1, intermittent in P2, occasional in P3, rare in P4 and absent in P5.

Discussion and conclusions

The aim of this systematic review was to analyze how observational methodology was implemented in 236 primary studies on football. After highlighting several findings of the descriptive analysis, it delved into aspects of the proportion analysis. Finally, it interprets the methodological quality profiles derived from the two-step cluster analysis according to the MQSOM criteria.

Spain was the most common affiliation country among the first authors in the sample (153 papers; 64.8%), which is consistent with existing literature and highlights the country's position as a leader in football research and observational methodology (Anguera, 2021, 2023; Lapresa et al., 2021). This may also explain the rising number of publications in scientific journals by researchers affiliated with Spanish institutions (Preciado et al., 2019).

Frontiers in Psychology published the highest quantity of the primary studies included in the sample (20; 8.47%), followed by *Apunts* (19; 8.05%), *Cuadernos de Psicología del Deporte* (18; 7.63%), and the *Journal of Sports Sciences* (16; 6.78%). These findings highlight Spain's pivotal role in the exponential rise in observational methodology studies on football. Previous works that considered primary studies until 2017 and 2019, respectively (Preciado et al., 2019, 2021), found that the *Journal of Sports Sciences* (Q2; IF = 2.3) published the highest

quantity of primary studies. This article reveals that in recent years, this journal has been surpassed in the publication of football research based on observational methodology not only by an international multidisciplinary journal (*Frontiers in Psychology*, Q2; IF = 2.6) but also by two Spanish journals specialized in exercise and sports, namely *Apunts* (Q2; IF = 1.6) and the non-indexed *Cuadernos de Psicología del Deporte*.

The most frequently studied topic was *combination plays*, with 82 studies. The most observed events were the *Spanish league*, with 55 studies, and the *FIFA World Cup*, with the 51 studies. These findings are consistent with those of previous studies, which have identified Spain as a pivotal nation in the development of observational methodology in football, and the World Cup as a salient event in this context, featuring the best of the elite players and the most extensive media coverage (Preciado et al., 2019).

This systematic review represents the first application of the MQSOM to primary works based on observational methodology that examine a wide range of phenomena in football. The analysis of proportions revealed significant differences in the type of observation, coding manual, data type, observation instrument, recording software, quality control software, analysis software, type of parameter, data quality control and type of analysis, a finding that concurs with previous literature (Preciado et al., 2019, 2021).

No significant differences were identified in the observational design (48% of the studies had one). However, this can be considered a positive indicator, since previous studies (Preciado et al., 2019, 2021) reported fewer primary works with an observational design (16% in futsal and 22% in football). This suggests a trend towards research procedures that ensure better methodological quality.

The cluster analysis enabled a clear distinction of methodological consistency by incorporating both the MQSOM and dimensions scores along with the research procedure characteristics when identifying the methodological quality profiles. The resulting five-cluster model matched the number of clusters identified in the futsal literature (Preciado et al., 2021) and exceeded the four clusters identified in the football literature (Preciado et al., 2019). Furthermore, this original contribution represents substantial progress over previous research procedure rankings, as it ranks the profiles in ascending order of quality criteria according to the MQSOM score and enables a characterization of each profile through the dimensions' scores.

Through this innovative application of the MQSOM to rank methodological quality profiles, this study provides a comprehensive overview of the literature on observational methodology applied in football research. Regarding the quality of design, P1 is the only profile that exhibits a high quality level (.78). As for the quality of measurement and analysis, P1, P2 and P3 exhibited a high quality level (.89, .83 and .79, respectively). These findings suggest a discrepancy between the conceptualization of a program based on observational methodology in football and its subsequent implementation. This may be attributed to the relative lack of popularity of low-intervention mixed-method designs compared to high-intervention quantitative designs, such as observational studies or quasi-experimental designs (Madill & Gough, 2008). Conversely, when the measurement and analysis are described, the requisite aspects for assessing methodological quality are typically present, as this stage aligns with usual methodological practices in terms of instrument, procedure and data analysis delimitations (Sanduvete-Chaves et. al., 2025).

In this sense, it is noteworthy that a plethora of the studies were based on observational methodology but did not use this term. Instead, they were labelled as *performance analysis* (e.g., Rampinini et al., 2007), *notational analysis* (e.g., James et al., 2002), or *time-motion analysis* (e.g., Di Salvo et al., 2007). Moreover, several studies employing ex post facto designs can be categorized as observational methodology studies with direct (e.g., Sgrò et al., 2016) or indirect observation (e.g., Thomas et al., 2006). Consequently, there appears to be a lack of networking or consensus between the quantitative and the mixed-methods perspectives in the scientific literature on football. Further work may thus be necessary to integrate both perspectives into a shared understanding of key elements encompassed in the rigorous methodological framework of systematic observation, such as the taxonomy of observational designs, the typology of qualitative data and quantitative parameters, and the endemic procedures for conducting data quality control and analysis. For social scientists, this meticulous integration would facilitate the sequential decision-making that underpins the design of studies based on observational methodology, contributing to the advancement and replicability of football studies.

