

Article

Confirmatory Factor Analysis and Validity of the Sexual Harassment Scale in Football Refereeing

Josefa Sánchez ^{1*}, Sara Serrat ¹, Estefanía Castillo ¹ and Alberto Nuviala ²

¹ Department of Integrated Didactics, area of Body Expression, School of Sport Education, Psychology and Sciences, University of Huelva, El Carmen Campus, Avenida de las Fuerzas Armadas S/N, Huelva, Spain; E-mails: josefa.sanchez@ddi.uhu.es (J.S.); sara.serrat@alu.uhu.es (S.S.); estefania.castillo@dempc.uhu.es (E.C.).

² Department of Sport Informatics, area of Physical and Sport Education, School of Sport Sciences, Pablo de Olavide University, Ctra. De Utrera, km 1, Seville, Spain; E-mail: anuvnuv@upo.es (A.N.)

* Corresponding author: josefa.sanchez@ddi.uhu.es. Telephone: +34664480295 (J.S.)

Abstract: Inequalities between men and women in the workplace are reflected in professional sports, specifically in football refereeing. This phenomenon sometimes becomes sexual harassment, since it is a stereotypically considered male profession, in which women are a minority. In order to measure that behaviour, it is necessary to count on valid and reliable tools. Therefore, the goal of this study was to determine the factorial structure and the discriminant and convergent validity of the 'sexual experiences questionnaire', version of the Department of Defence (SEQ-DoD). Eighty-nine male football referees and ninety-four female football referees, with a mean age of 23.30 ± 4.85 years, participated in this studio conducted in Andalusia, Spain. A confirmatory factor analysis was performed using the robust maximum-likelihood estimation method. Goodness of fit was assessed and the factorial invariance was calculated to determine the stability of the model. Subsequently, the validity was confirmed. The results corroborated the validity and reliability of the questionnaire adapted to the population studied. Therefore it can be used as a research instrument.

Keywords: sexual harassment, mobbing, football referee, gender inequality, sport.

1. Introduction

Inequality between men and women has been a permanent fact [1,2] occurring at the workplace and associated with physical and psychological violence [3,4]. Despite the proposed equality policies, there is still a gender division at work [5], as reflected by the fact that women's salaries are generally lower than men's [6], or that women tend to have part-time jobs [7]. This is sometimes due to the preconceived idea that women do not have the same interest in doing more competitive jobs [8], and there is bias against women occupying management or important positions [9]. This stereotype is particularly noticeable in the workplace of professional sport, where men have a much greater presence than women [10], and occupy greater number of positions than them [11,12,13].

These differences lead to the emergence of gender discrimination in the workplace. This situation is in many occasions aggravated by physical and psychological violence, with the latter occasionally including sexual harassment [14]. This behaviour threatens the dignity of the persons. Men feel superiority and dominance over women in the workplace and outside it, sometimes leading to cases of sexual violence [15].

Sexual harassment can be divided into three dimensions that feature differences in their concepts, although they are related to each other [16]. These dimensions are: sexual coercion; unwanted sexual attention; and gender-based harassment. Sexual coercion is sexual cooperation in exchange for certain considerations, making itself visible through bribes, threats, and sexual

blackmail. Unwanted sexual attention refers to verbal and non-verbal unwanted offensive behaviours, without being reciprocal, such as invitations to dates despite saying no, inappropriate touching, or non-consensual sex. Gender-based harassment refers to verbal and non-verbal behaviours with the purpose of insulting, harassing, and degrading through disrespect, sexist comments, distinctive treatment, comments about the body, debate about their sexual life, gestures, and inappropriate exhibitions [17].

Sexual harassment victims suffer violation of their dignity. They lose the right to enjoy their jobs, and may have negative consequences on their physical and mental health. In addition, their performance and productivity may decrease [18]. Victims of sexual harassment in the workplace are afraid of the negative consequences it may have on them, even fearing the loss of their jobs [19,20]. Consequently, very few women dare to report these incidents [21,22]. This phenomenon is not foreign to the sport world; it also occurs within this scenario, where there are also power relations [23].

