THE PSYCHOLOGY OF SPORT, PERFORMANCE AND ETHICS

EDITED BY: Yair Galily, Roy David Samuel, Gershon Tenenbaum and Edson Filho
PUBLISHED IN: Frontiers in Psychology and Frontiers in Sports and Active Living

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88966-679-9
DOI 10.3389/978-2-88966-679-9

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

THE PSYCHOLOGY OF SPORT, PERFORMANCE AND ETHICS

Topic Editors:
Yair Galily, Interdisciplinary Center Herzliya, Israel
Roy David Samuel, Interdisciplinary Center Herzliya, Israel
Gershon Tenenbaum, Interdisciplinary Center Herzliya, Israel
Edson Filho, Boston University, United States

Citation: Galily, Y., Samuel, R. D., Tenenbaum, G., Filho, E., eds. (2021). The Psychology of Sport, Performance And Ethics. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-679-9

Table of Contents

CHAPTER 4
PARENTAL STYLE AND INVOLVEMENT IN SPORT

frontiers in Psychology

EDITORIAL
published: 05 March 2021
doi: 10.3389/fpsyg.2021.658457

Editorial: The Psychology of Sport, Performance and Ethics

Yair Galily[1*], *Roy D. Samuel*[1], *Edson Filho*[2] *and Gershon Tenenbaum*[1]

[1] Interdisciplinary Center Herzliya, Herzliya, Israel, [2] Wheelock College of Education and Human Development, Boston University, Boston, MA, United States

Keywords: ethics, exercise, coronavirus, sports, athlete, coaches

Editorial on the Research Topic

The Psychology of Sport, Performance and Ethics

Sport and exercise are considered a significant factor in the lives of many people (see Simon, 2018; Galily, 2019, 2021). While many people exercise and are engaged in several physical activities, others view and watch sport competitions and events. Indeed, evenw those who are unconcerned with sporting games, and are critical of athletic rivalry, are often affected by them due to their relationships with other enthusiasts, and the impact of sport on human language, thought, and culture. The current limitations posed on exercise engagement and sport participation due to the COVID-19 global pandemic further emphasized the importance of exercise and sport for physical and mental health, and interpersonal interactions (Samuel et al., 2020).

Because sport is a meaningful form of social activity that affects politics, economy, and behavioral norms, it brings up a wide range of issues, some of which attract research and scientific inquiry (Simon, 2018: p. 3). While research in various domains of sport pertain to contextual and social factors, some were devoted to the factors which promote or hinder dishonesty. In addition, advances in technology and the growth of social media usage, along with substantial financial investment in sport, raise significant ethical questions worth elaboration and inquiry.

The aim of the current compendium on the psychology of sport, performance, and ethics was to assemble both theoretical and applied research from experts within the field of sport psychology, sociology, performance, and exercise. Twelve articles, written by researchers from Brazil, China, Germany, Israel, the Netherlands, Norway, Portugal, Spain, Taiwan, and the United States, were divided into four chapters. The first chapter, "Decision-Making Challenges in Dynamic Sporting Environments" holds four articles: Gershgoren et al. introduce the chapter *Perceived Performance in Team Sports Questionnaire* to capture the team members' perception of their team's performance. Samuel et al.'s case study adopted an intrinsic mixed-methods methodology to investigate the implementation of the video assistant (soccer) referee system within the Israeli Premier League context. Johansen and Erikstad investigated elite referees' positioning in the field of play (distance, angle, and insight) when making correct and erroneous decisions in potential penalty situations. Finally, Del Campo and Martin assessed the effects of manipulating video speeds on visual behavior and decision accuracy of 10 amateur football assistant referees when watching video sequences of 24 possible offside actions.

The second chapter, entitled "Integrity and Ethical Issues in Sport Psychology," comprises three articles: Englert and Schweizer tested the capability of individuals to judge correctly whether athletes are lying or telling the truth, and suggested that participants can distinguish between true and false statements, but only for some clips and not for others; indicating that some players were better at deceiving than others. Tamir and Bar-eli argued that despite fierce objections and extensive criticism, the video assistant referee system represents an important turning point in modern professional soccer, and moreover, it accomplishes a moral revolution in the evolution of the sport

Edited and reviewed by:
Guy Cheron,
Université libre de Bruxelles, Belgium

****Correspondence:***
Yair Galily
ygalily@idc.ac.il

Specialty section:
This article was submitted to Movement Science and Sport Psychology, a section of the journal Frontiers in Psychology

Received: *25 January 2021*
Accepted: *11 February 2021*
Published: *05 March 2021*

Citation:
Galily Y, Samuel RD, Filho E and Tenenbaum G (2021) Editorial: The Psychology of Sport, Performance and Ethics. Front. Psychol. 12:658457. doi: 10.3389/fpsyg.2021.658457

domain. The newly developed technologies enhanced the sport's professional standards and its public image and prestige, and especially the moral standard of fair play. Lastly, to examine effective communication and coordination within the context of referee teams, Sinval et al. developed the Referee Shared Mental Models Measure (RSMMM).

The third chapter, "Performance Enhancement and Aggressive Behavior" congregates three issues. Wang et al. offered support for the meshed control theory and indicated the dynamic nature of neuromotor processes for the superior performance of athletes in challenging situations. Strenge et al. outlined how the domain of cognitive interaction technology research addresses ethical issues and presented an empirical study in the context of a new measurement and assessment system for training in karate. In two experiments, Geng et al. explored the implicit link between the color red and aggressiveness as well as the color blue and agreeableness in students and in Taekwondo athletes.

The fourth chapter labeled "Parental Style and Involvement in Sport," includes two studies: Yaffe et al. tested the relationships between sport type (team or individual) and parenting styles (authoritative vs. non-authoritative), and moral decision-making in sport and sport values. Correspondingly, Lev et al. explored the nature of parental involvement in youth basketball in Israel with regards to parenting style and in the context of dilemmas and ethical issues.

In line with Sly et al.'s (2020) notion, our practice is increasingly characterized by a complex interaction among clientele, roles, services, and competencies. Hence, it is our hope that insights of this diverse collection of articles further expands the boundaries of issues related to athletes and coaches and the whole sporting eco-system. Acknowledgment of the recent diversification of sport psychology provision has been evidenced by Division 47 (Sport, Exercise, and Performance Psychology; SEP) of the American Psychological Association, advocating that SEP psychology must be conceptualized as a subdiscipline of performance psychology; that is, a domain of study and practice concerned with the identification, development, and execution of skills and abilities required to achieve excellence within a series of diverse performance domains, such as business, military, health care, education, and the performing arts (Sly et al., 2020: p. 87). In our humble effort to create a more intelligible picture of what characterizes and bounds our discipline, we conclude by saying: Let the games (and reading) begin!

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

The guest editors would like to personally thank all the authors and reviewers who contributed to this research topic. We also greatly appreciate the help of the Frontiers editorial office, as well as the kind assistance of Ilaria Prete.

REFERENCES

Galily, Y. (2019). "Shut up and dribble!"? Athletes activism in the age of twittersphere: the case of LeBron James. *Technol. Soc.* 58:101109. doi: 10.1016/j.techsoc.2019.01.002

Galily, Y. (2021). The coronavirus pandemic as a game-changer: when NBA players forced America to think. Again. *Front. Psychol.* 11:600267. doi: 10.3389/fpsyg.2020.600267

Samuel, R. D., Tenenbaum, G., and Galily, Y. (2020). The 2020 coronavirus pandemic as a change-event in sport performers' careers: conceptual and applied practice considerations. *Front. Psychol.* 11:567966. doi: 10.3389/fpsyg.2020.567966

Simon, R. L. (2018). *Fair Play: The Ethics of Sport*. London: Routledge.

Sly, D., Mellalieu, S. D., and Wagstaff, C. R. (2020). "It's psychology Jim, but not as we know it!": the changing face of applied sport psychology. *Sport Exerc. Perform. Psychol.* 9, 87–101. doi: 10.1037/spy0000163

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Check for updates

Does the League Table Lie? The Development and Validation of the Perceived Performance in Team Sports Questionnaire (PPTSQ)

Lael Gershgoren[1], Asaf Blatt[1], Tal Sela[2] and Gershon Tenenbaum[3,4]*

[1] *School of Behavioral Sciences, The College of Management Academic Studies, Rishon LeZion, Israel,* [2] *Department of Behavioral Sciences, School of Social Sciences and Humanities, Kinneret Academic College on the Sea of Galilee, Emek HaYarden, Israel,* [3] *Ivcher School of Psychology, Interdisciplinary Center (IDC), Herzliya, Israel,* [4] *College of Education, Department of Educational Psychology and Learning Systems, Florida State University, Tallahassee, FL, United States*

OPEN ACCESS

Edited by:
*Donatella Di Corrado,
Kore University of Enna, Italy*

Reviewed by:
*Lori Dithurbide,
Dalhousie University, Canada
Itay Basevitch,
Anglia Ruskin University,
United Kingdom*

***Correspondence:**
*Lael Gershgoren
laelg75@gmail.com*

Specialty section:
*This article was submitted to
Movement Science and Sport
Psychology,
a section of the journal
Frontiers in Psychology*

Received: *07 October 2020*
Accepted: *08 December 2020*
Published: *21 January 2021*

Citation:
*Gershgoren L, Blatt A, Sela T and
Tenenbaum G (2021) Does
the League Table Lie?
The Development and Validation
of the Perceived Performance in Team
Sports Questionnaire (PPTSQ).
Front. Psychol. 11:615018.
doi: 10.3389/fpsyg.2020.615018*

Objective performance measures are vastly used in sport psychology despite their inherent limitations (e.g., unaccounted baseline differences). Founded on the nature of group goals in team sports, we aimed at developing the Perceived Performance in Team Sports Questionnaire (PPTSQ) to capture the team members' perception of their team's performance. Accordingly, three dimensions were hypothesized: *effort investment, skills execution*, and *perceived outcome*. To measure these dimensions, items were generated to address the players' perception of their team performance as a whole. Four samples of athletes were used to test the psychometric properties of the PPTSQ: professional (n = 231), collegiate (n = 222), professional—retest (n = 89), and mixed professional–collegiate (n = 139). Exploratory and confirmatory factor analyses were used to estimate construct and content validities. These procedures revealed a better data fit to a two-dimensional model that consists of effort investment and perceived outcome. The reliability analyses for the PPTSQ provide satisfactory evidence that the questionnaire is a reliable measure of perceived performance in team sport. Adequate internal consistency emerged for both dimensions ($0.75 < \omega < 0.89$). Furthermore, a high correlation was obtained for temporal stability. Concurrent validity was addressed by correlating the PPTSQ scores with the Group Environment Questionnaire and the Team Assessment Diagnostic Instrument. Correlational analysis between the PPTSQ and an objective measure of performance was used to test its predictive validity. The correlations strongly support the concurrent and predictive validities of the PPTSQ. We conclude that our perceived performance questionnaire can address various objective measures shortcomings (e.g., considering base-rate biases) resulting in a more meaningful team performance metric. Implication of the PPTSQ for sport psychology research and applied work enhancement are discussed in detail.

Keywords: sport psychology, perceived performance, objective outcome, team operation, effort, base-rate bias

INTRODUCTION

Being the end result, performance is one of the most used variables in sport psychology research. Several examples include the relationship between performance and anxiety (e.g., Craft et al., 2003), motivation (e.g., Gershgoren et al., 2011), self-efficacy (e.g., Moritz et al., 2000), and emotions (e.g., Lazarus, 2000). In team sports, the relationship between team performance and cohesion as well as Shared Mental Models (SMM) has been established both conceptually (e.g., Carron et al., 1985; Eccles and Tenenbaum, 2004, Eccles and Tenenbaum, 2007) and empirically (e.g., Gershgoren et al., 2013; Filho et al., 2014). Nevertheless, despite its empirical and methodological importance, neither a chapter on performance measures nor on perceived performance were published in the books on measurement in sport and exercise psychology of both Duda (1998) and Tenenbaum et al. (2012). This may represent a deficiency in performance measures in the sport psychology literature in general and sport psychology research in particular.

Exploring the sport domain, one can clearly identify that actual outcome measures such as win–loss percentage, points gained, and ranking are more commonly used than perceived ones. The common notion that "the league table doesn't lie," may be true in deterministic terms. However, "relative to expectations" terms can provide additional reliable and valuable data. A few anecdotal evidences are next presented to support the necessity of a valid and reliable perceived performance questionnaire in team sports being the purpose of this study.

In deterministic terms, a successful objective outcome, such as gaining 82 points out of 114 (i.e., 72% success), is better than gaining 56 points (i.e., 49% success; ESPN, 2020). However, subjectively, FC Barcelona considered the 2019–2020 season as a failure because the team failed to gain the championship from the rivalry team Real Madrid CF, although investing 255 million Euros in transferring players before the season of 2019–2020 (Transfermarket, 2020a). Consequently, two head coaches were fired from the team at that season, and its superstar, Lionel Messi, requested to leave the club. On the other hand, with only a 49% success rate, Granada CF ended the same season at the seventh place in the Spanish premier league as their most successful season since 1974 (Transfermarket, 2020b).

Another example comes from the first game of 2018 FIFA World Cup group stage E, where the national soccer team of Brazil, aspiring to win its sixth world cup title, tied with Switzerland who made effort to rebuild itself after the failure of disqualifying to the quarter final in the UEFA Euro 2016. At the post-game press conference, Switzerland's head coach, Vladimir Petkovic, noted: "we were able to do it well, and it is an excellent starting position. . . I'm very proud and pleased with the discipline and with the way we played" (BBC Sport, 2018). In contrast, Tite, Brazil's head coach, claimed: "there is some more work to be done." When asked about the matches' score, he replied, "my expectation, of course, was to get a victory, and of course, I'm not happy with the result." Having the objective result being a tie, one may ask why the head coaches of both national teams differed as much in reviewing the result of this game.

Using an objective outcome may be appropriate assuming that all teams are similar in capabilities such as physical, technical, mental, and tactical. However, this rarely is the case in sports. Indeed, all teams start the season with no points, no win–loss percentage, and with a similar ranking. Nevertheless, teams' capabilities are not equal causing objective outcomes to misrepresent base-rate information and thus teams' performances expectations. Because equal baseline cannot be presumed, perceived performance must be considered aside the teams' objective outcome and standing. Perceived performance measures are administered post-performance but may also incorporate pre-performance information such as prior expectations (i.e., was the actual outcome better or worse than expected?). The idea of unequal baseline is also reflected in sports gambling, as the winning/losing odds allocated for each team in any sports competition are rarely even. Furthermore, these odds are dynamic and are updated as the season progresses, meaning that the baseline information is complex and constantly changing. We suggest that considering inequalities among teams at baseline has meaningful consequences of how the teams' performance is conceived, and thus must be accordingly operationalized.

To account for inequality in the team's initial capacities, and consequently its performance expectations, one must consider "baseline rates" in order to evaluate and assess the teams' performances. In this vein, Kahneman and Tversky (1973) laid the foundations for investigating the human tendency to neglect a prior statistical probability in favor of the most representative case. The notion of base rate bias was developed to account for judgment predictions that violate the logic of statistical likelihood (Bar-Hillel, 1980) and is reflected in peoples' tendency to judge the likelihood of a situation without considering all relevant data. On a broader view, this notion suggests that a judgment that relies merely on "the bottom line" is fairly over simplistic and hence, may be, in cases, misrepresentative. For instance, Medvec et al. (1995) have examined the Barcelona Olympic Games and found that, despite their ranking, bronze medal winners (3rd place) expressed significantly more positive emotions than silver medal winners (2nd place) on the podium.

The Perceived Performance in Team Sports Questionnaire (PPTSQ), which we have developed herein, aims at capturing the team members' perception of their team performance at any stage throughout the season. The conceptual framework guiding the development of the PPTSQ stems from Brawley et al. (1992) findings pertaining to team goal achievement in collegiate team sports. Their results revealed that goals for competitions were effort, skill, or outcome related. Accordingly, the PPTSQ was designed to capture the team members' perception of three factors: *effort investment*, *skills execution*, and *perceived outcome*. The conceptual three-factor model of the PPTSQ is presented in **Figure 1**. To measure these factors, items were generated to address the players' perception of their team performance as a whole. This approach was supported by Feltz et al. (2008) claiming that team aspects can rightfully be measured through each member's appraisal of the team.

Because objective measures of performance may overlook fundamental differences among teams, only a moderate positive

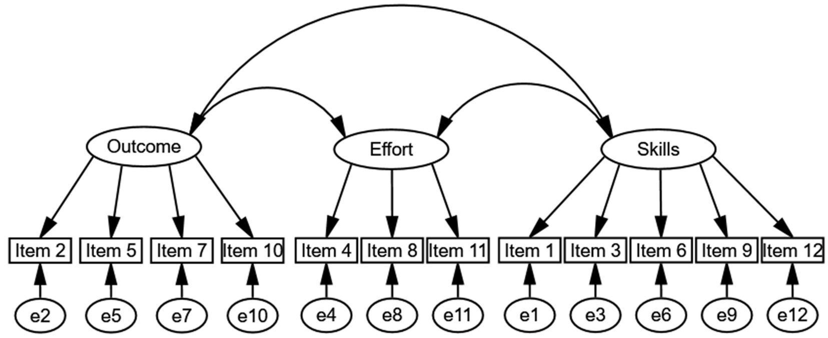

FIGURE 1 | A conceptual three-factor model of the Perceived Performance in Team Sports Questionnaire (PPTSQ).

correlation was expected among the PPTSQ dimensions and the objective performance measure. Consisting of the conceptual relationship between cohesion and performance (e.g., Widmeyer et al., 1985) and the overall effect sizes reported by the meta-analysis of Filho et al. (2014, *ES* = 0.34), a moderate and positive correlation was anticipated between the PPTSQ and the cohesion measure being used. Furthermore, stemming from Eccles and Tenenbaum's (2004, 2007) conceptual framework of shared cognitions in expert teams and the correlation ($r = 0.36$) reported by Webber et al. (2000) between SMM and performance, a positive and moderate correlation was expected between the PPTSQ and the SMM questionnaire being administered to team members.