Limitations and further research

One possible limitation of this study is the selection by language. Although searching for primary works in English on electronic databases may ensure a certain level of overall quality in the works retrieved, it could also introduce bias by potentially excluding the gray literature of a certain methodological quality simply because it was written in another language or is unpublished (Sánchez-Meca, 2022). Additionally, this study is specifically focused on the application of observational methodology in football. Although this field has the largest scientific production based on observational methodology, it would be beneficial to conduct systematic reviews in other sports to verify the findings of this work.

Conclusions

In summary, this study presents the first application of MQSOM to primary works based on observational methodology. The study yielded five quality profiles, which were ranked according to the scale's criteria. Results also revealed a methodological discrepancy between the dimensions of design, and measurement and analysis in the scientific literature based on observational methodology in football. Therefore,

the MQSOM should be disseminated among researchers, journals and international initiatives such as the EQUATOR Network to reinforce systematic observation as a methodological strategy. These findings support the MQSOM's capacity to measure the methodological quality of studies based on observational methodology (Sanduvete-Chaves et al., 2025). The use of MQSOM is recommended as a prescriptive tool for fellow applied researchers to effectively assess the design, measurement, and analysis of their own studies based on observational methodology.

Complementary information

Funding statement. This article was funded by the research project PID2020-115486GB-I00, Ministerio de Ciencia, Innovación y Universidades, MICIU/AEI/10.13039/501100011033, Government of Spain.
Conflicts of interest. The authors declare no conflict of interest.

Acknowledgement. We would like to thank the support of a Chilean government project FONDECYT Regular 1250316 funded by the National Fund for Scientific and Technological Development, ANID. We gratefully acknowledge the support of the Generalitat de Catalunya Research Group (GRUP DE RECERCA E INNOVACIÓ EN DISSENYS [GRID]) Tecnología i aplicació multimedia i digital als dissenys observacionals (Mª Teresa Anguera; grant no. 2017 SGR 1405). Additionally, authors would like to thank Isabella Fioravante for her collaboration applying the instrument to primary studies. Finally, we thank Mª Teresa Anguera for her work in observational methodology, what supposed a significant growing in this area. She really was a guide and an example to follow, as a researcher and as a person for her kind nature. May rest in peace.
Authorship declaration: The authors declare no conflict of interest. This manuscript is original, has not been published before and is not currently being considered for publication elsewhere. We confirm that the manuscript has been read and approved by all named authors and that there is no other person who satisfied the criteria for authorship but are not listed. We also confirm that the order of authors listed in the manuscript has been approved by all of us.

References

Anguera, M. T. (2003). Observational methods (General). In R. Fernández-Ballesteros (Ed.), *Encyclopedia of psychological assessment, Vol. 2* (pp. 632-637). Sage.

Anguera, M. T. (2021). Desarrollando la observación indirecta: Alcance, proceso, y habilidades metodológicas en el análisis de textos [Development of indirect observation: Scope, process and methodological abilities in textual analysis]. In C. Santoyo-Velasco, & L. Colmenares-Vázquez (Eds.), *Patrones de habilidades metodológicas y conceptuales de análisis, evaluación e intervención en ciencias del comportamiento* [Patterns of methodological and conceptual skills of analysis, evaluation and intervention in behavioral sciences]. Libros UNAM.

Anguera, M. T. (2023). Revisitando las revisiones sistemáticas desde la perspectiva metodológica [Revisiting systematic reviews from a methodological perspective]. *RELIEVE, 29*(1). https://doi.org/10.30827/relieve.v29i1.27758

Anguera, M. T., & Blanco-Villaseñor, A. (2005). Registro y codificación en el comportamiento deportivo [Recording and coding in sports behavior]. In A. Hernández-Mendo (Ed.). *Psicología del deporte (Vol. II): metodología* [Sport Psychology (Vol. II): Methodology] (pp. 6–34). Wanceulen.