Professional sports are workplaces, where there are various relationships between its workers. This fact also occurs in the most sought-after and popular sport of the moment, i.e., football [24]. There are numerous professions within football, and some of them are important although hardly valued, such as refereeing [25, 26, 27]. It is a minor group and there are very few studies addressing this public [28] and a very low number of sports publications focusing on the figure of referees [29] despite the importance of this group in all sports [30]. However, the number of studies addressing this population is currently increasing [31,32]. The world of arbitration, with a work environment in which both men and women meet, is a highly masculinised world [33,34], where the number of women is significantly lower than that of men [35], this number being equivalent to 1% of all football referees in Spain [36]. This proportion creates barriers that stop female football referees' professional development [37]. These barriers are related to the greater pressure suffered by women [38], the difference in recognition of their work compared to men [35], and even to aesthetic aspects, highlighting the later over their sporting successes [39]. In addition, there is discredit of women's refereeing, as if they were not equally valid or capable of directing a football match [40]. Aspects like these make it impossible for women to dedicate themselves exclusively to refereeing, finding themselves in a situation of double discrimination for being a football referees and for being women [41].

Despite the desire for professional growth that some women have in their professional facet [8], there are harassers who take advantage of their power and women's desire to grow to exceed their attitudes and behaviours in exchange for a reward [33]. In addition, this sexual harassment against female football referees is also perpetrated on the pitch by fans [37], who do not go to the stadiums precisely to enjoy the shows, but to take part in violent acts [30]. This fact is accentuated when it comes to women directing a football match, simply because they are women [35], sometimes even denigrating their dignity [33]. In previous studies [37], several female football referees have referred to the increased scrutiny their decisions suffer because of their sex, and the sexist comments to which they are exposed each time they perform in a match.

Given the need to investigate sexual harassment in the football refereeing world, the purpose of the present study was to adapt the 'sexual experiences questionnaire', version of the Department of Defense (SEQ-DoD) [17]. The goals were to determine the factorial structure and the discriminant and convergent validity of the SEQ-DoD.

2. Materials and Methods

2.1. Participants

The participants of the present study were 183 football referees from different categories, of which 51.4% were women. The mean age was 23.30 ± 4.85 years, and the average experience in refereeing was 5.32 ± 4.80 years.

2.2. Instrument

The SEQ-DoD [17] was the instrument used in the present study. It is a measurement tool used to know offensive sexual experiences. This questionnaire was reviewed and adapted to the football refereeing environment. The SEQ-DoD, in its original version, consisted of four factors, namely: sexist hostility; sexual hostility; unwanted sexual attention; and sexual coercion. All the questions shared a common root: "In the last 12 months, have you observed or been a victim of some type of behaviour described below, perpetrated by others in your work as a football referee?" The body of each element described behaviours that the interviewee might have experienced. The reliability of the instrument, after field work, measured with Cronbach's alpha, was 0.934. Responses were given in a Likert-type scale, ranging from 1 (never) to 5 (very often). Various sociodemographic questions were added to the questionnaire, such as sex, age, experience in football refereeing, and refereeing category.

2.3. Procedure

First, the organisation responsible for the football refereeing that participated in the study was informed about it. Participating referees were asked for permission to request their informed consent. The study was conducted after approval. The design took into account the principles established in the Declaration of Helsinki [42]. In the same way, we took into consideration the current Spanish legal regulations that normalize the protection of personal data [43]. The field work was carried out by means of a self-administered questionnaire with the presence of an interviewer, which lasted about ten minutes.

2.4. Stadistic analysis

Firstly, we performed a confirmatory factor analysis. The method used was the robust maximum-likelihood estimation. In order to determine the goodness of fit, we reviewed the indicators, namely: the Chi-square value divided by the degrees of freedom (χ^2/gf); root mean square error of approximation (RMSEA); and comparative fit index (CFI) [44,45]. In addition, in order to follow Byrne's indications [46], we added the Akaike information criterion and the expected cross-validation index. Subsequently, the factorial invariance was calculated in order to determine the stability of the model in different populations.