MATERIALS AND METHODS

Participants

Four samples were used to estimate the statistical properties of the PPTSQ: professional athletes ($n = 231$), collegiate athletes ($n = 222$), professional athletes—retest ($n = 89$), and mixed professional–collegiate athletes ($n = 139$). Demographic data of these samples are presented in **Table 1**. The first sample contained both male and female ($n = 154$ and 77, respectively) professional athletes from 25 teams with an average of 9.2 athletes per team ($SD = 3.7$). The second sample, collegiate athletes, was composed of both genders (126 male athletes and 96 female athletes) participating in 32 teams ($M = 6.9$ athletes per team, $SD = 3.7$). Third, 89 professional male and female athletes ($n = 48$ and 41, respectively) from the first sample completed a retest of the PPTSQ. This sample was explored in relation to their first administration (i.e., temporal stability) and independently (i.e., as an additional confirmatory analysis). The last sample, mixed professional ($n = 62$) and collegiate ($n = 77$) athletes, also included both male and female athletes ($n = 83$ and 56, respectively) from 39 teams with an average of 3.6 athletes per team ($SD = 2.8$). This sample was used only to reconfirm the structural model and was composed of participants with missing data from samples 1 and 2. We refrained from using these participants in the original exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) analyses to perform a full-case analysis to establish and validate our structural model.

TABLE 1 | Samples and demographic data.

Sample	Sample type	*N*	Gender (Male/Female)	Age (*SD*)
Sample 1	Pro	231	154/77	24.61 (4.80)
Sample 2	Col	222	126/96	24.95 (2.97)
Sample 3—retest	Pro	89	48/41	24.82 (3.03)
Sample 4—mixed	Pro + Col	139	83/56	23.80 (3.59)

Pro, Professional level; Col, College level.

Four criteria and justifications were employed for participation in this study. First, participants were active members in professional or collegiate level teams from various sports such as soccer, basketball, futsal, volleyball, rugby, water polo, and team handball. Second, all the athletes were 18 years of age and above. Team strategies and tactics are mostly trained and best acquired at least at the late stages of adolescence (i.e., collegiate and professional levels). Third, to provide meaningful data pertaining to team performance, team members were required to share three or more competitive experiences (i.e., played at least three games together). Last, in accordance with the institutional review board (IRB) committee guidelines, participants were, regardless of their original nationalities, active members of teams that competed in Israeli leagues/tournaments.

Instrumentation

Three questionnaires were administered in this study. The PPTSQ was administered to examine the members' perceptions pertaining to various performance-related components. In addition, team cohesion and SMM questionnaires were administered to validate the PPTSQ. Furthermore, objective data regarding the team's outcome scores were collected. To capture the sample's characteristics, a demographic form was administered.

Since this study took place in Israel, in which Hebrew is the native spoken language, a back-translation procedure was conducted. Hence, all instruments were translated to Hebrew and then back-translated to English. No noteworthy misfits were observed. These translations were made by two separate individuals who were very familiar with both the Hebrew and the

English languages. A back-translation procedure is commonly used in such instances and is supported in the extant literature (see Brislin, 1986). Eventually, both the English and the Hebrew versions were available to the athletes and were provided according to their request.

Perceived Performance in Team Sports Questionnaire

The development and validation of the PPTSQ followed a commonly used procedure (e.g., Chatterji et al., 2002; Johnson et al., 2007) and consisted of six stages.

Items Generation. This stage involved a thorough literature review pertaining to performance measures in sport. In addition, an examination of team-related goals afforded the emergence of a performance conceptual framework that was used to generate 12 items under three factors pertaining to performance perception in team sports. Accordingly, the *effort investment* scale included items 4, 8, and 11; the *skills execution* scale contained items 1, 3, 6, 9, and 12; and the *outcome* scale included items 2, 5, 7, and 10. Items were scored on a 5-point bipolar Likert scale where each item contained two unique contradicting/opposite statements at the continuum ends and a statement at its mid-point. All the items of the original PPTSQ are presented in **Supplementary Appendix A**.

Content Validity. New measures must adequately capture the variable under examination eliminating superfluous content (Hinkin, 1995). Hence, the items' *relevance* and *representativeness* to the domain of interest are at the forefront of content validity assessment (Vaughn and Daniel, 2012). This judgment procedure is usually performed by experts in the domain of interest. In the current study, content/face validity of the PPTSQ was obtained by two respected scholars in the sport psychology domain in general and in team processes in particular.

Descriptive Statistics. Descriptive statistics aims at calculating central tendencies and distribution variables (i.e., mean, *SD*, skewness, and kurtosis) of each item, items within factors, and the entire questionnaire. This stage was implemented following data collection.

Construct Validity. Construct validity is aimed at accounting for as much as possible the distinctive variance of each of the variables while grouping them into clusters. Because the PPTSQ does not rely on an *a priori* model but only on a conceptual framework, the model fit to the data was examined through an EFA procedure followed by several confirmatory procedures.

Reliability. Reliability (i.e., internal consistency) was obtained using McDonald's omega (ω) for each scale. Recently, Hayes and Coutts (2020) have presented McDonald's omega (ω; see also McDonald, 1999; Hancock and An, 2020) as a superior reliability measure to the traditional Cronbach's alpha (α). Temporal stability was also estimated, as the PPTSQ was readministered to a subset sample of 89 professional athletes with an average of 8 days apart between the two administrations.

Concurrent and Predictive Validity. Concurrent validity was addressed by correlating the PPTSQ scores with the Group Environment Questionnaire (GEQ; Carron et al., 1985). Furthermore, concurrent validity was also tested by correlating the PPTSQ with the Team Assessment Diagnostic Instrument (TADI; Johnson et al., 2007), which is an SMM questionnaire. Correlational analysis between the PPTSQ and an objective measure of performance was used to test its predictive validity.

The Group Environment Questionnaire

The GEQ (Carron et al., 1985) was designed to measure group cohesiveness. The conceptual framework, which guided the development of the GEQ, was derived from the notions of task and social cohesions stemming from individual and team level perspectives (see Widmeyer et al., 1985). Social cohesion pertains to the degree to which group members are bonded or close to one another while interacting socially. Task cohesion refers to the degree to which group members remain united while striving to complete their task or achieve their performance-related goals (see Festinger et al., 1963; Carron, 1984).

Four subscales are used to capture team cohesion: *interpersonal attraction—social (ATGS)*; *interpersonal attraction—task (ATG-T)*; *group integration—social (GI-S)*; and *group integration—task (GI-T)*. The interpersonal attraction subscale consists of nine items (i.e., five social cohesion items and four task cohesion items). The group integration subscale also includes nine items (i.e., four social and five task cohesion items). All items are scored on a 9-point Likert-type scale ranging from *strongly disagree* (1) to *strongly agree* (9). Twelve items (i.e., 1, 2, 3, 4, 6, 7, 8, 11, 13, 14, 17, and 18) are negatively worded and reversed in scoring.

Cronbach alpha coefficients for the ATG-S, ATG-T, GI-S, and GI-T were found to be $r = 0.64$, 0.75, 0.76, and 0.70, respectively (Carron et al., 1985). Using CFA, Li and Harmer (1996) have confirmed the four-factor model of the GEQ as initially suggested by Carron et al. Criterion validity of the GEQ was obtained by identifying a moderate correlation with the team perception scale of the Sport Cohesiveness Questionnaire (Martens et al., 1971) and moderate correlations between the task subscales in team sports and the roles related items of the Team Climate Questionnaire (Grand and Carron, 1982). As predicted, a non-significant correlation was obtained between the GEQ and the Bass Orientation Inventory (Bass, 1962), suggesting that one's personal motivation is unrelated to one's appraisal of his/her team's cohesion. In conclusion, the GEQ is by far the most widely used cohesion questionnaire in sport (Carron et al., 2002).

Team Assessment Diagnostic Instrument

The TADI (Johnson et al., 2007) is a measure of team SMM content. Five team processes (i.e., SMM factors) emerged from the EFA analysis accounting for 82% of the variance: *General Task and Team Knowledge* (GTTK), *Communication Skills (CS)*, *Attitude Toward Teammates and Task* (ATTT), *Team Dynamics and Interactions* (TDI), and *Team Resources and Working Environment* (TRWE). The Cronbach's alpha coefficients of the SMM factors (i.e., 0.76, 0.89, 0.75, 0.81, 0.85, respectively) suggest adequate reliability (i.e., internal consistency). Furthermore, at a later stage of their study, Johnson et al. have verified the five-structure model through a CFA procedure.

The TADI includes 15 items in total, 3 for each team process, rated on a 5-point Likert scale. Items response is anchored by *strongly disagree* (1) and *strongly agree* (5). A team score is calculated by averaging all the items' scores for each team member followed by averaging all the team members' mean scores. This score was interpreted as the perception of the team members pertaining to their overall level of SMM. Higher mean score indicates higher levels of sharedness. Similar mathematical reasoning was used for each team process independently. Such calculation afforded capturing the sharedness level of each of the SMM factors separately.

A Standardized Objective Performance Score

Because our teams performed in different types of competitions (i.e., league vs. tournament) and under different scoring systems (e.g., soccer where a win grants 3 points and a loss grants 0 vs. basketball where a win grants 2 points and a loss grants 1 point), the performance scores were standardized. Thus, the objective outcome score was calculated as the percentage of points earned from the maximum possible points when a loss, a tie, and a win grant 0, 1, and 2 points, respectively. For example, if a futsal team won three games, tied two, and lost two, a total of 8 points was considered. This value was transferred to a percentage value resulting in a 57.1% (8 points earned divided by 14 possible multiplied by 100) success rate. Success rate values were calculated only for the official records of the professional leagues/tournaments.

Procedure

All teams completed the questionnaire at the end of a tournament or the end of their regular season. At first, permission to approach the athletes was obtained from a team representative (e.g., athletic director, head coach, team's owner). Then, a convenient time for administration was coordinated with the head coach either at the tournament venue or at the team's site. When repeated measures took place, both dates were scheduled in advance with approximately 1 week apart. Following the IRB-approved protocol, informed consents were obtained from the athletes, and confidentiality was verbally announced prior to data collection. Specifically, participants were informed that only the researchers will have access to the players' responses. The athletes were asked to complete the demographics form first followed by a battery of questionnaires including the PPTSQ, GEQ, and the TADI. A counter-balanced order was used to eliminate order effects due to fatigue or loss of interest, which can negatively affect the response quality and consequently increase the measurement error. The retest administration included only the PPTSQ. Prior to the retest, confidentiality was announced again. At the end of each administration, once the athletes completed the questionnaire, they were thanked and released.

Statistical Analysis

The analytic approach consisted of several stages. First, using sample 1 (professional athletes), we fitted different exploratory models. We started with a three-factor model based on our conceptual framework that included 12 items composing three latent factors: effort investment, skills execution, and perceived outcome. We evaluated this model alongside other models using Kline's (2015) fit indices: Standardized Root Mean square Residual (SRMR < 0.08), Adjusted Goodness of Fit (AGFI ≥ 0.90), Comparative Fit Index (CFI ≥ 0.90), Root Mean Square Error of Approximation (RMSEA < 0.08), and Bayes Information Criterion (BIC), to establish the best-fitted model. Second, after establishing the best-fitted model, we reexamined this model by a CFA using three additional samples (college-athletes sample, professional athletes retest sample, and mixed sample). Next, we used bivariate correlations to estimate the concurrent validity between PPTSQ factors and TADI and GEQ scores. Finally, the predictive validity of the PPTSQ was obtained by correlating its factors' scores with the scores of the standardized objective performance measure. Reliability (i.e., internal consistency) was obtained using McDonald's omega (ω; Hayes and Coutts, 2020). Omega estimates reliability more accurately than Cronbach's alpha when the tau-equivalence assumption (equal factor loadings of all items) is violated (as frequently is the case; Trizano-Hermosilla and Alvarado, 2016). Furthermore, Cronbach's alpha can be high even if a set of items measures more than one construct because it only *assumes* a construct's unidimensionality (Green et al., 1977; Graham, 2006) but fails in measuring it directly as such. Omega, on the other hand, overcomes these limitations and does not rely on the tau-equivalence assumption and consist of a single-factor solution extracted from a factor analysis (see Hancock and An, 2020; Hayes and Coutts, 2020). Temporal stability was estimated via Pearson correlation. Data were analyzed using IBM SPSS Statistics for Windows, Version 25.0 (IBM Corp., Armonk, NY) and IBM SPSS Amos Version 24.

RESULTS

The development and validation of the PPTSQ followed a six-stage procedure. These stages included item generation, content validity, descriptive statistics, construct validity, reliability, and concurrent and predictive validity. The first two stages are related

TABLE 2 | Descriptive statistics of the original PPTSQ total score, scales, and items.

Scale	Mean	*SD*	Skewness	Kurtosis
Total score	3.21	0.86	−0.32	−0.36
Effort investment	3.18	0.91	−0.26	−0.44
Skills execution	2.94	1.06	−0.19	−0.71
Perceived outcome	3.61	0.91	−0.52	0.15
Item 1	3.39	1.10	−0.50	−0.24
Item 2	2.90	1.27	0.00	−0.94
Item 3	3.16	1.16	−0.36	−0.50
Item 4	3.43	1.08	−0.41	−0.13
Item 5	3.19	1.19	−0.35	−0.56
Item 6	2.97	1.15	−0.18	−0.74
Item 7	2.87	1.15	−0.08	−0.67
Item 8	3.74	1.06	−0.59	−0.12
Item 9	3.10	0.99	−0.28	−0.19
Item 10	2.81	1.18	0.03	−0.72
Item 11	3.64	1.10	−0.55	−0.23
Item 12	3.26	0.98	−0.27	−0.03

to the development of the PPTSQ and were conducted before data collection. These stages were elaborated on in the method section and, hence, will not be repeated in this section.

Descriptive Statistics

Prior to testing the study's hypotheses, central tendencies and distribution statistics (i.e., mean, *SD*, skewness, and kurtosis) of all the PPTSQ items and scales were examined in the first sample (professional level; $N = 231$; see **Table 2**). Items' mean ranged from 2.81 (item 10) to 3.74 (item 8); *SD* ranged from 0.98 (item 12) to 1.27 (item 2). Skewness and kurtosis values for all the items ranged between –0.94 and 0.03, suggesting no substantial deviations from normality in the item-response distributions.

Construct Validly (Convergent and Discriminant Validity)

Several EFA models were tested to identify the best fit to the data (see **Table 3** models 1–3). First, we fitted a three-factor model based on theoretical foundation that included 12 items composing three latent factors: effort investment, skills execution, and perceived outcome. As depicted in **Table 3**, this model failed to share an adequate fit to the data. Furthermore, the correlations between skills execution and the other two factors were very high (0.91 with the perceived outcome factor and 0.77 with effort investment). Thus, a two-factor model that included only the perceived outcome and effort investment factors (see **Table 3**, model 2) was tested. This model demonstrated a good fit to the data. Nonetheless, the omission of one item (item 5) enhanced model fit indices (see **Table 3**, model 3). This model was, then, determined as the best-fitted model to the data.

Figure 2A presents the EFA model, which includes two factors, effort investment and perceived outcome. Furthermore, the structural model was then confirmed using three additional samples. **Table 3** and **Figure 2B** present the CFA of a two-factor model based on sample 2—the college sample. This model demonstrated a very good fit to the data. Lastly, the two-factor model was confirmed by two separate samples: (a) the professional athletes retest sample and (b) the mixed professional–collegiate sample (see **Table 3** and **Figures 2C,D**, respectively). **Table 4** presents the descriptive statistics of the PPTSQ total score and scales for models 1–4. Means ranged from 2.86 (perceived outcome, model 1) to 3.61 (effort investment, model 1); *SD* ranged from 0.66 (total score, model 4) to 1.09 (perceived outcome, model 1). Skewness and kurtosis coefficients for the total score and the scales ranged between −0.70 and 1.1. Thus, the assumption of distribution normality was conformed.

Considering the entire study's sample (363 male and 229 female athletes), we examined possible gender differences for mean PPTSQ scores for both PPTSQ dimensions using independent samples *t*-tests. Non-significant gender difference emerged for male and female athletes in effort investment scores ($M = 3.54$, $SD = 0.89$ vs. $M = 3.50$, $SD = 0.80$, respectively), $t(590) = 0.553$, $p = 0.58$, two-tailed. Moreover, non-significant gender difference was obtained for perceived outcome ($M = 2.89$, $SD = 1.00$ vs. $M = 2.94$, $SD = 0.84$, respectively), $t(543.565) = –0.622$, $p = 0.53$, two-tailed. Finally, non-significant gender effect for PPTSQ total score was revealed ($M = 3.22$, $SD = 0.81$ vs. $M = 3.22$, $SD = 0.70$, respectively), $t(537.554) = –0.60$, $p = 0.95$.