Anguera, M. T., Blanco-Villaseñor, A., Hernández-Mendo, A., & Losada, J. L. (2011). Diseños observacionales: ajuste y aplicación en psicología del deporte [Observational designs: their suitability and application in sports psychology]. *Cuadernos de Psicología del Deporte, 11*, 63–76. https://revistas.um.es/cpd/article/view/133241

Anguera, M. T., Blanco-Villaseñor, A., Losada, J. L., & Portell, M. (2018). Guidelines for designing and conducting a study that applies observational methodology. *Anuario de Psicología, 48*(1), 9-17. https://doi.org/10.1016/j.anpsic.2018.02.001

Anguera, M. T., Blanco-Villaseñor, A., Losada, J. L., & Sánchez-Algarra, P. (2020). Integración de elementos cualitativos y cuantitativos en metodología observacional [Integration of qualitative and quantitative elements in observational methodology]. *Ámbitos: Revista Internacional de Comunicación, 49*, 49-70. https://doi.org/10.12795/Ambitos.2020.i49.04

Anguera, M. T., Camerino, O., Castañer, M., Sánchez-Algarra, P., & Onwuegbuzie, A. J. (2017). 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. *Frontiers in Psychology, 8*, Article 2196. https://doi.org/10.3389/fpsyg.2017.02196

Anguera, M. T., Magnusson, M. S., & Jonsson, G. K. (2007). Instrumentos no estándar: planteamiento, desarrollo y posibilidades [Non-standard instruments: Approach, development and possibilities]. *Avances en Medición, 5*, 63–82. http://bit.ly/473pXVj

Anguera, M. T., Portell, M., Chacón-Moscoso, S., & Sanduvete-Chaves, S. (2018). Indirect observation in everyday contexts: Concepts and methodological guidelines within a mixed methods framework. *Frontiers in Psychology, 9*, Article 13. https://doi.org/10.3389/fpsyg.2018.00013

Bakeman, R. (1978). Untangling streams of behavior: Sequential analysis of observation data. In G. P. Sackett (Ed.), *Observing Behavior: Vol. 2. Data collection and analysis methods* (pp. 63-78). University of Park Press.

Blanco-Villaseñor, A., Losada, J. L., & Anguera, M. T. (2001). Diseños observacionales, cuestión clave en el proceso de la metodología observacional [Observational designs, a key issue in the process of observational methodology]. *Metodología de las Ciencias del Comportamiento*, *3*(2), 135-160.

Carling, C., Reilly, T., & Williams, A. (2009). *Performance assessment for field sports*. Routledge.

Carling, C., Williams, A., & Reilly, T. (2005). *The handbook of soccer match analysis: A systematic approach to improving performance*. Routledge.

Chacón-Moscoso, S., Anguera, M. T., Sanduvete-Chaves, S., Losada, J. L., & Portell, M. (2019). Methodological quality checklist for studies based on observational methodology (MQCOM). *Psicothema*, *31*(4), 458-464. https://doi.org/10.7334/psicothema2019.116

Chiu, T., Fang, D., Chen, J., Wang, Y., & Jeris, C. (2001). A robust and scalable clustering algorithm for mixed type attributes in large database environment. In D. Lee, M. Schkolnick, F. Provost, & R. Srikant (Eds.), *Proceedings of the seventh ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 263-268). ACM. https://doi.org/10.1145/502512.502549

Cochran, W. G., & Chambers, S. P. (1965). The planning of observational studies of human populations. *Journal of the Royal Statistical Society. Series A (General), 128*(2), 234-266. https://doi.org/10.2307/2344179

Cresswell, J. W., & Plano-Clark, V. (2017). *Designing and conducting mixed methods research*. Sage.

Di Salvo, V., Baron, R., Tschan, H., Montero, F. C., Bachl, N., & Pigozzi, F. (2007). Performance characteristics according to playing position in elite soccer. *International Journal of Sports Medicine*, *28*(3), 222-227. https://doi.org/10.1055/s-2006-924294

Holgado-Tello, F. P. (2015). Análisis de los ítems [Analysis of the items]. In M. I. Barbero-García, E. Vila-Abad, & F. P. Holgado-Tello (Eds.), *Psicometría* [Psychometrics] (pp. 407-468). Sanz y Torres.

Hollander, M., Wolfe, D. A., & Chicken, E. (2013). The dichotomous data problem. In M. Hollander, D. A. Wolfe, & E. Chicken (Eds.), *Nonparametric statistical methods* (pp. 11-38). Wiley.

Hughes, M., & Franks, I. (1997). *Notational analysis of sport*. E&F Spon.