Convergent validity tests were performed by calculating correlations between factors and composite reliability. Finally, we determined the discriminant validity using three different procedures: calculation of correlations between factors and comparison with the square root of the average variance extracted (AVE); estimation of alternative models; and construction of confidence intervals for factors correlation with 95% confidence interval. The statistical analyses were performed using the statistical packages SPSS and AMOS v23.

3. Results

In order to confirm whether or not the scale met the expected factorial structure, we performed a confirmatory factor analysis. The adequacy of the model under test (model 0), which consisted of

four factors and twenty-four items, was carried out through a joint assessment of a group of indices. Table 1 contains the information provided by the adjustment indices, and it can be concluded that it was a correct model.

The factorial invariance of the model was contrasted by comparing two groups of football referees, which were selected at random among the population object of the present study. We considered the differences in χ^2 between the models without restrictions (model 1) and the rest of the models with restrictions, observing differences between models 1 vs. 2, and 2 vs. 3 (Table 1). The CFI value of the models indicated that all of them had very similar values, with a difference between them equal to -0.01. Similarly, the Akaike information criterion and the expected cross-validation index indicated that the differences in the adjustments were minimal; therefore, the different models exhibited very similar values. These results suggest the factorial invariance of the model.

Table 1. *Statistics adjustment for the SEQ-DoD scale model; comparison between models using model 1 as correct.*

Goodness-of-fit indices and model comparisons of the tested models								
Model	CMIN	DF	<i>p</i>	CMIN/DF	CFI	RMSEA	ECVI	AIC
Model 0	332.997	238	0.000	1.399	0.930	0.063	4.570	456.997
Model 1	596.642	476	0.000	1.253	0.913	0.051	8.532	844.642
Model 2	630.864	496	0.000	1.272	0.903	0.052	8.473	838.864
Model 3	640.836	506	0.000	1.266	0.903	0.052	8.372	828.836
Comparisons of conditions using measurement invariance procedures								
Model			Dif. DF	Dif. CMIN		<i>p</i>		
Assuming that model 1 is correct	2	3	20	34.222		0.025		
Assuming that model 2 is correct	3	3	30	44.194		0.046		
Assuming that model 3 is correct	3	3	10	9.972		0.443		

Note. CMIN: minimum discrepancy; DF: degrees of freedom; CFI: comparative fit index; RMSEA: root mean square error of approximation; ECVI: expected cross-validation index; AIC: Akaike information criterion; Model 1 had no restrictions; Model 2 had restrictions relating to the weight measurement; Model 3 had weight measurement and covariance restricted; Dif. CMIN: difference between model 1 and the rest of the models; Dif. DF: difference between model 1 and the rest of the models; *p*: significance level between models.

The convergent validity was confirmed by the calculation of the correlations between the factors of the SEQ-DoD. The results indicated positive and significant correlations between the factors of the scale. Similarly, the composite reliability values obtained for each dimension suggested the existence of this type of validity.

Table 2. *Means, correlations between factors, and square roots of average variance extracted (in the diagonal); Cronbach's alpha; Composite reliability.*

	Total mean	Sexist hostility	Sexual hostility	Unwanted sexual attention	Sexual coercion	α	CR
Sexist hostility	2.46±1.04	(0.716)	0.612**	0.485**	0.427**	0.803	0.806
Sexual hostility	2.23±0.95		(0.735)	0.602**	0.525**	0.899	0.901
Unwanted sexual attention	1.44±0.60			(0.629)	0.599**	0.825	0.817
Sexual coercion	1.22±0.54				(0.714)	0.830	0.837

Note. * $p < 0.05$; ** $p < 0.01$; CR: composite reliability.