Reliability

Table 5 presents the McDonald's omega coefficients for effort investment and perceived outcome. Omega coefficients were adequate across all the samples ranging from 0.75 to 0.84 for effort investment and from 0.78 to 0.89 for perceived outcome. In addition, temporal stability for both factors was good (effort investment dimension: $r = 0.85$, $p < 0.001$; perceived outcome dimension: $r = 0.80$, $p < 0.001$). Item-scale correlations were very high (ranged from 0.79 to 0.92), demonstrating a very strong item-scale relationship. Inter-item correlations ranged from 0.42 to 0.77, suggesting both inter relationship and differentiation among the items within the scales. Item-scale and inter-item correlations for each scale are separately presented in **Tables 6**, **7**. Considering the entire study's sample ($N = 592$; samples 1, 2, and

TABLE 3 | Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) structural equation modeling (SEM) models for the Perceived Performance in Team Sports Questionnaire (PPTSQ)—attached in a separate file.

	Model	Items	Sample	*N*	Analysis type	Chi-square (CMIN)	DF	Probability level	CMIN/DF	RMR	SRMR	GFI	AGFI	CFI	RMSEA	BIC
1	Model 1: 3-factor model	12	Pro, Sample 1	231	EFA	165.562	51	p < 0.001	3.246	0.056	0.047	0.891	0.833	0.943	0.099	312.507
2	Model 2: 2-factor model	7	Pro, Sample 1	231	EFA	30.391	13	*p* < 0.005	2.338	0.042	0.0329	0.965	0.925	0.981	0.076	112.027
3	Model 3: 2-factor model*	6	Pro, Sample 1	231	EFA	15.482	8	*p* = 0.05	1.935	0.028	0.022	0.979	0.944	0.989	0.064	86.234
4	Model 3	6	Col, Sample 2	222	CFA 1	14.578	8	*p* = 0.068	1.822	0.024	0.0248	0.979	0.946	0.987	0.061	84.813
5	Model 3	6	Pro, Re-test, Sample 3	89	CFA 2	13.32	8	*p* = 0.101	1.665	0.048	0.056	0.957	0.887	0.98	0.087	71.672
6	Model 3	6	Pro + Col, Sample 4	139	CFA 3	10.575	9	*p* = 0.306	1.175	NA				0.99	0.036	NA

*EFA and CFA SEM Models for the PPTSQ.*Best EFA fitted model; SRMR, Standardized root mean square residual; AGFI, Adjusted Goodness of Fit; CFI, Comparative Fit Index; RMSEA, Root Mean Square Error of Approximation; BIC, Bayes Information Criterion. Pro, Professional level, Col, College level.*

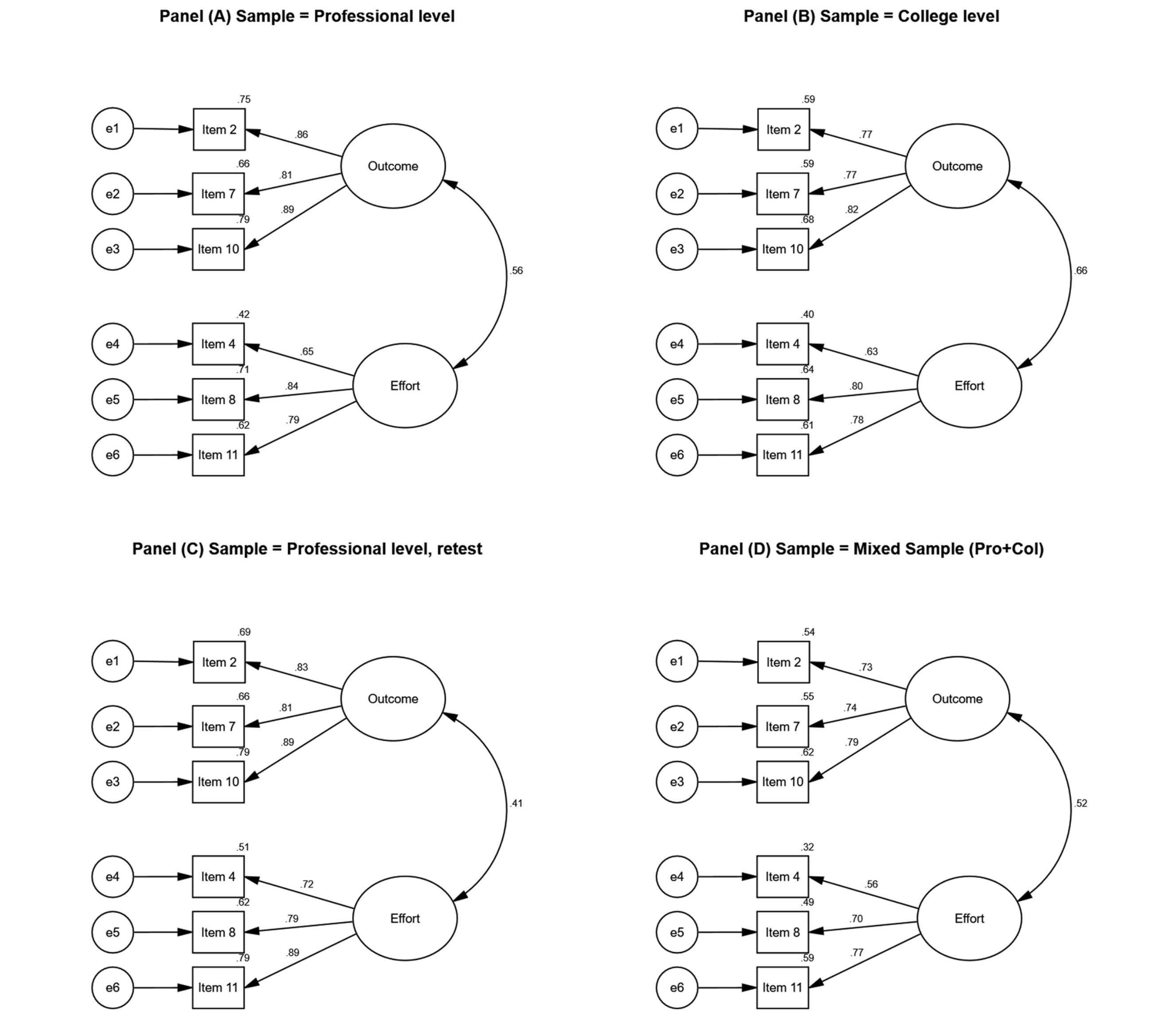

FIGURE 2 | Best fitted model (Model #3 - 2-factor model) for samples 1–4. Standardized estimates are presented alongside squared multiple correlation for each item (upper left corner). All estimates are significant at the $p < 0.001$. **Panel (A)** presents results from the EFA analysis (Professional sample, $N = 231$); **Panel (B)**, presents results from the CFA analysis (College sample, $N = 222$). **Panel (C)** presents results from the CFA analysis (Professional retest sample, $N = 89$); **Panel (D)**, presents results from the CFA analysis (mixed sample, $N = 139$).

4), items composing each factor were highly correlated with the total factor score ranging from 0.8 to 0.9 and only moderately related to the total score of the other factor, with correlations ranging from 0.35 to 0.46.

Concurrent and Predictive Validity

The correlations presented in **Table 8** indicate that the PPTSQ and its two dimensions correlated significantly with TADI and GEQ scores. The PPTSQ total score was moderately correlated with the team cognitive measure (i.e., TADI; $r = 0.56$). Moderate correlative pattern was evident between the PPTSQ total score and all the TADI scales (GTTK, $r = 0.47$; CS, $r = 0.46$, ATTT, $r = 0.48$; TDI, $r = 0.46$; TRWE, $r = 0.42$). A low, yet significant, correlation was revealed between the PPTSQ and the team social construct of cohesion (i.e., GEQ; $r = 0.24$). Within the GEQ, a higher correlation was found between the PPTSQ and the GEQ group-integration scales ($r = 0.36$ with both the task and social scales) than with the attraction-to-group scales (task, $r = 0.12$; social, $r = 0.16$). Separately, the effort investment dimension moderately correlated with the TADI ($r = 0.58$), while the perceived outcome dimension demonstrated a slightly lower, yet still moderate, correlation ($r = 0.44$) with it. Similarly, the effort investment dimension shared a higher correlation with the GEQ then the Perceived outcome factor ($r = 0.27$ and 0.17, respectively). These findings support the concurrent validity of the PPTSQ measure

TABLE 4 | Descriptive statistics of the modified Perceived Performance in Team Sports Questionnaire (PPTSQ) scales and total scores.

Sample #	Scale	Mean	*SD*	Skewness	Kurtosis
1	Effort investment	3.61	0.91	−0.52	0.15
	Perceived outcome	2.86	1.09	−0.08	−0.70
	Total score	3.23	0.86	−0.27	−0.14
2	Effort investment	3.47	0.85	−0.13	−0.24
	Perceived outcome	2.98	0.87	0.20	0.11
	Total score	3.23	0.75	0.07	−0.04
3	Effort investment	3.46	0.82	−0.54	0.93
	Perceived outcome	2.92	0.78	−0.10	0.58
	Total score	3.19	0.66	0.07	1.10
4	Effort investment	3.48	0.74	−0.18	0.36
	Perceived outcome	2.88	0.77	−0.36	0.13
	Total score	3.18	0.63	−0.15	0.13

TABLE 5 | McDonald's omega coefficient (ω) for the Perceived Performance in Team Sports Questionnaire (PPTSQ) scales in the four samples.

	Sample			
Scale	**1**	**2**	**3**	**4**
Effort investment	0.81	0.78	0.84	0.75
Perceived outcome	0.89	0.83	0.88	0.78

TABLE 6 | Item-scale and inter-item correlation matrix for the effort investment scale.

Sample		Item 4	Item 8	Item 11
1	Item 4	1		
	Item 8	0.539**	1	
	Item 11	0.507**	0.671**	1
	EI total score	0.806**	0.869**	0.861**
2	Item 4	1		
	Item 8	0.497**	1	
	Item 11	0.490**	0.628**	1
	EI total score	0.809**	0.845**	0.842**
3	Item 4	1		
	Item 8	0.523**	1	
	Item 11	0.651**	0.707**	1
	EI total score	0.850**	0.842**	0.907**
4	Item 4	1		
	Item 8	0.417**	1	
	Item 11	0.475**	0.581**	1
	EI total score	0.788**	0.797**	0.851**

$^{**}p < 0.001$.EI, effort investment.

and its dimensions. Interestingly, using a partial correlation analysis, we found that, when controlling for the perceived outcome scores, the correlation between TADI scores and objective performance diminished ($r = 0.02$, $p = 0.80$), as well as the correlation between GEQ and objective performance ($r = -0.003$, $p = 0.97$).

To test its predictive validity, the PPTSQ was correlated with an objective performance score. A moderate correlation ($r = 0.47$) was identified between these measures. Moreover,

TABLE 7 | Item-scale and inter-item correlation matrix for the perceived outcome scale.

Sample		Item 2	Item 7	Item 10
1	Item 2	1		
	Item 7	0.697**	1	
	Item 10	0.769**	0.723**	1
	PO total score	0.914**	0.886**	0.917**
2	Item 2	1		
	Item 7	0.610**	1	
	Item 10	0.628**	0.623**	1
	PO total score	0.872**	0.846**	0.874**
3	Item 2	1		
	Item 7	0.685**	1	
	Item 10	0.729**	0.727**	1
	PO total score	0.909**	0.881**	0.907**
4	Item 2	1		
	Item 7	0.513**	1	
	Item 10	0.565**	0.544**	1
	PO total score	0.833**	0.816**	0.849**

$^{**}p < 0.001$. PO, perceived outcome.

TABLE 8 | Pearson correlations between Perceived Performance in Team Sports Questionnaire (PPTSQ), Group Environment Questionnaire (GEQ), Team Assessment Diagnostic Instrument (TADI), and Objective Performance for samples 1 and 2 (combined; *N* = 453).

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Mean PPTSQ	1					
(2) Effort investment	0.914**	1				
(3) Perceived outcome	0.918**	0.678**	1			
(4) Mean GEQ	0.240**	0.271**	0.170**	1		
(5) Mean TADI	0.556*	0.585**	0.436**	0.288**	1	
(6) Objective performanceˆ	0.467*	0.310**	0.527**	0.174*	0.258**	1

ˆN = 180 for correlation between objective performance and all other variables.$^{*}p < 0.05$, $^{**}p < 0.001$.

a moderate correlation ($r = 0.53$) was obtained between the perceived and the objective outcome scores. Such a correlation suggests that these variables are related yet distinct. Finally, effort investment correlated low-moderately ($r = 0.31$) with the objective outcome score. Noteworthy, the PPTSQ has demonstrated stronger relationships with the SMM measure (the TADI; $r = 0.56$) than the objective performance measure ($r = 0.26$). These results were also evident with the cohesion measure. The PPTSQ shared a higher correlation with the GEQ ($r = 0.24$) than the objective measure ($r = 0.17$).

DISCUSSION

It is often assumed that athletes' perceived performance is equivalent to the objective outcome. However, psychological mechanisms of information processing are involved in players' appraisals of their teams' operations (e.g., Micklewright et al., 2009). These mechanisms are idiosyncratic and consist of relative analyses of information that are abstract in nature such as perceived effort (e.g., did the players exert sufficient amount of effort?) or absolute such as the score (e.g., are the players happy with this score considering the opponent's skill level or how has the game unfolded?). Yet, only a few scientific studies have addressed this enigmatic notion.

In the current study, we challenged this traditional belief, where the objective performance score merely determines perceived performance in team sports. This study centered on the development and validation of the PPTSQ, a self-report measure that assesses team sports members' perceptions about previous team performances. The PPTSQ development consisted of a phenomenological analysis that considered the nature of team goals in sports teams (Brawley et al., 1992) and by each player's perceptions of his/her team performance as a whole (Feltz et al., 2008).

Specifically, the EFA of the PPTSQ retained six items (the final version of the PPTSQ and its six items are presented in **Supplementary Appendix B**) and revealed a two-factor structure: *effort investment* and *perceived outcome*. The factor of skills execution was omitted due to its very high correlation with the perceived outcome factor. Confirmatory factor analyses of three samples confirmed the PPTSQ's EFA factor structure obtained by several fit indices (e.g., AGFI, SRMR; Kline, 2015). Importantly, the structure of the PPTSQ was confirmed on different samples of professional and collegiate athletes, supporting the generalizability (i.e., external validity) of the final structural dimensions of the measure.

Overall, the reliability analyses of the PPTSQ provide satisfactory evidence that the questionnaire is a reliable measure of perceived performance in team sport. Both internal consistency and temporal stability were examined to support the PPTSQ reliability. High internal consistency emerged for both factors of perceived effort investment and perceived outcome ($0.75 < \omega < 0.89$[1]). Furthermore, a high correlation was obtained for temporal stability, with an average of 8 days between administrations. The temporal stability coefficients of both the perceived effort investment scale and the perceived outcome scale were deemed sufficient exceeding 0.80 (see Vaughn and Daniel, 2012).

Finally, the PPTSQ and its dimensions, perceived effort investment and perceived outcome, correlated positively and significantly with the TADI (Johnson et al., 2007) and the GEQ (Carron et al., 1985) scores. Specifically, as hypothesized, the PPTSQ and its two dimensions correlated moderately or higher with TADI. Significant and positive association, although slightly lower than hypothesized, emerged between the PPTSQ total score and its dimensions and GEQ.

The overall results of these correlational analyses support the concurrent validity of the PPTSQ. In addition, the PPTSQ was correlated with the objective measure of performance to support its predictive validity. Because the objective measure of performance was assumed to overlook fundamental differences on the one hand and play an important role on the other, a moderate correlation was expected between the introspective and objective performance measures. Indeed, the objective measure of performance demonstrated a moderate correlation with the PPTSQ total score. Yet, a slightly lower correlation was identified for the effort investment dimension, and a slightly higher correlation was obtained for the perceived outcome dimension. Overall, these findings support the predictive validity of the PPTSQ and, at the same time, emphasize its unique contribution to performance measurement.

Theoretical Considerations

Our study pertained to how athletes perceive the performance of their teams. The "base rate" of the team (e.g., team capacities) plays a major role in the establishment of performance expectations; yet, in performance evaluation, this factor was overlooked. Herein, Base Rate Fallacy is a human tendency to neglect the base rates of a case in favor of the event-specific information (Bar-Hillel, 1980). For instance, million people play the lottery despite the scant odds against winning the jackpot. In a similar vein, in sport, one can assume that a draw in a game represents an equivalent performance satisfaction for both contenders. However, such performance satisfaction is often determined by a stochastic concept rather than a deterministic one. Stochastic analyses consider several variables (e.g., squad, budget, expectations, how the game unfolded, etc.) and, although cognitively complex, are performed by athletes and coaches almost automatically. This, almost automatic, analysis pattern of relative, multidimensional and complex information is familiar to athletes, as it characterizes their on-field decision-making processes (Tenenbaum, 2003). The controversy of judging sport's outcome calls for theoretical principles and empirical data that concentrate on the mechanisms that account for this phenomenon. Therefore, the PPTSQ was proposed to examine team players' appraisals following actual competitions. Overall, the current study was aimed at establishing a perceived performance questionnaire in sport and examining its psychometric properties within adequate samples.