Huwiler-Müntener, K., Jüni, P., Junker, C., & Egger, M. (2002). Quality of reporting of randomized trials as a measure of methodologic quality. *Journal of the American Medical Association, 287*(21), 2801-2804. https://doi.org/10.1001/jama.287.21.2801

IBM. (2025). *TwoStep cluster analysis. IBM SPSS Statistics*. https://www.ibm.com/docs/en/spss-statistics/30.0.0?topic=features-twostep-cluster-analysis

James, N., Mellalieu, S. D., & Hollely, C. (2002). Analysis of strategies in soccer as a function of European and domestic competition. *International Journal of Performance Analysis in Sport, 2*(1), 85-103. https://doi.org/10.1080/24748668.2002.11868263

Kalkbrenner, M. T. (2023). Alpha, omega, and H internal consistency reliability estimates: Reviewing these options and when to use them. *Counseling Outcome Research and Evaluation, 14*(1), 77-88. https://doi.org/10.1080/21501378.2021.1940118

Lapresa, D., Otero, A., Arana, J., Álvarez, I., & Anguera, M. T. (2021). Concordancia consensuada en metodología observacional: efectos del tamaño del grupo en el tiempo y la calidad del registro [The consensus agreement method in observational methodology: Impact of sample size on time requirements and dataset quality]. *Cuadernos de Psicología del Deporte, 21*(2), 47-58. https://doi.org/10.6018/cpd.467701

Madill, A., & Gough, B. (2008). Qualitative research and its place in psychological science. *Psychological Methods, 13*(3), 254-271. https://doi.org/10.1037/a0013220

Norusis, M. J. (2012). *IBM SPSS statistics 19: Advanced statistical procedures companion*. Prentice Hall.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery, 88*, Article 105906. https://doi.org/10.1016/j.ijsu.2021.105906

Portell, M., Anguera, M. T., Chacón-Moscoso, S., & Sanduvete-Chaves, S. (2015). Guidelines for reporting evaluations based on observational methodology. *Psicothema, 27*(3), 283-289. https://doi.org/10.7334/psicothema2014.276

Portney, L., & Watkins, M. (2000). *Foundations of clinical research: Applications to practice*. Prentice Hall.

Preciado, M., Anguera, M. T., Olarte, M., & Lapresa, D. (2019). Observational studies in male elite football: A systematic mixed study review. *Frontiers in Psychology, 10*, Article 77. https://doi.org/10.3389/fpsyg.2019.02077

Preciado, M., Anguera, M. T., Olarte, M., & Lapresa, D. (2021). Revisión sistemática en fútbol sala desde los mixed methods [Systematic review in futsal from mixed methods]. *Revista de Psicología del Deporte, 30*(1), 75-96. https://rpd-online.com/index.php/rpd/article/view/291

Rampinini, E., Coutts, A. J., Castagna, C., Sassi, R., & Impellizzeri, F. M. (2007). Variation in top level soccer match performance. *International Journal of Sports Medicine*, *28*(12), 1018-1024. https://doi.org/10.1055/s-2007-965158

Sánchez-Algarra, P., & Anguera, M. T. (2013). Qualitative/quantitative integration in the inductive observational study of interactive behaviour: Impact of recording and coding predominating perspectives. *Quality & Quantity, 47*, 1237–1257. https://doi.org/10.1007/s11135-012-9764-6

Sánchez-Meca, J. (2022). Revisiones sistemáticas y meta-análisis en educación: un tutorial [Systematic reviews and meta-analyses in Education: A tutorial]. *RiiTE - Revista Interuniversitaria de Investigación en Tecnología Educativa, 13*, 5-40. https://doi.org/10.6018/riite.545451

Sanduvete-Chaves, S., López-Arenas, D., Anguera, M. T., & Chacón-Moscoso, S. (2025). Methodological Quality of studies based on Observational Methodology: A scale for its evaluation. *Psicothema, 37*(1), 1-10. https://doi.org/10.70478/psicothema.2025.37.01

Sarmento, H., Marcelino, R., Anguera, M. T., Campaniço, J., Matos, N., & Leitão, J. (2014). Match analysis in football: A systematic review. *Journal of Sport Sciences, 32*, 1831–1843. https://doi.org/10.1080/02640414.2014.898852

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461-464. https://doi.org/10.1214/aos/1176344136

Sgrò, F., Aiello, F., Casella, A., & Lipoma, M. (2016). Offensive strategies in the European football championship 2012. *Perceptual and Motor Skills*, *123*(3), 792-809. https://doi.org/10.1177/0031512516667455

Thomas, S., Reeves, C., & Smith, A. (2006). English soccer teams' aggressive behavior when playing away from home. *Perceptual and Motor Skills, 102*, 317-320. https://doi.org/10.2466/pms.102.2.317-320