In order to determine the discriminant validity, the square root of the AVE was compared with the correlation between both constructs. Table 2 shows this correlation and, in the diagonal, the square root of the AVE, which was superior to the correlation between the different constructs of the questionnaire. Considering these results, it can be affirmed that there was discriminant validity.

As a second discriminant validity test, alternative models were estimated in such a way that a restriction in all of them, i.e., the correlation between each pair of dimensions should be equal to 1. In addition, the chi-square test was performed with each one to compare the models in order to assess whether or not they were significantly different. Table 3 shows how the difference between the chi-square test values was always significant. This way, the dimensions of the scale were significantly different from each other, thus confirming the discriminant validity.

As a third way to confirm this type of validity, we calculated the possible correlations between the factors. This procedure allowed the construction of the confidence interval relating to the correlations between the dimensions. Table 3 shows that the discriminant validity of the scale could be confirmed, since none of the confidence intervals of these correlations contained value 1 at 95% confidence.

Table 3. Test of χ^2 differences; confidence intervals of correlations between dimensions.

	χ^2 differences (gl)	p	Confidence interval
Sexist Hostility / Sexual Hostility	$334.129(239)-332.997(238) = 1.132$ (1)	0.000	(0.679 - 815)
Sexist Hostility / Unwanted Sexual Attention	$351.000(239)-332.997(238) = 18.003$ (1)	0.000	(0.351 - .600)
Sexist Hostility / Sexual Coercion	$360.712(239)-332.997(238) = 27.715$ (1)	0.000	(0.286 - .553)
Sexual Hostility/ Unwanted Sexual Attention	$344.250(239)-332.997(238) = 11.253$ (1)	0.000	(0.577 - .750)
Sexual Hostility/ Sexual Coercion	$355.737(239)-332.997(238) = 22.74$ (1)	0.000	(0.400 - .631)
Unwanted Sexual Attention/ Sexual Coercion	$367.236(239)-332.997(238) = 34.239$ (1)	0.000	(0.503 - .828)

4. Discussion

The goals of the present study were to determine the factorial structure and the discriminant and convergent validity of the SEQ-DoD [17] in Spanish football referees. The results confirmed the validity and reliability of the adaptation of the questionnaire to the population under study. The resulting latent variables were the same as those in the original questionnaire.

The purpose of our study was to determine the fit of the original model to the data obtained from a sample of Spanish football referees. To that end, we performed a confirmatory factor analysis. The parameters were estimated using the maximum likelihood method [47]. In order to assess the adequacy of the model under test, we performed a joint assessment of a group of indices. Some of the most used adjustment indices were selected considering values above 0.90 acceptable in the case of the CFI. In the case of RMSEA, the model would exhibit an acceptable fit if the value was <0.07 [48], and values ≤ 0.06 would indicate a good fit [49]. Regarding the values of the quotient between χ^2 and gl, in a model considered perfect, the value would be 1.00, and ratios below 2.00 would be considered a very good fit of the model, whereas values below 5.00 would be considered acceptable [49,50,51]. Finally, due to the convenience of comparing the fit of the model, we added two specially developed indices, namely: the Akaike information criterion, i.e., a comparative index between models, having to choose the model that presents a lower value [52] (values closer to zero indicate a better fit); and the expected cross-validation index, which measures the discrepancy between the covariance matrix involved in the analysed sample and the expected covariance matrix of another sample of the same size. When models are being compared, the lower expected cross-validation index value indicates the model with the best fit [52].

The results of the different fit indices of the original model can be considered acceptable. Therefore, the model can be considered correct for the population of football referees assessed in the present study. Furthermore, the reliability of the resulting instrument measured with Cronbach's alpha was 0.934, which indicated good internal consistency.