Following the EFA procedure, the modified PPTSQ contained two scales, *perceived outcome* and *effort investment*. The former factor incorporates the expectations and the objectives of the athletes concerning team performance. Consistent with attribution theory (Weiner, 1974), these aspects are more related to one's appraisal of ability and task difficulty and, thus, are relatively external and only partially under one's control. Yet, they are vital to perceived performance, as they encompass the athlete's initial expectations and appraisals of fundamental differences. The second dimension captures effort investment properties, such as commitment and intensity. These properties are relatively internal and under the player's control (Weiner, 1974) although they can also be compared to others. Obviously, these two factors, although distinct, are interrelated. Team members who met or exceeded their performance expectations reported being highly focused and committed; vice versa, falling short of team objectives resulted in a lower perceived effort (Gould et al., 1999). Indeed, the results of the current study revealed a strong correlation (r = 0.68) between the PPTSQ components.

[1]There is no universally accepted guideline for acceptable or adequate omega reliability (nor to alpha; Lance et al., 2006). The ω coefficient should meet the same standards as alpha coefficients and a minimum of 0.75 would be preferred although this is merely a subjective guideline as Reise et al. (2013) suggested.

However, such a correlation, although high, suggests that each dimension contributes uniquely and sufficiently to a complete understanding of the athlete's perception of his/her performance. This conclusion is further supported by the different correlation each dimension shared with the GEQ, TADI, and objective performance measures.

Pertaining to the deletion of the *skills execution* dimension, the results indicated an extremely high correlation ($r > 0.90$) between skill execution and perceived outcome. It appears that athletes already consider skill abilities in their initial expectations and, hence, assimilate them into their appraisal of the outcome. Further work is required to verify if these two dimensions are confounded or if there are some conditions in which it is possible to dissociate them.

Interestingly, this study revealed that the PPTSQ is tied more closely to the TADI than to the GEQ. This finding indicates that perceived performance is more associated with SMM, the cognitive aspect of team performance, rather than with cohesion that, as aforementioned, is inherently a social oriented psychological construct. SMM relates to collective cognitive schemas that dynamically govern team coordination through team-related decision-making processes. Performance wise, these processes enable synchronization among the players' actions under specific task demands (Cannon-Bowers et al., 1993; Tenenbaum and Gershgoren, 2014).

Concerning predictive validity, both the TADI and the GEQ demonstrated stronger correlations with the introspective measure of performance than the objective outcome. Since team coordination (e.g., superior SMM) and cohesion underly team performance (Gershgoren et al., 2013; Filho et al., 2014), our new measure of performance may be considered a more suitable indicator of team performance. Furthermore, when controlling for the variable *perceived performance*, the statistical relationship between the objective measure and either the TADI or the GEQ diminished. Altogether, these findings suggest that the two performance measures, objective and introspective, are confounded.

Akin with our postulations, outcome result must be addressed as context dependent (e.g., ranking); otherwise, it may be misleading. To exemplify, assume that a cohesive underdog team ties a superior non-cohesive team. The high perceived performance score of the underdog team and the low perceived performance score of the superior team will provide a correlative support to the cohesion-performance linkage that otherwise could have been remained unnoticed.

When addressing performance measurement, one must distinguish between expected and unexpected competitive outcomes. In the case of expected results (for instance, a high-ranked team dominates the scoreboard against a low-ranked team or when two relatively equal-ranked teams tie), both objective and introspective measures are reliable indicators of performance level. However, under unexpected results (e.g., when a low-ranked team ties or beats a high-ranked team, or when one team dominates a relatively equal-ranked team), only the introspective measure represents reliably the performance level of both teams. That is, the introspective measure of perceived performance remains reliable under both expected and unexpected conditions. This pattern suggests that the objective performance measure is inherent within the broader concept of perceived performance.

We believe that the PPTSQ provides an additional important contribution in facilitating group metacognition (Hinsz, 2004). Teammates engage in post-process coordination that comprises metacognition behaviors pertaining to the team performance. These include, among others, verbal discussions and video analyses (Eccles and Tenenbaum, 2004). Post-event reflection (Chow and Luzzeri, 2019) is a post-competition evaluation procedure that aims to enhance self-monitoring, self-correction, and emotional regulation skills among team athletes. Indeed, teams that engaged in task execution monitoring followed by future strategies establishment demonstrated improved performance compared to teams who did not participated in such processes (Rasker et al., 2000).

In a sense, the PPTSQ requires the subject to perform a metacognitive evaluation process on a team level. The athlete is required to reflect on how the team performed during competitions in terms of effort and whether the outcome was in line with prior expectations. Metacognition refers to a person's knowledge of his cognitive system (Flavell, 1979). It involves reflective thinking in which the thought process itself becomes the object of observation. Different theoretical accounts put forward two main metacognitive processes: *monitoring* and *control.* Monitoring involves a subjective assessment of the quality of task performance, while control is the decision made following monitoring (Nelson and Narens, 1990). We suggest that the two dimensions retained from the PPTSQ, effort investment and perceived outcome, reflect the two metacognition processes. Effort investment is related to the monitoring processes, since the athlete reflects on how the team functioned during the competition in terms of effort, intensity, and commitment. On the other hand, the perceived outcome dimension is related to metacognitive control. Here, the athlete is faced with outcome items and he/she is forced to decide explicitly about the end result compared to prior expectations. This task demands active monitoring over information (i.e., control) rather than merely monitoring.

Limitations, Future Directions, and Implications

In self-serving bias, cognitive processes may be impaired by one's need to preserve self-esteem (Campbell and Sedikides, 1999). Thus, under some circumstances, an unreliable response may be produced to rationalize or justify a certain outcome. As the PPTSQ aims at capturing high-order cognitive processes, the usefulness of the responses being provided in this study might be questioned. However, the reliability and validity of the PPTSQ provide evidence that professional and collegiate athletes can truly report on their team's cognitive processes, possibly because reaching higher levels of performance requires strong abilities of analyzing cognitive processes (Breivik, 2013). To fully capture the utility of a subjective performance inventory, future PPTSQ studies must incorporate an inclusive approach that considers samples from novel or moderate skill levels as well as youth populations. Furthermore, future research may include additional performance measures (e.g., GPS and video analyses) alongside the PPTSQ and the outcome score. This line of investigation is especially prominent for validating the PPTSQ for a single match or a specific tournament or league.

From a methodological standpoint, the dataset was mostly based on Israeli participants, limiting the generalizability of the results. While the Israeli culture is mostly western, the use of the PPTSQ should be examined in other western cultures, such as in the US and western Europe, as well as in Asian, Eastern European, African, and other cultures. From a linguistic standpoint, although the common back-translation procedure (Brislin, 1986) was employed to accurately translate the PPTSQ, the questionnaire being directly validated in this study is the Hebrew version of the PPTSQ. Consequently, the original, English version of the PPTSQ should be examined. Such an examination can be conducted in future research.

Future directions may address the need for a valid and reliable measure for individual sports. As individual sports differ significantly from team sports in various psychological aspects (e.g., anxiety level), such a measure can facilitate the understating of one's experience as he/she performs individually. Moreover, this measure can be used to better comprehend the relationship between psychological variables and individual performance.

Despite the prevalence of teams in the exercise and sport contexts, there is relatively little research involving teams' mechanisms compared with other psychological topics in our field (Eys et al., 2019). Thus, the development of a validated tool that assesses the performance appraisals of sports teams intertwines the emotions and the pressures that exist in the competitive sport environment. These findings support the notion that addressing sport subjective performance inventory uncovers some blind spots in the understanding of team performance. Therefore, the literature of sport and exercise psychology can enrich its competencies. For instance, sport psychology scholars and practitioners can operationalize subjective team performance knowledge via the PPTSQ. Likewise, the completion of post-performance inventory can tailor the delivery of mental sessions following sport events and facilitate the preparation for upcoming competitions.

The development of the PPTSQ and the importance of perceived performance measure alongside an objective one supports an applied line of research on the importance of resourcefulness (e.g., Kennett, 1994). Thus, resourcefulness skills can enrich one's adaptive abilities (i.e., tailoring appropriate solutions and utilizing self-management techniques) with minimal to no expenses (Goff, 2011). In cases of resourcefulness, the team can overcome barriers and exceed expectations (i.e., positive subjective performance evaluation) even if other, more resourced teams, do objectively better (e.g., are ranked higher).

Metacognition can facilitate team information processing in distinguishing good from poor information (Hinsz, 2004). In the long run, group-level metacognitive deliberation, as offered by the PPTSQ, can benefit team performance, as players and practitioners can learn from experience and evolve from successes and failures. The positive correlation between PPTSQ and TADI scores reflects this idea. Yet, further effort must establish a causal relationship between the two constructs.

CONCLUSION

To conclude, the current study aimed at developing a validated tool of post-performance team members' appraisals, namely, the PPTSQ. Objective measures of performance may overlook fundamental psychological mechanisms; hence, an integration of such a subjective inventory may provide valuable information in terms of interpreting the notion *team performance*. The content of the PPTSQ was driven by concepts from groups' goal achievement in team sports (Brawley et al., 1992) and endorsed by two respected sport psychology scholars (i.e., content validity). The construct validity of the modified PPTSQ has been established in this study including two scales (i.e., effort investment and perceived outcome) with a demonstration of high reliability. Furthermore, the analyses confirmed the concurrent validity between PPTSQ factors and TADI and GEQ scores and the predictive validity between PPTSQ factors and objective measures of performance. Future research should address cultural, skill-level, and linguistic-related limitations in this study. As a valid and reliable performance measure that provides a more accurate understanding of the athletes' competitive experience than the objective outcome measure commonly used, the PPTSQ has a great potential to contribute to sport psychology research. Moreover, the value of intervention programs, being tailored based on the PPTSQ's results, must be investigated for the benefit of team performance.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Florida State University IRB Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LG designed the study, collected and analyzed the data, and wrote the manuscript. AB helped with data collection and the write-up of the manuscript. TS analyzed the data and contributed to the write-up of the manuscript. GT supervised the project, contributed to the design of the study, data collection and analysis, and the write-up of the manuscript. All authors actively contributed to the writing process of the manuscript.

FUNDING

LG wishes to express his gratitude to the Research Fund of the Research Authority of the College of Management Academic Studies, Rishon LeZion, Israel, for the financial support provided for this publication. TS was supported by Kinneret Academic College Research Fund.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.615018/full#supplementary-material

REFERENCES

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychol.* 44, 211–233. doi: 10.1016/0001-6918(80)90046-3

Bass, B. M. (1962). *The Orientation Inventory*. Palo Alto, CA: Consulting Psychologists Press.

BBC Sport (2018). *Catch-up: Brazil v Switzerland*. Available online at: https://www.bbc.com/sport/live/football/43974754 (accessed September 6, 2020).

Brawley, L. R., Carron, A. V., and Widmeyer, W. N. (1992). The nature of group goals in sport teams: a phenomenological analysis. *Sport Psychol.* 6, 323–333. doi: 10.1123/tsp.6.4.323

Breivik, G. (2013). Zombie-like or superconscious? a phenomenological and conceptual analysis of consciousness in elite sport. *J. Philos. Sport* 40, 85–106. doi: 10.1080/00948705.2012.725890

Brislin, R. W. (1986). "The wording of translation of research instruments," in *Field Methods in Cross-Cultural Research*, eds W. J. Lonner and J. W. Berry (Beverly Hills, CA: Sage), 137–164

Campbell, W. K., and Sedikides, C. (1999). Self-threat magnifies the self-serving bias: a meta-analytic integration. *Rev. Gen. Psychol.* 3, 23–43. doi: 10.1037/1089-2680.3.1.23

Cannon-Bowers, J.A., Salas, E., & Converse, S.A. (1993). "Shared mental models in expert team decision making," in *Individual and Group Decision Making,* ed N.J. Castellan (Hillsdale, NJ: Lawrence Erlbaum Associates), 221–246.

Carron, A. V. (1984). "Cohesion in sport teams," in *Psychological Foundations of Sport*, eds J. M. Silva & R. S. Weinberg (Champaign, IL, Human Kinetics).

Carron, A. V., Colman, M. M., Wheeler, J., and Stevens, D. (2002). Cohesion and performance in sport: a meta-analysis. *J. Sport Exerc. Psychol.* 24, 168–188. doi: 10.1123/jsep.24.2.168

Carron, A. V., Widmeyer, W. N., and Brawley, L. R. (1985). The development of an instrument to assess cohesion in sport teams: the group environment questionnaire. *J. Sport Psychol.* 7, 244–266. doi: 10.1123/jsp.7.3.244

Chatterji, M., Sentovich, C., Ferron, J., and Rendina-Gobioff, G. (2002). Using an iterative model to conceptualize, pilot-test, and validate a teacher measure of reform readiness. *Educ. Psychol. Meas.* 62, 444–465. doi: 10.1177/001316402128774905

Chow, G. M., and Luzzeri, M. (2019). Post-event reflection: a tool to facilitate self-awareness, self-monitoring, and self-regulation in athletes. *J. Sport Psychol. Action* 10, 106–118. doi: 10.1080/21520704.2018.1555565

Craft, L. L., Magyar, T. M., Bechker, B. J., and Feltz, D. L. (2003). The relationship between the competitive state anxiety inventory-2 and sport performance: a meta-analysis. *J. Sport Exerc. Psychol.* 26, 44–65. doi: 10.1123/jsep.25.1.44

Duda, J. L. (1998). *Advances in Sport and Exercise Psychology Measurement*. Morgantown: Fitness Information Technology.

Eccles, D. W., and Tenenbaum, G. (2007). "A social-cognitive perspective on team functioning in sport," in *Handbook of Sport Psychology* 3rd ed. eds G. Tenenbaum & R. Eklund (New York: Wiley), 264–283. doi: 10.1002/9781118270011.ch12

Eccles, D. W., and Tenenbaum, G. (2004). Why an expert team is more than a team of experts: a social-cognitive conceptualization of team coordination and communication in sport. *J. Sport Exerc. Psychol.* 26, 542–560. doi: 10.1123/jsep.26.4.542

ESPN (2020). *Spanish Primera División Table 2019-20*. Available online at: https://www.espn.com/soccer/standings/_/league/ESP.1/season/2019 (accessed September 6, 2020).

Eys, M., Bruner, M. W., and Martin, L. J. (2019). The dynamic group environment in sport and exercise. *Psychol. Sport Exerc.* 42, 40–47. doi: 10.1016/j.psychsport.2018.11.001

Feltz, D. L., Chow, G. M., & Hepler, T. J. (2008). Path analysis of self-efficacy and diving performance revisited. *J. Sport Exerc. Psychol.* 30, 401–411. doi: 10.1123/jsep.30.3.401

Festinger, L., Schachter, S., and Back, K. (1963). *Social Pressures in Informal Groups*. Stanford, CA: Stanford University Press.

Filho, E. S. M., Dobersek, U., Gershgoren, L., Becker, B., and Tenenbaum, G. (2014). The cohesion-performance relationship in sport: A 10-year retrospective meta analysis. *Sport Sci. Health* 10, 165–177. doi: 10.1007/s11332-014-0188-7

Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive–developmental inquiry. *Am. Psychol.* 34, 906–911. doi: 10.1037/0003-066x.34.10.906

Gershgoren, L., Filho, E. M., Tenenbaum, G., and Schinke, R. J. (2013). Coaching shared mental models in soccer: a longitudinal case study. *J. Clin. Sport Psychol.* 7, 293–312. doi: 10.1123/jcsp.7.4.293

Gershgoren, L., Tenenbaum, G., Gershgoren, A., and Eklund, R. C. (2011). The effect of parental feedback on young athletes' perceived motivational climate, goal involvement, goal orientation, and performance. *Psychol. Sport Exerc.* 12, 481–489. doi: 10.1016/j.psychsport.2011.05.003

Goff, A. M. (2011). Stressors, academic performance, and learned resourcefulness in baccalaureate nursing students. *IntL. J. Nurs. Educ. Scholarsh.* 8:1.

Gould, D., Guinan, D., Greenleaf, C., Medbery, R., and Peterson, K. (1999). Factors affecting Olympic performance: perceptions of athletes and coaches from more and less successful teams. *Sport Psychol.* 13, 371–394. doi: 10.1123/tsp.13.4.371

Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educ. Psychol. Meas.* 66, 930–944. doi: 10.1177/0013164406288165

Grand, R. R., & Carron, A. V. (1982). "Development of a team climate questionnaire," in *Psychology of Sport and Motor Behavior: Research and Practice*, eds L. M. Wankel & R. B. Wilberg (Alberta: University of Alberta).

Green, S. B., Lissitz, R. W., and Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educ. Psychol. Meas.* 37, 827–838 doi: 10.1177/001316447703700403

Hancock, G. R., and An, J. (2020). A closed-form alternative for estimating ω reliability under unidimensionality. *Meas. Interdiscip. Res. Perspect.* 18, 1–14. doi: 10.1080/15366367.2019.1656049

Hayes, A. F., and Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. but. *Commun. Methods Meas.* 14, 1–24. doi: 10.1080/19312458.2020.1718629

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *J. Manag.* 21, 967–988. doi: 10.1177/014920639502100509

Hinsz, V. B. (2004). "Metacognition and mental models in groups: an illustration with metamemory of group recognition memory," in *Team cogntion: Understanding the factors that drive process and performance* eds E. Salas & S. M. Fiore (Washington, DC: American Psychological Association), 33–58. doi: 10.1037/10690-003

Johnson, T. E., Lee, Y., Lee, M., O'Connor, D. L., Khalil, M. K., and Huang, X. (2007). Measuring sharedness of team-related knowledge: design and validation of a shared mental model instrument. *Hum. Resour. Dev. Intl.* 10, 437–454. doi: 10.1080/13678860701723802

Kahneman, D., and Tversky, A. (1973). On the psychology of prediction. *Psychol. Rev.* 80, 237–251.