Subsequently, we assessed the invariance of the factorial structure through multi-group análisis [53]. To that end, the group was divided into two subgroups at random. The aim was to confirm that there were no significant differences between a model without invariance and different models with invariance in some parameters. We found significant differences in chi-square values between the unrestricted model (Model 1) and the rest of the models. However, given that the chi-square coefficient is sensitive to sample size, we also used the criterion proposed by Cheung and Rensvold [54] with respect to ΔCFI . According to these authors, ΔCFI values lower than or equal to -0.01 indicate that the null hypothesis of invariance cannot be rejected. The ΔCFI values found in the present study, in the comparison of the unrestricted model with the rest of the models, suggest that the factorial structure of the scale was invariant.

The convergent validity was determined by the correlations between the SEQ-DoD factors using Pearson's correlation coefficient. The correlations between them were positive, being high in some cases, which can give an idea of the similarity of the constructs. The results of the correlations demonstrated this type of validity, since the results were within the criteria proposed by Devon et al. [55] for this type of validity. A second test of convergent validity of the instrument was determined by composite reliability. Acceptable values are >0.6 [56,57]. Both tests indicated the existence of this type of validity.

The discriminant validity of the scale was expressed by the contrast between the different factors that composed it. This type of validity occurs if the concepts that comprise it are really different and, at the same time, related to each other [58]. In order to confirm this validity, these concepts were assessed in various ways. The first consisted of comparing the square root of the AVE with the correlation between the constructs of the scale [56]. The square root of the AVE should be higher than the correlation between the constructs, so that there is discriminant validity between them. Considering the results of the correlations and the AVE values, it can be affirmed that there was discriminant validity.

This type of validity can also be confirmed in two other ways. The first has been proposed by Burnkrant and Page [59]. It attempts to estimate alternative models in such a way that a restriction is included in all of them, i.e., the correlation between each pair of dimensions should be equal to 1. In the other, each model should be subject to a chi-square test to compare them and assess whether they are significantly different. Our results have proven that the difference between the chi-squared values was always significant. Therefore, the dimensions of the SEQ-DoD were different from each other, thus confirming the discriminant validity.

The third way consists in calculating the possible correlations between the factors and constructing the confidence intervals of the correlations between all the dimensions. The results of the present study also indicated the occurrence of this type of validity, since none of the confidence intervals of these correlations contained the value 1 at 95% confidence [60].

5. Conclusions

In conclusion, the SEQ-DoD has proven to be valid and reliable. However, it is still in an early stage. The limitations related to the number of football referees assessed in the field work and the lack of bibliography on the subject makes it necessary to conduct further studies in depth and improve, if possible, this instrument.