Kennett, D. J. (1994). Academic self-management counselling: preliminary evidence for the importance of learned resourcefulness on program success. *Stud. High. Educ.* 19, 295–307. doi: 10.1080/03075079412331381890

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling* 4th ed. New York: The Guilford Press

Lance, C. E., Butts, M. M., and Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: what did they really say? *Organ. Res. Methods* 9, 202–220. doi: 10.1177/1094428105284919

Lazarus, R. S. (2000). How emotions influence performance in competitive sport. *Sport Psychol.* 14, 229–252. doi: 10.1123/tsp.14.3.229

Li, F., and Harmer, P. (1996). Confirmatory factor analysis of the group environment questionnaire with an intercollegiate sample. *J. Sport Exerc. Psychol.* 18, 49–63. doi: 10.1123/jsep.18.1.49

Martens, R., Landers, D. M., and Loy, J. W. (1971). *Sport Cohesiveness Questionnaire*. Champaign, IL: University of Illinois.

McDonald, R. P. (1999). *Test Theory: A unified Approach*. Mahwah, NJ: Erlbaum.

Medvec, V. H., Madey, S. F., and Gilovich, T. (1995). When less is more: counterfactual thinking and satisfaction among Olympic medalists. *J. Pers. Soc. Psychol.* 69, 603–610. doi: 10.1037/0022-3514.69.4.603

Micklewright, D., Papadopoulou, E., Parry, D., Hew-Butler, T., Tam, N., and Noakes, T. (2009). Perceived exertion influences pacing among ultramarathon runners but post-race mood change is associated with performance expectancy. *South Afr. J. Sports Med.* 21, 167–172.

Moritz, S. E., Feltz, D. L., Fahrbach, K. R., and Mack, D. E. (2000). The relation of self-efficacy measures to sport performance: a meta-analytic review. *Res. Q. Exerc. Sport* 71, 280–294. doi: 10.1080/02701367.2000.10608908

Nelson, T., and Narens, L. (1990). Metamemory: a theoretical framework and new findings. *Psychol. Learn. Motiv.* 26, 125–173. doi: 10.1016/s0079-7421(08)60053-5

Rasker, P. C., Post, W. M., and Schraagen, J. M. C. (2000). Effects of two types of intra-team feedback on developing a shared mental model in command and control teams. *Ergonomics* 43, 1167–1189. doi: 10.1080/00140130050084932

Reise, S. P., Bonifay, W. E., and Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *J. Pers. Assess.* 95, 129–140. doi: 10.1080/00223891.2012.725437

Tenenbaum, G. (2003). "Expert athletes: an integrated approach to decision making," in *Expert Performance in Sports* eds J.L. Starkes & K.A. Ericsson (Champaign, IL: Human Kinetics), 191–218

Tenenbaum, G. E., Eklund, R. C., and Kamata, A. E. (2012). *Measurement in Sport and Exercise Psychology*. Champaign, IL: Human Kinetics.

Tenenbaum, G., and Gershgoren, L. (2014). "Decision making," in *Encyclopedia of Sport & Exercise Psychology*, Vol. 4, eds R. Eklund and G. Tenenbaum (Thousand Oaks, CA: SAGE Publications, Inc.), 190–192.

Transfermarket (2020a). *FC Barcelona Transfer Record 2019–20*. Available online at: https://www.transfermarkt.com/fc-barcelona/transfers/verein/131/saison_id/2019 (accessed September 6, 2020).

Transfermarket (2020b). *Granada Cf Historical Rankings*. Available online at: https://www.transfermarkt.com/fc-granada/platzierungen/verein/16795 (accessed September 6, 2020).

Trizano-Hermosilla, I., and Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: congeneric and asymmetrical measurements. *Front. Psychol.* 7:769.

Vaughn, B. K., & Daniel, S. R. (2012). "Conceptualizing validity," in *Measurement in Sport and Exercise Psychology* eds G. Tenenbaum, R. C. Eklund, & A. Kamata (Champaign, IL: Human Kinetics), 33–39

Webber, S. S., Chen, G., Payne, S. C., Marsh, S. M., and Zaccaro, S. J. (2000). Enhancing team mental model measurement with performance appraisal practices. *Organ. Res. Methods* 3, 307–322. doi: 10.1177/109442810034001

Weiner, B. (Ed.). (1974). *Achievement Motivation and Attribution Theory*. New York: General Learning Press.?

Widmeyer, W. N., Brawley, L. R., & Carron, A. V. (1985). *The Measurement Of Cohesion in Sport Teams: the Group Environment Questionnaire*. London, CA: Sports Dynamics.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 30 October 2020
doi: 10.3389/fpsyg.2020.564855

Check for updates

Implementation of the Video Assistant Referee (VAR) as a Career Change-Event: The Israeli Premier League Case Study

Roy David Samuel[1†], Yair Galily[1†], Edson Filho[2†] and Gershon Tenenbaum[1†]*

[1] Interdisciplinary Center Herzliya, Herzliya, Israel, [2] Wheelock College of Education and Human Development Boston University, Boston, MA, United States

The inclusion of the video assistant referee (VAR) in the Laws of the Game [International Football Association Board (IFAB)] reflects a historic action in the world of soccer. The VAR was designed to reduce critical errors in soccer referees' decision-making (DM), thereby increasing the social perceptions of justice. From the referees' perspective, the implementation of the VAR represents a technical–technological career change-event. This case study adopted an intrinsic mixed-methods methodology to investigate the implementation of the VAR system within the Israeli Premier League context. The results indicated that the initial VAR performance rates of the Israeli referees were not high compared with data from other countries (e.g., Italy). The Israeli referees perceived the VAR implementation as a moderate change-event in their careers. The largest effects were in pre-match preparation, players' management, public perception, and DM. The referees felt that their perceived pressure during the matches decreased. This change-event produced off-field demands mostly related to the educational process and on-field demands mainly related to developing VAR proficiency. The referees also experienced off-field (e.g., lack of clear goals and rewards system) and on-field barriers (e.g., errors of the VARs). To effectively cope with these new demands and barriers, most of them initially consulted with others and then made a decision to apply all necessary adjustments in response to the new situation. They also received considerable support yet perceived their cooperation with the professional committee as moderate. The referees showed some professional improvements; still they experienced scrutiny from the clubs, the media, and the Referee Union. Thus, three of them perceived the outcome of this change process favorably, four naturally, and four negatively. The discussion presents reflections of these findings in light of the recently emerging literature on technological officiating aids in sport. Recommendations are provided for referee unions who aspire to integrate the VAR system into their operation.

Keywords: football, officials, decision-making, video replay, transition, technology

Edited by:
Bruno Travassos,
University of Beira Interior, Portugal

Reviewed by:
João Aragão e Pina,
University Institute of Lisbon, Portugal
Diogo Monteiro,
Polytechnic of Leiria, Portugal

***Correspondence:**
Yair Galily
ygalily@idc.ac.il

†ORCID:
Roy David Samuel
orcid.org/0000-0001-9914-3754
Yair Galily
orcid.org/0000-0001-5308-5675
Edson Filho
orcid.org/0000-0002-8548-4651
Gershon Tenenbaum
orcid.org/0000-0001-6357-873X

Specialty section:
This article was submitted to Movement Science and Sport Psychology, a section of the journal Frontiers in Psychology

Received: *22 May 2020*
Accepted: *07 October 2020*
Published: *30 October 2020*

Citation:
Samuel RD, Galily Y, Filho E and Tenenbau G (2020) Implementation of the Video Assistant Referee (VAR) as a Career Change-Event: The Israeli Premier League Case Study. Front. Psychol. 11:564855. doi: 10.3389/fpsyg.2020.564855

INTRODUCTION

I was left with the decision I had taken with no independent evidence that I'd got it wrong other than a gut feeling, and I was just hoping that Ronaldo would miss the penalty. But he didn't (Bradbury, 2020).

This quote of former English Premier League referee, Howard Webb, reflects one of the key aspects of soccer (association football) refereeing, namely, decision-making (DM). In this context, the inclusion of the video assistant referee (VAR) in the Laws of the Game (International Football Association Board [IFAB], 2018b) and its introduction in the 2018 World Cup represent a historic action in the world of soccer (Simón, 2020). This system is aimed to reduce critical DM errors in soccer refereeing, focusing on four areas: approval/disapproval of goals, penalty decisions, direct red card decisions, and mistaken identity in awarding a red or yellow card (International Football Association Board [IFAB], 2017). Within these four areas, the VAR system reviews the various match events, conducting a silent check. In case that the video referee identifies a potential clear error, s/he can then communicate with the on-field head referee, who in turn can (a) change the call on the advice of the VAR, or (b) conduct an *on-field review* (OFR), using a designated monitor station. There are specific protocols for the usage of the system, including when intervention is allowed, how to signal that an intervention is in order, and various player management rules (International Football Association Board [IFAB], 2017). In terms of the refereeing DM sequence, VAR intervention appears at the end of the process, when the on-field referee has already made a decision and receives external feedback from the refereeing team. S/he is then faced by a subsequent decision, namely, to keep or change the original decision (Samuel et al., 2020b).

Arias et al. (2011) reviewed the literature on rule modification in sport and identified that most rule modifications were intended to improve sports performance, attract spectators, and attend to commercial pressures and interests. The VAR system rule modification was mainly due to pressure from club owners and the media to improve refereeing performance and, consequently, justice in the game (Simón, 2020).

The VAR is a technological system that includes a human factor (i.e., the video referee and assistant referee, the operator), a technological factor (the video system, the audio system), a human–technological interface (Sánchez Cid and García García, 2020), and a social–ethical factor (e.g., referee team dynamics, perceptions of fairness and justice; de Dios Crespo, 2020). As such, its introduction presented major modifications in the way referees train, prepare for, and officiate matches (Armenteros and Webb, 2020; Fernández Ruiz et al., 2020) as well as in refereeing psychology (de la Vega and Fuentes, 2020). Furthermore, Kolbinger and Lames (2017) reviewed the literature regarding the use of technological officiating aids in various game sports. They identified seven major issues: the underlying phenomena, usage patterns, accuracy, standard of review, influence on the nature of the game, material as well as immaterial costs, and the amount of authority that is granted to the officiating aid. It could be argued that VAR is a technological aid with important ramifications for all these dimensions.

Among semi-elite Australian rugby league referees, for example, the inclusion of the video referee technology had led to various stressors pertaining to training, DM, media, coaches and players, and communication issues (Baldwin, 2014). These changes resulted in significant challenges in terms of referees' coping and motivation. Yet, currently, there is a scarcity of scientific publications concerning the usefulness of the VAR system, as well as its potential influence on referees' motivation, performance, and development (Samuel et al., 2020b). The present case study, therefore, examined the implementation of the VAR system as a technical–technological change-event within Israeli referees' careers (Samuel et al., 2017).

The Emergence of the Video Assistant Referee System

The VAR system was initially trialed by various national associations, like the Italian Serie A (Simón, 2020) and the German Bundesliga (Kolbinger, 2020). An analysis that was based on these initial trials (i.e., 804 matches) concluded that, in most of the matches (68.8%), there was no review and, on average, in every three matches, there was one clear and obvious error. The system improved the total accuracy in key match decisions (i.e., in the four categories that VAR intervenes in) from 93 to 98.9%. Moreover, in 8% of all matches, the VAR had a decisive impact on the outcome of the match (International Football Association Board [IFAB], 2018b). As a result of these positive analysis outcomes, on March 3, 2018, the IFAB wrote the VARs into the Laws of the Game on a permanent basis, presenting its decision as a "historic step for greater fairness in football" (International Football Association Board [IFAB], 2018a).

An analysis of 1,024 matches played in the Italian Serie A and the German Bundesliga league during the seasons before and after the implementation of the VAR system revealed a decrease in the number of offsides, fouls, and yellow cards after the implementation of the VAR (Lago-Peñas et al., 2019). There was an increase in the number of minutes added to the playing time in the first half and the full game, but not in the second half. Finally, there were no significant changes in the number of penalties or red cards per match (Lago-Peñas et al., 2019). It was further proposed that using technological officiating aids in sport might undermine the authority of the referee, interfere with the flow of the match, and even create injustice when on-field decisions are overruled by the video review (Kolbinger and Lames, 2017). Moreover, better visual coverage of the field and match events can result in detecting more minor incidents, previously overlooked by referees, players, and fans. This can be considered as one of the potential threats of including such a technology aid (Kolbinger and Lames, 2017; de Dios Crespo, 2020).

Another issue that might raise concern is the effectiveness of communication between the on-field referee and the VAR through the headset (Baldwin, 2014; Sánchez Cid and García García, 2020). The referee is expected to trust the judgment of the VAR in critical match incidents and his/her professional skills in deciding on the correct decision. This might entail a significant change in the authority of the head referee to officiate the match autonomously. Considering the role of the VAR, this task is new, requiring the referees to engage in video DM, in most cases out of context and under much pressure and public scrutiny.

Video Assistant Referee as a Career Change-Event

The scheme of change for sport psychology practice (SCSPP; Samuel and Tenenbaum, 2011a) is a conceptual framework describing typical characteristics of change processes in sport performers' careers. The change process begins with the emergence of a change-event that disrupts the career *status quo* and creates emotional and cognitive imbalance. Each change-event is characterized by unique demands (i.e., perceived challenges associated with the various emerged changes). Performers then evaluate the meaning and significance of the event in the context of their careers, considering existing resources of coping and potential solutions. Change-events are characterized by unique *emotional profiles* pertaining to perceived significance, perceived severity, perception of others, emotional and cognitive reactions, and perceived control over the situation. This appraisal process leads to active decisions; initially, *a strategic decision* concerning preliminary response to the new situation (i.e., deny/ignore it, cope independently, consult with others, consult with a sport psychologist) and a subsequent decision to avoid change or to *a decision to change* (i.e., apply all required modifications to effectively cope with the new situation). It is assumed that if referees decide to change, and also have a possibility to implement the change, they will feel efficient in their coping efforts, in control over the situation, and accept responsibility for initiating the change. They would then better cope with the change process barriers (i.e., factors that debilitate effective adaptation). As a result, the change process outcome will be perceived more favorably (Samuel and Tenenbaum, 2011a; Samuel et al., 2017; Samuel, 2019).

Previous research indicated a highly dynamic soccer refereeing career, comprising of various transitions and change-events (Slack et al., 2013; Samuel et al., 2017; Samuel, 2019). Using the SCSPP as a conceptual framework, Samuel et al. (2017) examined the career experiences of 154 Israeli referees from several professional levels. The referees experienced over 10 types of change-events in their careers, including *a transition to a higher league*, *excelling in a big match*, and *a very poor performance or a decision error in a big match*. Half of the sample reported experiencing *the initiation of the communication system*, which could be classified as a technical–technological modification. This study further indicated that most referees made an initial strategic decision to consult with others in response to the initiation of a change-event and a subsequent decision to change.

The emergence of the VAR system in soccer refereeing can be conceptualized as a quasi-normative transition (i.e., applies to a selected group of elite referees; Stambulova and Samuel, 2020). The VAR system challenges referees to modify their training and performance. This might influence their motivation for refereeing directly and indirectly through effects on performance and professional and public credibility and reputation. Anecdotal evidence suggests a mixture of opinions regarding the usefulness and necessity of the VAR system among referee union leaders (Simón, 2020). In this context, VAR implementation was challenging in most relevant countries, emphasizing the difficulty in integrating such a major change.

Study Purpose and Objectives

How do soccer referees perceive and respond to the implementation of the VAR system? How does this system influence their training, performance, and career status? How do they cope with these changes? To date, no empirical study examined how referees perceive the integration of the VAR system as a change-event in the context of their careers. As this system might induce various types of changes in referees' training, performance, and career status, it is important to study its potential effects on referees' careers. Therefore, the present study used the SCSPP (Samuel and Tenenbaum, 2011a) framework to examine elite referees' perceptions and responses to this change-event, as well as their coping efforts (including conscious coping decisions) and support resources. As each country has its own unique referee union structure and cultural environment and a specific process in which VAR was initiated, it is important to carefully consider context. Thus, we applied a case study methodology to examine the implementation of the VAR system within the context of the Israeli Referee Union and the Israeli Premier League—"Ligat Ha'Al." We begin by presenting the official data concerning the general VAR-related performance of the Israeli referees. Subsequently, our focus is on understanding the referees': (a) perceptions of this change-event, (b) evaluations of the major areas of change (i.e., demands and barriers), (c) conscious decisions and coping, (d) support resources, and (d) the change process outcomes (including effects on motivation).

MATERIALS AND METHODS

Design

An intrinsic mixed-methods case study methodology was applied (Yin, 2014; Hodge and Sharp, 2016). Specifically, a pragmatic approach was adopted with hopes to holistically describe the referees' change process in their unique context (Patton, 2002; Sparkes, 2015). Qualitative data (i.e., interview notes, referees' feedback) were collected over a 19-month period, lasting from the beginning of the VAR implementation process (September 2018) to the end of the regular season (March 2020). In addition, quantitative data [i.e., the Change-Event Inventory (CEI); Samuel and Tenenbaum, 2011b] were collected in a single administration point, during January–February 2020 (i.e., middle of the 2019–2020 season). This time point was selected to ensure the referees already had sufficient experience with the VAR system but were still experiencing the change process (i.e., they were still attempting to adjust to the VAR). A similar approach was previously used to examine referees' change process to the Premier League (Samuel, 2019). The referees' experiences regarding the implementation of the VAR system were analyzed within the Israeli Premier League context as they were advancing throughout their change processes. Data sources complemented each other.