Author Contributions: Conceptualization, S.S.; methodology, J.S. ; validation, S.S.; formal analysis, A.N. ; investigation, E.C.; resources, J.S.; data curation, A.N.; writing—original draft preparation, S.S.; writing— review and editing, E.C.; visualization, E.C.; supervision, A.N.; project administration, J.S. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. González, M. Género y trabajo en los contact centers. *Revista de Ciencias Sociales*. **2019**, 7, 261-273. <https://doi.org/10.17502/m.rcs.v7i2.270>
2. Hontangas, J.; Mestre, J.A.; Orts, F. *Género y deporte (El camino hacia la igualdad)*, Madrid, España: Reus. **2018**.
3. Cuesta, O. El discurso femenino en el mercado laboral. *Prisma Social. Revista de Ciencias Sociales*. **2011**, 7, 296-320.
4. De Oliveira, O. Reflexiones acerca de las desigualdades sociales y el género. *Estudios sociológicos*. **2007**, 25, 805-812.
5. Torns, T. El trabajo y el cuidado: cuestiones teórico-metodológicas desde la perspectiva de género. *EMPIRIA: Revista de Metodología de Ciencias sociales*. (2008). **2008** 15, 53-73. <https://doi.org/10.5944/empiria.15.2008.1199>
6. Anghel, B.; Conde-Ruiz, J.I.; Marra de Artiñano, I. Brechas salariales de género en España. *Hacienda Pública Española*. **2018**, 229, 87-119. <https://doi.org/10.7866/hpe-rpe.19.2.4>
7. Díaz, C.; Díaz, C.; Dema, S.; Ibáñez, M. Dinero, amor e individualización: las relaciones económicas en las parejas/familias contemporáneas. *Revista Española de Investigaciones Sociológicas*. **2004**, 115, 370-373.
8. Donoso, T.; Figuera, P.; Roig, M.L. Barreras de género en el desarrollo profesional de la mujer universitaria1 Gender barriers in the professional development of women university graduates. *Revista de educación*. **2011**, 355, 187-212. <https://doi.org/10.4438/1988-592X-RE-2011-355-021>
9. Cuadrado, I. Valores y rasgos estereotípicos de género de mujeres líderes. *Psicothema*. **2004**, 16, 270- 275.
10. Chalabaev, A.; Sarrazin, P.; Fontayne, P.; Boiché, J.; Clément-Guillotin, C. The influence of sex stereotypes and gender roles on participation and performance in sport and exercise: Review and future directions. *Psychology of Sport and Exercise*. **2013**, 14, 136-144. <https://doi.org/10.1016/j.psychsport.2012.10.005>
11. Hoel, H.; Faragher, B.; Cooper, C. L. Bullying is detrimental to health, but all bullying behaviours are not necessarily equally damaging. *British Journal of Guidance and Counselling*. **2004**, 32, 367-387. <https://doi.org/10.1080/03069880410001723594>
12. Koenig, A. M.; Eagly, A. H.; Mitchell, A. A.; Ristikari, T. Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*. **2011**, 137, 616-642. <https://doi.org/10.1037/a0023557>
13. Pérez-Rivases, A.; Torregrosa, M.; Viladrich, C.; Pallarés, S. Women occupying management positions in top-level sport organizations: a self-determination perspective. *Anales de Psicología*. **2017**, 33, 102-113. <https://doi.org/10.6018/analesps.33.1.229171>
14. Raver, J. L.; Hisae, L. Once, twice, or three times as harmful? Ethnic Harassment, gender harassment, and generalized workplace harassment. *Journal of Applied Psychology*. **2010**, 95, 236-254. <https://doi.org/10.1037/a0018377>
15. Carballido, P. Movimientos sociales y medios de comunicación: el cambio en el tratamiento de la violencia contra las mujeres. *RECERCA. Revista de Pensament y Anàlisi*. **2007**, 7, 211-240.
16. Gelfand, M. J.; Fitzgerald, L. F.; Drasgow, F. The structure of sexual harassment: A confirmatory analysis across cultures and settings. *Journal of Vocational Behavior*. **1995**, 47, 164-177. <https://doi.org/10.1006/jvbe.1995.1033>
17. Fitzgerald, L. F.; Magley, V. J.; Drasgow, F.; Waldo, C. R. Measuring sexual harassment in the military: the sexual experiences questionnaire (SEQ—DoD). *Military Psychology*. **1999**, 11, 243-263. https://doi.org/10.1207/s15327876mp1103_3
18. Dias, I. Violência contra as mulheres no trabalho: O caso do assédio sexual. *Sociologia, Problemas e Práticas*. **2008**, 57, 11-23.
19. Lampman, C.; Phelps, A.; Bancroft, S.; Beneke, M. Contrapower harassment in Academia: A survey of faculty experience with student incivility, bullying and sexual attention. *Sex Roles*. **2009**, 60, 331-346. <https://doi.org/10.1007/s11199-008-9560-x>

20. Macik-Frey, M.; Campbell, J.; Nelson, D. L. Advances in occupational health: From a stressful beginning to a positive future. *Journal of Management*. 2007, 33, 809-840. <https://doi.org/10.1177/0149206307307634>

21. Bergman, M. E.; Langhout, R. D.; Palmieri, P. A.; Cortina, L. M.; Fitzgerald, L. F. The (un) reasonableness of reporting: Antecedents and consequences of reporting sexual harassment. *Journal of Applied Psychology*. 2002, 87, 230-242. <https://doi.org/10.1037/0021-9010.87.2.230>

22. Herrera, M.C.; Herrera, A.; Expósito, F. To confront versus not to confront: Women's perception of sexual harassment. *European journal of psychology applied to legal context*. 2017, 10, 1-7. <https://doi.org/10.1016/j.ejpal.2017.04.002>

23. Leruite, M. T.; Martos, P.; Zabala, M. Análisis del deporte femenino español de competición desde la perspectiva de protagonistas clave. *Retos: nuevas tendencias en educación física, deporte y recreación*. 2015, 28, 3-8. <https://doi.org/10.47197/retos.v0i28.34821>

24. Úbeda, J. Fútbol contra el enemigo. *Cultura, Ciencia y Deporte*. 2013, 8, 157-158.

25. González-Oya, J. Aproximación a la violencia en el fútbol y en el arbitraje. *Revista Iberoamericana de Psicología del Ejercicio y el Deporte*. 2006, 1, 29-44.

26. Moriconi, M.; Teixeira-Diniz, R. Mitos, percepciones y actitudes frente a los resultados combinados en Portugal: un estudio con énfasis en los árbitros y los hinchas. *Sociología, Problemas e Práticas*. 2016, 82, 145-164. <https://doi.org/10.7458/SPP2016827159>

27. Parsons, T.; Bairner, A. You want the buzz of having done well in a game that wasn't easy: A sociological examination of the job commitment of English football referees. *Movement & Sport Sciences- Science & Motricité*. 2015, 87, 41-52. <https://doi.org/10.1051/sm/2014017>

28. Marrero, G.; Gutiérrez, C. Las motivaciones de los árbitros de fútbol. *Revista de Psicología Del Deporte*. 2002, 11, 69-82.

29. Soriano, G.; Ramis, Y.; Torregrosa, M.; Cruz, J. Apoyo contra el estrés: ¿Cómo podemos ayudar a los árbitros? *Kronos*. 2017, 16, 1-10.

30. Alonso-Arbiol, I.; Arratibel, N.; Gómez, E. La motivación del colectivo arbitral en fútbol: un estudio cualitativo. *Revista de Psicología del Deporte*. 2008, 17, 187-203.

31. Kellett, P.; Shilbury, D. Umpire participation: is abuse really the issue? *Sport Management Review*. 2007, 10, 209-229. [https://doi.org/10.1016/S1441-3523\(07\)70012-8](https://doi.org/10.1016/S1441-3523(07)70012-8)

32. Kellett, P.; Warner, S. Creating communities that lead to retention: The social worlds and communities of umpires. *European Sport Management Quarterly*. 2011, 11, 471-494. <https://doi.org/10.1080/16184742.2011.624109>

33. Cuenca-Piqueras, C. Factores precipitantes del acoso sexual laboral en España. *Revista mexicana de sociología*. 2015, 77, 525-554.

34. Fink, J.S. Gender and sex diversity in sport organizations: Concluding comments. *Sex Roles*. 2008, 58 (1-2), 146-147.

35. Santillán, M.; Gantús, F. Transgresiones femeninas: fútbol. Una mirada desde la caricatura de la prensa, México 1970-1971. *Tzintzun. Revista de Estudios Históricos*. 2010, 52, 143-176.

36. Federación Internacional de Fútbol Asociado. Professional Football Report 2019. 2019. Available online: <https://img.fifa.com/image/upload/jlr5corccbsef4n4brde.pdf>

37. Forbes, A.; Edwards, L.; Fleming, S. 'Women can't referee': exploring the experiences of female football officials within UK football culture. *Soccer and society*. 2015, 16, 521-539. <https://doi.org/10.1080/14660970.2014.882829>

38. Messner, M. Boyhood, Organized Sports and the Construction of Masculinities. *Journal of Contemporary Ethnography*. 1990, 18, 416-444. <https://doi.org/10.1177/089124190018004003>

39. Oliveira, M. T.; Mendes, A.; Jensen, L. Olhos masculinos nascidos para a contemplação do belo: a relação entre esporte e mulher na crônica esportiva brasileira. *Revista Brasileira de Ciências do Esporte*. 2017, 39, 355-361. <https://doi.org/10.1016/j.rbce.2017.09.001>

40. Selva, C.; Pallarès, S.; González, M.D. Una mirada a la conciliación a través de las mujeres deportistas. *Revista de Psicología del Deporte*. 2013, 22, 69-76.