Case Study Context—Video Assistant Referee in the Israeli Premier League

Following the IFAB's decision to implement the VAR in various federations beginning of March 2018 and the successful 2018 World Cup, the major European countries integrated VAR in their regular leagues. Israel is affiliated to the Union of European Football Associations (UEFA). The Israel Football Association, together with the Israeli Professional Football Leagues Directory, has therefore, began considering implementing a VAR project in Israel. The Directory received a green light from the teams and a financial assurance for 5 years' term. In fact, Israel was somewhat a pioneer in this context, as one of the small–medium size countries to implement such a complex project. In September 2018, a public announcement was made that the VAR would already be integrated in the same season's playoff matches beginning mid-March 2019. This created a very tight schedule for such a complex implementation project—6 months compared with the IFAB's recommended 12-month period—resulting in high pressure and a very focused process. Two project managers were appointed in charge of setting the technological infrastructure as well as the professional instruction and certification. As the VAR project was novel, the professional manager had to individually strategize the implementation process, including the comprising of an instruction program to align with the IFAB's strict certification demands. On top of certifying the referees, VARs, assistant VARs, and operators, there was a need to technologically certify all hosting stadiums to receive official IFAB certifications. Furthermore, a training area was required to be modified to include two simulator rooms and an instruction classroom.

The theoretical and practical instruction of the referees during the active league season was challenging, as they were preoccupied with training and match officiating. There was a strict learning protocol that each referee went through, including (a) theoretical knowledge of the VAR system and protocol, (b) an offline simulator of clips and then full matches, and (c) live simulations. As a final step, the referees experienced offline performance in real Premier League matches. In total, the referees engaged in 45 h of instruction. Then, they initiated a pilot phase in the Premier League upper playoff of the 2018–2019 season (i.e., 30 matches in total). As part of the continued instruction, the project manager provided feedback to the VARs following each performance, including analysis of technical performance and teamwork between the VAR and the operator.

Participants

The participants were 11 male elite referees from the Israeli Premier League (M_{age} = 34.54 years, SD = 6.23), representing 65% of the designated population. Their refereeing experience ranged from 10 to 25 years (M = 18.14 years, SD = 5.41). Three of the referees were relatively new to the Premier League (i.e., 1–3 seasons), two had moderate experience (i.e., four seasons), and six were seasoned referees.

Procedure

We followed the American Psychological Association's ethical guidelines concerning the formulation of a case study with existing clients (APA, 2017). Specifically, genuine voluntariness was maintained as the participants could choose to not participate in this case study or withdraw from it at any point. Also, they could decide whether to exclude any of their related information. In addition, the participants provided their written informed consent. Finally, the first author engaged in an open dialogue with the participants concerning the formulation and presentation of the case study data and debriefed them concerning the results and conclusions (APA, 2017). As the data were obtained as part of an ongoing consultation process, no institutional ethical approval was required.

Initially, the participants were informed of the study's overarching purpose, i.e., a case study concerning professional referees' experiences with the newly introduced VAR system. They provided their preliminary verbal consent on a voluntary basis. Qualitative data sources (e.g., interview notes, open-ended inventory data) were collected as part of an ongoing consultation process involving the first author and the referees. Additionally, the VAR project manager was contacted and asked to provide his account of the VAR implementation process and the official data on VAR-related performance. Also, the first author visited the VAR instruction center and the Video Operations Room (VOR) during a live match. Moreover, the referees were provided an explanation as to how to complete the CEI; they were asked to reflect on their individual VAR implementation experience and to complete the CEI at home. Upon manuscript completion, all participants received a copy of it and provided their informed consent for their anonymized data to be presented. They were asked to read the manuscript, providing their insights and suggestions for modifications. Data collection and analysis ensured, to the possible degree, the anonymity and confidentiality of the participants' information.

Data Collection and Treatment

Qualitative Data Collection

There were three main qualitative data sources. First, the participants received sport psychology support as an ongoing service provided by the Israel Referee Union. In these semimonthly 1-h sessions, the referees consulted with the first author on an array of issues pertaining to their refereeing engagement, including psychological preparation for matches and match analysis (see Samuel, 2015), organizational stress, and career management (see Samuel et al., 2017; Samuel, 2019). Since the introduction of the VAR system, much attention was dedicated to assessing the referees' perceptions of the system, how it impacted them in terms of performance and organizational demands, and how to effectively cope with this change-event. During these sessions, interview notes were taken pertaining to main ideas and thoughts, direct quotes of the referees' words, as well as the consultant's own interpretations. The first author then read and reread these notes (i.e., familiarization), identifying all references made to the VAR system and collecting them into a single file (see Patton, 2002). This practice was used in previous change-event studies with mixed-methods designs (e.g., Samuel, 2019).

Second, as part of completing a measurement related to their transition experiences, the referees were asked to freely elaborate in writing on (1) the new demands (i.e., what an athlete wants/ought to achieve in going through the transition, Stambulova and Samuel, 2020) associated with officiating with the VAR system in professional, psychological, occupational, time, and training aspects; (2) the barriers and difficulties experienced throughout the VAR implementation process; (3) their coping efforts when addressing this change-event; and (4) any additional comments and recommendations they have. Finally, for the purpose of this study only, the VAR project manager of the Israel Referee Union was asked to provide a detailed internal report on the implementation process, including the decision to implement, the professional demands, the instructional process, main barriers, and the referees' performances. In addition, the VAR project manager provided the official data concerning all matches officiated with the VAR system in Israel (i.e., for all Premier League referees) throughout the 2018–2019 Pilot Phase and two thirds of the 2019–2020 Season.

Quantitative Data Measurement

To assess the referees' change process, they completed a *modified Change-Event Inventory* (Samuel and Tenenbaum, 2011b). The original CEI measures change-event experiences in a retrospective manner using a three-section format: (a) demographic information; (b) perception of and reaction to a change-event (i.e., the implementation of the VAR system), measured by 13 two-item Likert-type subscales (e.g., perceived significance of the event, perceived emotional severity of the event, perceived control over the event), ranging from 1 to 5 (1, *not at all/very negative*; 3, *moderate/neutral*; 5, *very much/very positive*); and (c) coping-related DM and availability of support resources. Previous research on a heterogeneous sample of competitive athletes indicated adequate psychometric properties of the CEI, including temporal stability, internal consistency (i.e., all Cronbach's α ranged between 0.68 and 0.89), and internal factorial structure (Samuel and Tenenbaum, 2011b).

For the purpose of this study, as there is no existing VAR-related measure, the CEI was modified to capture the referees' VAR implementation experience, and additional items were added to the inventory. First, a VAR-specific section was added, which included items assessing the participants' enjoyment of refereeing with the VAR system and how they evaluate their performance since it was introduced. Also, the degree in which they evaluate the VAR's impact in 15 areas related to performance (e.g., running patterns, DM, and game management), referees' emotional response (e.g., changes in self-confidence and in stress levels), and in organizational and public perceptions. Second, two items pertaining to satisfaction from the cooperation with the professional committee concerning the VAR were included in the second section of the inventory. Most two-item subscales indicated adequate internal consistency indices (i.e., α = 0.67–0.91), despite the small sample size. Low reliability indices (i.e., $\alpha \geq 0.57$) were found, however, in *past experience in similar events*, *satisfaction of coping*, *capacity for change*, *cognitive response*, and *the perception of others*.

Data Analysis

In line with Sparkes (2015) recommendations concerning quantitative and qualitative data sources in mixed-methods case studies, all data sources were integrated into a coherent account of the soccer referees' change process. Data were collected as part of psychological support services, so confidentiality considerations were pertinent (APA, 2017). Data are, therefore, presented with no identifying cues. The qualitative data were read several times. An indwelling posture was adopted by going through the data, immersing in it, and understanding the referees' point of view from an empathetic position (Holt and Sparkes, 2001). The lengthy consultation notes as well as the referees' open-ended inventory responses were summarized by identifying key events, thoughts, and issues pertaining to each referee's experiences within his change process. The analysis then focused on the condensed summary report created for each referee, as well as the referee's quantitative data. Yin (2014) suggested that the preferred strategy in a case study analysis is to follow a theoretical framework that accounts for "how" and "why" questions. Therefore, the SCSPP framework (Samuel and Tenenbaum, 2011a) was used to understand the referees' change process. Each of the 11 referees was considered a *unit of analysis*, and attention was given to the individual unique experience as well as to the shared experiences. Specifically, the analysis focused on realizing the on- and off-field demands and barriers associated with the VAR implementation, as well as the referees' perceptions of the change process in the context of their careers, their emotional and cognitive reactions, how they had coped with it (i.e., their DM, their coping efforts, and the existing available support), and the outcome of the change process in the context of their careers.

In addition, the quantitative data were analyzed using descriptive and inferential statistical procedures, including reliability indices, central tendency metrics, and correlations among variables. With a relatively small sample of 11 participants, we were limited in statistical power (Schweizer and Furley, 2016), and any generalization of the quantitative data should be considered with a degree of caution.

Rigor and Quality of the Case Study

Evaluating the rigor of this study, we have adopted a relativist stance and created our list of criteria (Sparkes, 2015; Smith and McGannon, 2018). First, we believe this is *a worthy topic* (Tracy, 2014), as not much research effort was devoted to examining referees' experiences with the VAR system. Moreover, the first author was well-positioned within the case study context, serving for many years as a sport psychologist of the Israel Referee Union. This experience facilitated the formation of clear research questions as well as the *credibility* of data collection and analysis (Patton, 1999; Tracy, 2014). The lead researcher also engaged in self-reflection, maintaining a critical mindset throughout the research process, including data collection and analysis. Still, we were aware of potential bias, and for this reason, the other team members were involved in the data analysis as "critical friends." Specifically, the research team engaged in a process of reflexive dialogue concerning the interpretations of the data (Sparkes and Smith, 2014). *Triangulation* was conducted by applying various types of data sources (i.e., qualitative and quantitative; Patton, 1999). Using the SCSPP (Samuel and Tenenbaum, 2011a) as a conceptual framework to guide the data analysis further increased credibility. In the "Results" section, we present a *coherent* and *comprehensive* narrative of the referees' experiences, in a unique context, by providing themes and associated direct quotes (i.e., a thick description; Patton, 2002). Finally, concerning the qualitative data, representational generalizability (Smith, 2018) was reached by recognizing similarities and differences to the results of previous athletic transition research. Using the SCSPP framework to analyze the referees' experiences, analytical generalizability was produced (Smith, 2018). In line with Smith and McGannon's (2018) recommendations, upon completion of the data analysis process, the manuscript was sent to the participants for *member reflections*. We asked them for their feelings and perspectives regarding the presented data and interpretation. The referees and the VAR project manager conveyed that the presented data were rich and accounted for their transitional experience. Noteworthy, this process of member reflections is also congruent with ethically sound research, as it allows participants to protect their well-being by identifying any misrepresentations of their data.

RESULTS

In this section, we initially provide data concerning the performance of the Israeli Premier League referee squad with the VAR system, as this information is important to contextualize this case study. Subsequently, we present the experiences (change process) of the study participants in detail, including perceptions, demands and barriers, decisions and coping, support resources, and change process outcomes.

The Israeli Video Assistant Referee Experience

As part of the data collection, the Israeli VAR project manager provided the official refereeing performance data of all 212 matches officiated with the VAR system in Israel until March 1, 2020 (**Table 1**). This period included the VAR system Pilot Phase in the 2018–2019 season playoff and two thirds of the 2019–2020 season (i.e., which was abrupted due to the Coronavirus pandemic). During this period, all 17 Premier League referees had officiated with the VAR system 6–17 matches each ($M = 12.40$, $SD = 3.92$). As shown in **Table 1**, across all 212 matches, there were 89 critical match errors made by the on-field referees. More specifically, during only two thirds of the 2019–2020 season (i.e., 182 matches in total), there were 81 critical errors. This relatively high error rate reflected a significant ($p < 0.05$) increase in the number of critical errors made by the on-field referees compared with the 2018–2019 season (69 critical errors), the 2017–2018 season (49 critical errors), and the 2016–2017 season (52 errors).

Sixty-nine of the 89 critical errors performed by the on-field referees (77.5%) were correctly rectified due to VAR intervention. In total, the referees conducted 80 on-field reviews (OFRs), which translate into one OFR every 2.65 matches on average. They modified their original decisions in 55 of the cases (68.75%). In nine of the cases, the on-field referee did not modify his original decision and was incorrect, whereas in three of the cases, the referee wrongly modified his original decision due to a VAR error. Furthermore, across all matches, the VARs intervened 98 times. As can be seen in **Table 1**, in 78 cases (79.60%), this intervention was accurate, whereas in 20 cases (20.40%), the intervention was incorrect. Also, in 11 cases, the VAR did not intervene when it should have.

Referees' Perceptions of the Video Assistant Referee

The referees' perceptions of the VAR system are presented in **Table 2** and in **Figure 1**. As shown in **Table 2**, the referees had initially wished the system to be implemented in the Israeli league. Yet, their current evaluation of enjoyment using the system and their appreciation of the system was somewhat moderate. These quantitative data correspond to the qualitative data concerning the progress of the VAR implementation in the Israeli context. In the beginning of the process, the system was perceived as new and innovative, and most attention was given to its initial learning and implementation through the IFAB instruction protocol. Thus, most of the referees were appreciative about implementing the system. There was also strong pressure from the clubs and fans to implement it, as they perceived the system as a professional tool to facilitate accuracy and, thereby, justice in refereeing. However, as the system got integrated in the 2018–2019 season playoff matches, the referees were still inexperienced, and therefore, criticism of the referees' performance had begun to emerge. This had significantly amplified in the 2019–2020 season. The referees felt that they had invested much effort in learning to officiate with VAR, yet still their performance was scrutinized. This situation had led to some negative change in the way the referees perceived the system, as one of them explained:

> It was the same with the additional assistant referees (AAR)—when it started, everyone thought it's a good thing that we have an extra pair of eyes. But then they began to acknowledge the problems (bad angle, fear of making decisions) and they started to drop this. Referees were afraid to be AARs. It's the same process with the VAR.

The data presented in **Figure 1** show that the referees perceived this change-event in the context of their careers with a moderate emotional profile. Specifically, they evaluated the significance of this change-event in their careers as moderate, its severity as neutral, and responded with moderate emotional and cognitive responses. Moreover, their perceived control over this new situation was also moderate. Expectedly, as the referees felt more control over the new situation, their cognitive reaction was less negative (i.e., less worried; $r = -0.86$, $p < 0.01$). Similarly,

TABLE 1 | The Israeli referees' performance with the VAR system throughout the 2018–2019 pilot phase and two thirds of the 2019–2020 season.

Factor	30 pilot matches	Matches 1–45	Matches 46–90	Matches 91–135	Matches 136–182
Overall critical errors	8	21	22	19	19
Rectified on-field critical errors	7	13	15	16	18
Correct VAR interventions	7	17	19	17	18
Incorrect VAR interventions	4	6	5	3	2
Missing VAR interventions	1	4	3	2	1
Total OFRs	10	21	20	14	15
Correct OFRs	6	15	15	11	13
Incorrect OFRs	4	6	5	3	2
Match time suspension during VAR check	199 s	198 s	150 s	164 s	155 s
Total match time suspension	116 s	108 s	92 s	82 s	79 s

OFR, on-field review; VAR, video assistant referee.

TABLE 2 | Descriptive statistics of VAR perceptions and associated refereeing modifications.

Perceptions	*N*	Min	Max	*M* (*SD*)
Wishing the VAR system to be implemented in Israel	11	3.00	5.00	4.37 (0.81)
Enjoying officiating with VAR as on-field referees	11	3.00	5.00	3.91 (0.83)
Enjoying officiating as VARs	11	2.00	5.00	3.78 (1.10)
Appreciating the VAR system as facilitating officiating	11	2.00	5.00	3.82 (0.87)
The influence of VAR on:				
Running pattern and locations	11	1.00	2.00	1.18 (0.40)
Foul evaluation	11	1.00	4.00	2.18 (0.87)
Decision making	11	1.00	4.00	2.54 (0.82)
Referee team communication	11	1.00	4.00	2.45 (1.21)
Players management and communication	11	1.00	5.00	2.64 (1.29)
Focus throughout the match	11	1.00	4.00	2.18 (1.08)
Technical aspects of refereeing	11	1.00	4.00	2.00 (1.00)
Pre-match preparation	11	1.00	5.00	2.82 (1.17)
Refereeing philosophy	11	1.00	4.00	2.18 (0.98)
Self-confidence	11	1.00	4.00	1.91 (1.14)
Motivation	11	1.00	3.00	1.73 (0.79)
Pressure increase	11	1.00	3.00	1.55 (0.69)
Pressure decrease	11	1.00	5.00	2.45 (1.29)
Changes in own public perception	11	2.00	3.00	2.55 (0.52)
Changes in own Referee Union status	11	1.00	3.00	2.28 (0.79)

VAR, video assistant referee.

as past experience in coping with such change-events was higher, they also felt less worried ($r = -0.75$, $p < 0.01$). Finally, the higher was their age, the more positive was their emotional reaction ($r = 0.60$, $p < 0.05$).

Demands and Barriers

Table 2 data revealed that the largest effects of the VAR system integration on refereeing were in pre-match preparation, players' management, public perception, and DM—in this order. Interestingly, the referees felt that the pressure they experienced during the matches decreased rather than increased. Thus, it is evident that the integration of the VAR system did not create extreme professional demands for the Israeli referees. As one of them commented: "the change is mostly mental to cope with an error during the match. From a professional standpoint, the change is minimal and requires small and insignificant modifications." Nevertheless, looking at the qualitative data, it seems that there were meaningful differences among the referees in the tone in which they related to this change process. While some referees perceived this change-event in a neutral manner, some perceived it favorably and others unfavorably. The following section, therefore, depicts the demands and barriers associated with this change-event, classified into off-field and on-field.