41. González, M.D.; Selva, C.; Torregrosa, M. Mujeres y árbitras: historias de vida de una doble minoría en el deporte. *Revista de Psicología del Deporte*. 2018, 27, 23-30.

42. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013, 310, 2191–2194. <https://doi.org/10.1001/jama.2013.281053>

43. Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Boletín Oficial del Estado, 294, de 06/12/2018. Available online: <https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673>

44. Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E.; Tatham, R.L. *Multivariate data analysis* (6^a Edición). Upper Saddle River, NJ: Pearson-Prentice-Hall. 2006.

45. Marôco, J. *Análise de Equações Estructurais. Fundamentos teóricos, Software & Aplicações*. Pero Pinheiro: Rolo y Filhos II, SA. 2010.

46. Byrne, B.M. *Structural Equation Modeling with AMOS. Basic Concepts, Applications, and Programmin*. Londres: LEA. 2001.

47. Thompson, B. *Exploratory and Confirmatory Factor Analysis. Understanding concepts and applications*. Washington, D.C.: American Psychological Association. 2004.

48. Steiger, J. H. Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences*. 2007, 42, 893-898. <https://doi.org/10.1016/j.paid.2006.09.017>

49. Hu, L.; Bentler, P. M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*. 1999, 6, 1-55. <https://doi.org/10.1080/10705519909540118>

50. MacCallum, R. C.; Widaman, K. F.; Preacher, K. J.; Hong, S. Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research*. 2001, 36, 611-637. https://doi.org/10.1207/S15327906MBR3604_06

51. Yuan, K. H. Fit indices versus test statistics. *Multivariate Behavioral Research*. 2005, 40, 115-148. https://doi.org/10.1207/s15327906mbr4001_5

52. Schermelleh-Engel, K.; Moosbrugger, H.; Müller, H. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research*. 2003, 8, 23-74.

53. Abalo, J.; Lévy, J.; Rial, A.; Varela, J. Invarianza factorial con muestras múltiples. En J. Lévy (Ed.), *Modelización con Estructuras de Covarianzas en Ciencias Sociales* (pp. 259-278). Madrid, España: Netbiblo. 2006.

54. Cheung, G. W., y Rensvold, R. B. Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling*. 2002, 9, 233-255. https://doi.org/10.1207/S15328007SEM0902_5

55. Devon, H. A.; Block, M. E.; Moyle-Wright, P.; Ernst, D. M.; Hayden, S. J.; Lazzara, D. J.; Savoy, S.M.; Kostas-Polston, E. A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship*. 2007, 39, 155-164. <https://doi.org/10.1111/j.1547-5069.2007.00161.x>

56. Fornell C.; Larcker, D. F. Evaluating structural model with unobserved variables and measurement errors. *Journal of Marketing Research*. 1981, 18, 39-50. <https://doi.org/10.2307/3151312>

57. Bagozzi, R.; Yi, Y. On the Evaluation of Structural Equation Models. *Journal of the Academy of Marketing Science*. 1988, 16, 74-94. <https://doi.org/10.1007/BF02723327>

58. Lehmann, D.R.; Gupta, S.; Steckel, J.H. *Marketing Research*. New York: Addison-Wesley. 1999.

59. Burnkrant, R.E.; Page, T.J. An examination of the convergent, discriminant and predictive validity of Fishbein's behavioral intention model. *Journal of Marketing Research*. 1982, 19, 526-550. <https://doi.org/10.1177/002224378201900414>

60. Anderson, J. C.; Gerbing, D. W. Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Psychological Bulletin*. 1988, 103, 411-423. <https://doi.org/10.1037/0033-2909.103.3.411>