Off-Field Demands

The initial instruction of the VAR system (September 2018–March 2019) in terms of following the IFAB protocol was highly demanding in both time and energy. The referees were required to travel weekly to the instruction facility located in the center of Israel. They received some monetary reward for these educational sessions, yet not necessarily perceived as satisfactory for the time

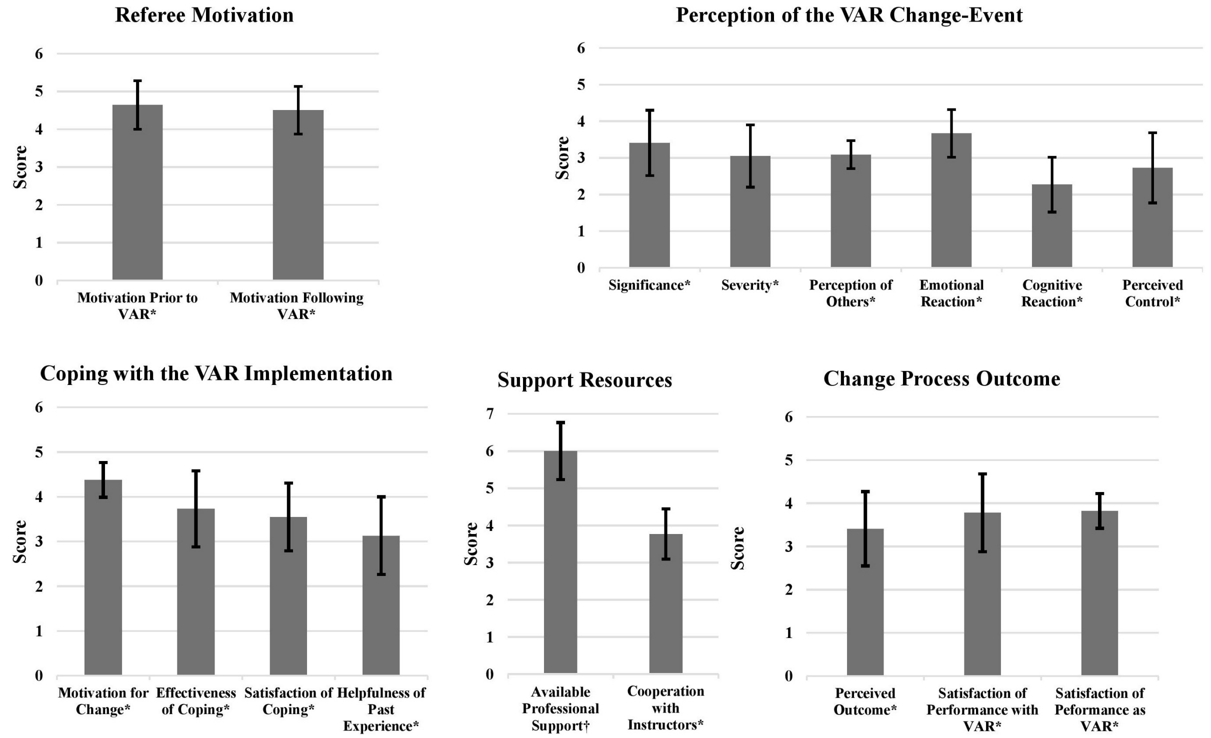

FIGURE 1 | Means and SDs for the referees' motivation **(upper left panel)**, perception of the video assistant referee (VAR) change-event **(upper right panel)**, coping with the VAR change-event **(lower left panel)**, support resources **(lower middle panel)**, and change process outcome **(lower right panel)**. *Scale is 1–5, †Scale is 1–7.

invested. They were informed, however, that the VAR integration would produce additional officiating options in the future, so the initial instruction investment was worthwhile.

During this period, which occurred in the active parts of the 2018–2019 season, the referees hardly received any regular refereeing instruction. This impacted their immediate performance, even before the actual application of the VAR system in live matches. From the project manager's standpoint, the main challenge was to uncover the ideal balance between the necessary knowledge required for a successful training and keeping the tight schedule. It was clear that a 12-month period was much more realistic and suitable for such a complex training process than 6 months. To develop proficiency, the VARs are required to watch many match clips and acknowledge the correct level of intervention. The referees felt a strong sense of duty and responsibility to "deliver back the match" to the Referee Union. As one of them reflected: "there is zero room for errors in the VAR, while on-field errors get more tolerance." Moreover, the VARs were somewhat "the protective shield," making sure the matches came to a safe conclusion, without any club or public scrutiny. As one of them acknowledged:

> The VAR is ungrateful, challenging. There is more pressure than on the field. . . I make decisions for someone else—it's more complicated. There is a gap in the attitude (toward the match), and there's a will not to screw the referee.

In addition, after several weeks in the 2019–2020 season, the referees were beginning to experience an officiating overload, as they were required to officiate as both head referees and as VARs in the same league round, in consecutive days. This created high mental demand; they lacked time to recover from one match and already were required to participate in another match.

Moreover, as the VAR is not only a professional refereeing system but also a social–ethical type of system, a need was raised to educate the "world of soccer" (i.e., the clubs, the teams, the media, the broadcasting staff, the commentators, and the fans and general public) about the new system; when and how it is being used. In this context, the Israel Referee Union was required to set the "correct" intervention level for the various types of match incidents in order to reach a degree of unity among all Premier League referees. The purpose was to create some balance between the existing UEFA intervention level and the one suitable for the domestic league. For example, in UEFA, the requirement is for minimum intervention (i.e., a very high intervention level), yet the Israeli clubs and media called for a low intervention level, as they wished to experience the system's full influence. Thus, the Referee Union decided to publish after each league round the professional committee's verdicts concerning the correct match decisions and the critical errors in which the VAR system was involved for better and worse. However, this did not reduce the pressure coming from the clubs and media when they perceived that an error was made. Alternatively, it stressed one of the main issues with the VAR system—that fairness is a relative term influenced by a sociocultural context. As one of the referees commented: "VAR is not an absolute justice machine, but a tool to prevent errors in critical decisions."

On-Field Demands

Taking the VAR role, the referees were required to learn how to manually operate the video and audio systems, how to quickly and efficiently identify match incidents, when and how to interfere within the match, and how to effectively communicate with the on-field referees. They were also required to adapt themselves to sitting quietly in the VOR (i.e., a very small and somewhat claustrophobic location) for a long duration and manage their arousal and stress differently than they were used to (i.e., they could not just run to reduce their arousal levels like they typically do on the field). In this context, the referees identified professional demands such as being highly familiar with the VAR protocol, having patience and restraint, communicating correctly in the audio system, being knowledgeable of the Laws of the Game specific criteria, quick use of the video system, not getting disconnected from the match spirit, understanding the match, picking the correct camera angle, making correct decisions, and being a fast learner to close gaps as quickly as possible. It is important to note that when VAR intervenes, the DM is critical, which further increases pressure. However, mere accurate decisions might not be enough, as one of the referees explained:

> In the van, you must have a macro vision—to be able to understand how this is being perceived by the public. Few low criteria are not sufficient; you must have lots of "meat" and a good reason for the decision to bring the story.

Another professional demand being a VAR was to be able to adjust to different referees in terms of level of intervention and communication style. Each referee expressed a different approach from the VAR, yet there were strict IFAB communication protocols. The VARs were not allowed to deliver certain information, which did not apply directly to their areas of responsibilities. Thus, serving as VARs, the referees were required to adjust between the demands of the specific head referee, the guidelines of the Referee Union professional committee, and their own refereeing philosophy to create the optimal level of intervention.

From the on-field referee standpoint, the referees must have learned the VAR protocol and accordingly adjust their refereeing. This meant modification of body language and new signing, as well as waiting after potential critical match incidents to hear from the VAR that the match could be continued. Also, they were required to manage a larger refereeing team, modifying their pre-match instructions to explain the level of intervention they were aspired to achieve. In addition, the referees were expected to adapt to having a "big brother" who watches their decisions and might intervene and correct them. They were required to adjust to a new situation in which every decision was not finite and could be corrected. Waiting for the VAR to check potential match incidents while being mobbed by the players was initially challenging, as the protocol was new for all stakeholders. They also must have adjusted to the OFR, which required them to go and check the match incident in the Referee Review Area (RRA) in a relatively short time period (i.e., up to 30 s is considered very good). In terms of DM quality, when the referee goes to the OFR, s/he shifts from an on-field DM to a video DM, which requires the referees to learn this new methodology. A further major demand was to first acknowledge an error in a critical decision and then to continue officiating the match with confidence and focus. Knowing that going to the OFR to correct a critical decision would result in a poor performance mark was also challenging.

Off-Field Barriers

There were also several *barriers* the referees experienced in their attempts to adapt to the VAR system. First, it is the relative lack of experience with the VAR system as well as the break formed in the implementation process (i.e., between the 2018–2019 playoffs and the beginning of the 2019–2020 season). Also, there were not enough VAR-related instruction clips, and the instructor lacked a definitive correct response for each VAR situation (i.e., unlike the typical refereeing instruction). This initially created frustration and reduced self-efficacy beliefs among the referees. As each country implemented the VAR system rather differently, it was difficult to learn much from the international colleagues' experience.

The referees also mentioned the soccer world's lack of understanding of the new system. In many cases, the referees performed according to the IFAB protocol, producing an accurate DM process (i.e., either to use the OFR or not, whether to modify their original decisions or not), yet still the soccer world perceived that the process to be incorrect or inaccurate. One of the referees commented on this issue:

> There is much focus and screen time that goes to us now, and we must perform this in the most efficient way because the media and the whole soccer world would be much less tolerant toward us if we maintained an error even after we watched [on video].

In addition, the VAR learning process did not include a gradual sequence comprised of clear goals, levels of difficulty, rewards, and immediate feedback. This stemmed from the system being novel not only to the referees but also to the Israel Referee Union and to its professional committee. Thus, the referees were unsure, at the beginning, how the match marks of both the on-field and the VAR performances would be calculated and what would be the rewards and punishments for successful and poor performances. In fact, as the system was highly pressured by the Israeli soccer world, excessive attention was given to punishments and scrutiny, and the referees were hardly commended and rewarded for their efforts. This led to much frustration and, in some cases, to reduced motivation to officiate both on-field and as VARs.

Finally, there were also technical–technological issues in some of the stadiums and matches. Unlike the major matches, in the minor matches, there were only four cameras involved in broadcasting, thus certain important view angles were missing. In two stadiums, there were also issues with the system's capacity to produce an accurate line grid of the field to enable valid video-based offside decisions. Also, in certain matches, the system could not be initiated on time, and the matches started without its use.

On-Field Barriers

From the on-field referee's standpoint, several referees wished to officiate the matches without getting corrected by the VAR. Moreover, they did not accept the VARs' corrections in certain cases and maintained their original decision. This had stemmed from the gaps in interpretation of key match incidents among the various referees. As one of them commented:

> I feel that the VARs are not synchronized with my level of foul decisions on the field. I tell the VAR—"if it takes you a long time then it's not clear." If I see in the OFR what I saw on the field, then I would not change my decision.

The discrepancies between the on-field referees and the VARs had also stemmed from the fact that being called to the OFR and changing one's decision resulted in a low match mark. Furthermore, once the referees were starting to contradict the VAR's ruling, the Israeli media began to claim that the referees had interpersonal issues and "ego problems," which affected their cooperation (Lipkin, 2020). As one of the referees commented: "every time that there is a wrong intervention of the VAR, it negatively affects the authority of the system." In certain cases, there were repercussions on behalf of the Referee Union professional committee, with the referee getting suspended from officiating 1–3 matches. This had both financial and morale impact on the referees.

Certain referees also modified their performance to accommodate the new system in an undesired manner. For example, they issued cards and called penalties not according to their own perceptions but rather to how these events would be perceived by the VAR. Some referees even began taking less critical decisions on-field, knowing that the VAR would correct them if necessary. This was exemplified in the following quote:

> VAR blurs the gaps between those who are good and those who are mediocre in refereeing. There are those who crash the matches and continue to officiate. You must encourage excellence and high skills. Now the situation is that referees run on the field just waiting for the VAR to make decisions for them.

In this context, certain referees had also begun stopping the match with every minor infringement to prevent "a problem" in the Attacking Possession Phase, in case of a potential VAR check. As one of them commented: "as a referee, you might change your officiating because of VAR—take decisions that prior to the VAR aera you wouldn't take."

In addition, there were significant gaps in interpretations between the referees and the professional committee. For example, the latter expected the referees to issue red cards in all "above ankle stamping" incidents, even if malice or excessive force was not evident (i.e., unlike what the official law criteria suggested). Certain referees, however, had not accepted these instructions, and when called to the OFR did not modify their original decisions. These actions, in turn, resulted in the referee being suspended. These incidents further intensified the lack of clarity as well as undermined the referees' belief in the new system: "in the OFR, I go into a dilemma—to stay with the red card that might not be an error; to change or not? The conflict is loss vs. profit."

There were also errors of the VARs. These stemmed from ineffective communication with the on-field referee, low skills with the video system, as well as performance anxiety. As one of the referees commented: "the errors stem from the VARs' desire to please the on-field referee, to not just call for the OFR for a long-time... and there is a lack of skill." In this context, the professional committee suspended VARs who made critical errors. This frustrated some of the referees, as reflected in the following quote:

> A freaky method in which they publicly suspend referees during the system's implementation process. This promotes "tricks" and unprofessional mode of actions on behalf of the referees... they had to decide ahead of time what would be the suspension method in cases of VAR errors.

Decisions and Coping

When the referees were firstly informed of the upcoming integration of the VAR, they were already aware of its effects in other countries. Thus, as part of their initial response to the new situation, nine of the referees decided to consult with others, mainly with colleagues, a refereeing mentor, the professional committee and the Referee Union, and a sport psychologist. Two of the referees decided to self-cope as part of their initial strategy. One of them explained that "initially, I tried to figure out what this means, evaluate the new situation, and check how much the world of refereeing is about to change." For another referee, "it was a challenge to be successful with the VAR system." Furthermore, nine of the referees made a decision to change (i.e., apply all necessary adjustments in response to the new situation). For example, one of them commented: "I contemplated and realized the changes and then "hugged" [the new system] as the right thing to do with many advantages that can assist me." Two of the referees decided to consult with others as their main coping strategy. This also reflected in their comments: "stick to the guidance of the professional committee; tips and recommendations of the sport psychologist."

Referring to **Figure 1**, it seems the referees maintained high motivation for change. Still, their effectiveness and satisfaction of coping were only moderate. They also felt that their experience had moderately facilitated their current coping efforts. There was a positive association between the helpfulness of experience in similar situations and the current coping effectiveness, $r = 0.70$, $p < 0.05$.

The referees reported various types of coping strategies, mostly adaptive: accepting the VAR system as a fact, taking advantage of the its benefits, consulting with professional resources, reevaluation of the refereeing role—"we are here to serve the teams and justice must be seen," mental preparation for matches, adapting oneself to the professional guidelines, becoming highly knowledgeable of the VAR protocols, watching many VAR clips, self-analysis of VAR performance and communication, and live performances.

Support Resources

Data from **Figure 1** suggest that the referees experienced high availability of professional support resources (e.g., sport psychologist, professional mentor, and refereeing coach). They also rated the helpfulness of the sport psychologist and professional mentor as high. Nevertheless, they rated their degree of cooperation with the professional committee as moderate. On one hand, the professional committee was responsible for providing the referees with the technical guidelines related to the VAR application. On the other hand, it monitored the referees' performances and sanctioned them when errors were made. This had created a degree of mistrust between the referees and the professional committee. The referees' cooperation with the professional committee correlated positively with the effectiveness of their coping ($r = 0.62$, $p < 0.05$). This indicates the importance of a positive cooperation in such a change-event, which involves new instructions and technical adaptation.

Change Process Outcomes

Frequency counts (per match) were computed for all variables of interest describing the referees' performance with the VAR system (**Figure 2**). The frequency of *Overall Critical Errors* was higher for Matches 1–45 compared to the pilot phase of the program (Matches 1–30); $\chi^2(3) = 20.35$, $p < 0.01$. The number of *Correct VAR Interventions* was also higher for Matches 1–45 compared to the pilot phase of the program (Matches 1–30); $\chi^2(3) = 18.92$, $p < 0.01$. Hence, the increase in *Overall Critical Errors* for Matches 1–45 was positively balanced by an increase in the number of *Correct VAR Interventions* for these same matches. No other statistical effects were observed, but the overall trend for all variables was positive over time (**Figure 2**). Specifically, Incorrect VAR Interventions, Missing VAR Interventions, and Incorrect OFRs showed the lowest values at the last phase of the program (Matches 136–182). *Total OFRs* and *Correct OFRs* increased after the pilot phase of the program and showed a mirrored fluctuating pattern afterward (**Figure 2**). Therefore, as the referees became more proficient with the VAR system, there was some improvement in performance. Also, the average match time suspension due to VAR interventions decreased as the season progressed, further indicating the referees' technical and DM proficiency. Nevertheless, these were small improvements not always recognized by the soccer world and the referees themselves.

On average, the referees perceived the outcome of this change-event as relatively neutral in the context of their careers (**Figure 1**). Yet, looking at the variance, it is evident that three referees perceived it positively, four perceived it neutrally, and four negatively. The referees were also moderately satisfied with their match performances both on-field and as VARs. Moreover, a positive correlation between the referees' satisfaction from their on-field performances and the perceived outcome of this change-event, $r = 0.67$, $p < 0.05$, emerged. This might indicate that the referees were more concerned about the effects of the VAR system on their on-field performances. Finally, their motivation for refereeing was high and remained so even after the VAR implementation (**Figure 1**).

DISCUSSION

With the growing recognition of soccer referees as performers on their own merit (Aragão e Pina et al., 2018), research attention was given to various aspects of their roles, including career development (Samuel et al., 2017). The introduction and implementation process of the VAR system within the Israeli Premier League reflected a technical–technological change-event (i.e., a quasi-normative transition) in the careers of the Israeli elite referees. This change-event introduced potential implications for training, performance, public status, and professional advancement (Armenteros and Webb, 2020; de la Vega and Fuentes, 2020).

Based on Kolbinger and Lames (2017) taxonomy, there are technologies that support the DM process of referees. Technologies are used to replace referees for a specific decision, and technological aids help the referee to enforce rules (e.g., the vanishing spray in soccer). The VAR system is supposedly related to the first classification, as it was designed to improve referees' DM by offering additional view angles as well as video replays. However, it might also set conditions for replacing the referee in certain basic decisions, such as offsides and goals. Furthermore, knowing that they are being watched and supervised by the VAR, the players are committing less fouls (Lago-Peñas et al., 2019), supporting the third classification as well.

The official data concerning the implementation of the VAR system in Israel indicated that, in 212 matches, there were 89 critical errors in total. This rate was higher than that in previous years, which could be attributed to the introduction of the VAR and the detection of more key match incidents, as well as to lower on-field refereeing quality due to the adaptation process. The initial use of the system was far from perfect, with error rates being higher than expected. In the Italian implementation phase, for comparison, the reported level of error was 1%, with 1,708 decisions being reviewed over 210 matches, which led to 60 decisions being corrected, 11 of those being wrong, including seven that influenced the outcome of the game (Simón, 2020).

The Israeli Premier League referees perceived the VAR implementation as a career change-event characterized by a moderate emotional profile. This change-event had mainly influenced the technical–technological aspect of their performance, with moderate effects on pre-match preparation, players' management and communication, public perception, and DM. These areas of refereeing might be associated with the referees' authority and credibility. Previous research suggested a potential threat of technological officiating aids to referees' authority (Kolbinger and Lames, 2017). Nevertheless, the effects on career development were minor. Moreover, research on referees' careers indicated that change-events related to career development were perceived as highly significant, whereas performance-related issues were perceived moderately (Samuel et al., 2017; Samuel, 2019). Also, a recent study on the regulations

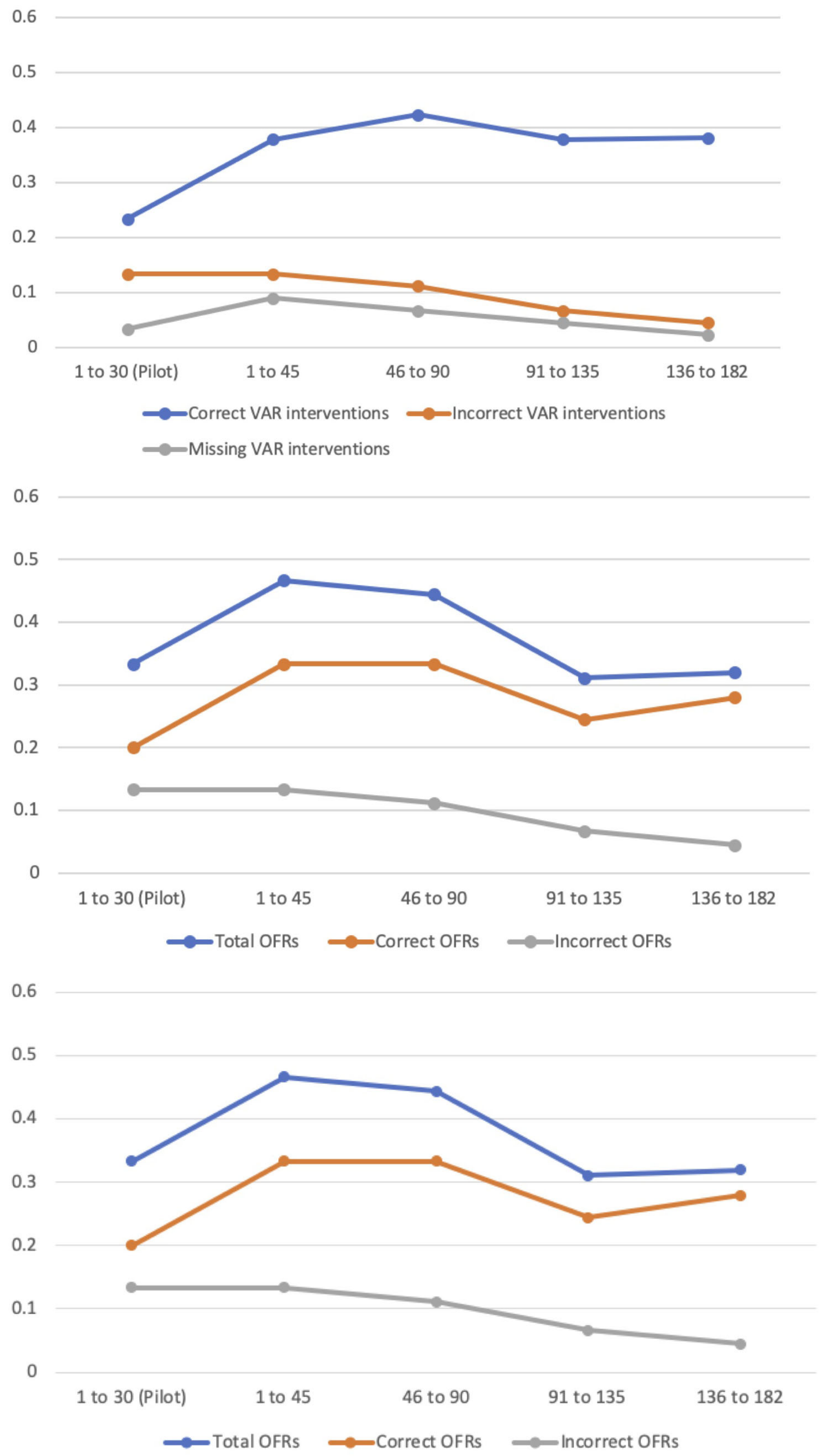

FIGURE 2 | Per match frequency count of Overall Critical Errors and Rectified On-Field Critical Errors **(upper panel)**; Correct Video Assistant Referee (VAR) Interventions, Incorrect VAR Interventions, and Missing VAR Interventions **(middle panel)**; and Total On-Field Reviews (OFRs), Correct OFRs, and Incorrect OFRs **(lower panel)**.

and refereeing modifications in competitive judo showed that the judokas and coaches perceived them with a moderate emotional profile (Samuel et al., 2020a). Thus, the findings of the present study are in line with research indicating that technical–technological modifications are being perceived with a moderate emotional profile, with modest effects on referees' performance as well as authority. These findings, therefore, provided support for the application of the SCSPP (Samuel and Tenenbaum, 2011a) in the context of referees' careers.

In addition, the emotional and cognitive perceptions of the referees were associated with their perceived control, their experience with similar situations, as well as their age. It seems that as the referees were older and more experienced, they perceived this change-event more favorably. In a sample of Swedish soccer referees, Folkesson et al. (2002) found that younger referees were shown to be the most prone to threat and aggression. It is possible that the younger Israeli referees in the present study were also more apprehensive of their ability to effectively manage players and make accurate decisions upon VAR implementation.

The findings indicated that most off-field demands were related to the instruction and operation of the VAR system. This change-event required the referees to quickly and efficiently integrate a new method of refereeing, accepting its importance for both the soccer world and the Referee Union. Fernández Ruiz et al. (2020) advised that as part of developing the VAR education curriculum, "learning activities for referees can be distributed over time, and with the aid of learning management systems and other applications, instructors can control this distribution, programming specific dates and completion times for the activities" (p. 337). However, in the Israeli case, the 12 months' IFAB protocol was shortened by half, presenting high pressure and demands. The referees were, therefore, under considerable workload while learning the new system and then attempting to adjust to it. They were asked to significantly increase the amount of active officiating in each league round, which further led to overload. This, in turn, resulted in lower refereeing quality and criticism from the Referee Union professional committee. These findings echo Baldwin's (2014) study of Australian rugby league referees, who reported much stress related to the education of the video replay technology.

Another important demand was to educate the soccer world about the new system. Although there was a standard IFAB protocol, each country implemented the VAR system somewhat differently (Armenteros and Webb, 2020). At first, the Israeli soccer world expected the VAR system to intervene in all decisions, which was not in line with the IFAB protocol. This created pressure on the referees and the professional committee. Thus, the referees and the VAR system were under much scrutiny in terms of what Kolbinger and Lames (2017) identified as *standard of review*—the influence of the initial call of the referee on the review process and its outcome.

In terms of the on-field demands, the referees were required to quickly learn a new method of officiating as VARs, which is significantly different from their typical on-field refereeing task. Adapting to this new system required forming a unique man–machine interface as well as developing specified DM skills. Fernández Ruiz et al. (2020) suggested that referees must apply 21st century skills (e.g., technical, information management, and critical thinking) while performing the VAR task. In this context, it is unclear what was the referees' degree of *digital competence and efficacy* (Janssen et al., 2013) when faced with the technological demands. Future research must, therefore, focus on analyzing the most effective DM process for the on-field and VAR refereeing task, considering the quick transformations between sequential DM and video-based DM. A critical aspect of these transformations is how to include the contextual match aspects to achieve not only correctness but also fairness (Samuel et al., 2020b).

Moreover, the referees were expected to accept that a "big brother" was watching their decisions, not necessarily intending to overrule them, but rather to support and assist them. This required a major shift in the referees' mindset. Samuel et al. (2020b) proposed that, in the final phase of the DM sequence, referees may decide to keep or change their decision. As previous research indicated, referees are individuals with high trait and state self-control (Samuel et al., 2018). Thus, asking them to incorporate the VAR into their on-field refereeing meant trusting the VARs' professional quality and personal integrity. This required "a leap of faith," that at least at the beginning of this process was highly demanding for most of them. Research findings indicated that *making a controversial call* and *making a wrong call* are considered high stressors for soccer referees (Voight, 2009), potentially leading to increased anxiety (Yun and Jeon, 2016). As the elite refereeing environment is highly competitive (Samuel et al., 2017), the referees were apprehensive about getting corrected by the VARs. Still, they reported that their pressure decreased as a result of the VAR integration, which might indicate that they preferred to end a match with a rectified critical error than letting such an error remain.

The barriers involved in this change-event corresponded to the on- and off-field demands. The off-field barriers included the educational process, gaps in interpretations with the world of soccer, lack of clear goals and rewards system, and technical–technological setbacks. All these barriers were previously identified in other VAR implantations (e.g., Baldwin, 2014; Fernández Ruiz et al., 2020; Kolbinger, 2020; Simón, 2020). For example, in the German Bundesliga VAR implementation, the head of the elite referees committee, Lutz Michael Fröhlich, concluded after the first five match days that "a lot of players and team officials still struggle to differentiate between apparently similar situations, which are in fact different and therefore must be treated differently by the VAR protocols" (Kolbinger, 2020, p. 233). Thus, discrepancies in interpretations concerning the application of the VAR were evident in other leading countries, not only in Israel.

The on-field barriers included gaps between the referees and the VARs, modifications of the referees' typical officiating style, gaps between the referees and the professional committee, and errors of the VARs. Some of these barriers were also previously identified (Baldwin, 2014; Armenteros and Webb, 2020; Bacigalupe, 2020). For example, the German elite referee committee decided that "for difficult calls, which can't be rated indisputably as 'obvious mistakes,' but the video assistant doubts the initial decision by a high degree, he shall immediately contact the referee to initiate an on-field review" (Kolbinger, 2020, p. 233). This quote suggests some gaps in interpretations between the German referees and the professional committee. Furthermore, Bacigalupe (2020) proposed that the VAR is under pressure "as they are the watchdog—and therefore ultimately responsible—who is there to make sure nothing that happens on the pitch goes unseen" (p. 201). Unlike the on-field referee who is given some latitude when erring, there is no tolerance for VAR errors. A challenge is not to fall into a trap of justifying the on-field referee's decisions.

There were also unique barriers for the Israeli referees. For example, the referees' lack of adherence to the VARs' calls, perhaps due to match mark interests or to ego-protecting interests, was not discussed in the literature so far. These motives are typical for the highly competitive Israeli elite referees (Samuel et al., 2017). In this context, the system of rewards for VAR officiating was unclear, with high focus on sanctions and scrutiny and too little praise. For almost all referees, this led to a negative motivational process and a heightened fear of failure (Samuel, 2019).

Also, the modification of officiating style has not been recognized by researchers yet. The Israeli referees modified their regular style both on-field and as VARs to decrease potential errors and increase their level of accountability. Once more, these effects are related to the unique Israeli context where the referees did not wish to be criticized by the professional committee as well as to sustain lower match marks.

Even though this change-event was not perceived as highly demanding, it required coping efforts. Most referees made active decisions as part of their coping, initially to consult with professional resources and subsequently to apply all necessary changes to effectively cope. These decisions manifested in application of adaptive coping strategies. These findings are in line with the literature on career change-events in referees (e.g., Samuel et al., 2017; Samuel, 2019). Thus, it seems that the referees assumed an active approach in their attempts to adapt to the VAR system. In addition, while the referees felt they have had high availability of professional support, their cooperation with the professional committee was lacking, negatively impacting the effectiveness of their coping. The professional committee had a dual role in the VAR implementation process, which included (a) establishing contextualized implementation standards for the Israeli league (i.e., an educational role) and (b) regulating misapplications of the system (i.e., a punitive role). These findings, therefore, emphasize the role of the domestic professional committees in providing the right support for referees who undergo a technical change. This support should be manifested in setting clear goals for the change process, being tolerant toward referees' errors and confusions and protecting the referees from scrutiny of the media and other stakeholders.

In terms of the change process outcomes, the Israeli referees showed small, yet important, improvements in their performance. However, they did not receive credit for them. On the contrary, in many cases, they experienced scrutiny from the clubs, the media, and the Referee Union's professional committee. Furthermore, the referees varied in their perception of this change process outcome, with three perceiving it positively, four neutrally, and four negatively. These findings further demonstrate the dynamic and probabilistic nature of the change process (Stambulova and Samuel, 2020). An important factor was their degree of satisfaction from their on-field performance since the VAR was integrated. This finding is on a par with Samuel et al.'s (2020a) study on the regulations and refereeing modifications in competitive judo; the judokas' perceptions of the outcome of that change-event were associated with their perceptions of professional achievements since these modifications took place. These findings are important, as they explain to the referee unions that, in the end, the VAR is here to serve the on-field refereeing, not the other way around. Referees are motivated to officiate the matches on the field, applying their long-learned skills and experience. Finally, the referees maintained high motivation following the VAR integration into the domain of refereeing.

To conclude, much of the above findings are in line with the emerging literature on technological officiating aids (Kolbinger and Lames, 2017; Samuel et al., 2020a), including VAR (Armenteros et al., 2020), as well as with the literature on referees' change-events (Samuel et al., 2017; Samuel, 2019). Generally, these findings emphasize that implementing the VAR presents a moderate change-event in soccer referees' careers that require a degree of adaptation in terms of active DM, coping, and support.

Study Limitations

We must consider limitations to the current study. First, while the sample was highly representative of the entire Israeli Premier League squad, not all referees participated in the study. Those who responded received psychological support, and it is possible that this had influenced their change process. Thus, sampling referees who do not receive psychological support throughout the VAR implementation is imperative. Also, the study is highly contextualized within the Israeli soccer culture. While contextualizing career-related studies is called for (Stambulova et al., 2020), it might also influence the potential for generalizing the findings to other populations. Moreover, all participants were males, and gender effects were not examined. Thus, examining how VAR is perceived and implemented in other countries and refereeing populations is of interest, while adopting an empathetic stance and realizing the particularities of each country. Also, there might be other potential factors not examined in this case study that might influence referees' perceptions of change processes, such as referees' salaries and monetary and professional rewards (Fernández Ruiz et al., 2020), as well as digital competence (Janssen et al., 2013). Finally, we did not collect quantitative data in several time points, as conducted in previous change-event studies (e.g., Samuel et al., 2015). It is possible that collecting data after a longer time period following the VAR integration would allow a different perspective on its effects on referees' careers. Future studies that accompany referees' change process are advised to apply a full longitudinal design. Also, as there were several CEI subscales with low reliability indices, related correlational data must be cautiously viewed.

Practical Recommendations

Referee unions should provide referees with the adequate support to be able to respond to this change favorably. First, to acknowledge the importance of the VAR and its correct role—not to replace on-field officiating, but to assist it. Second, to recognize the key refereeing areas, which such a change might impact. In the present case, the largest effects of the VAR system integration were in pre-match preparation and players' management. These are two areas, as well as DM, in which referees can progress, given the appropriate training support (e.g., Cunningham et al., 2014; Samuel, 2015). Still, referees in other countries might experience modifications in additional areas. It is advised to construct a gradual educational process, not as intense as in the Israeli case. This could be optimized by setting a clear system of goals, providing ample feedback, establishing rewards and sanctions, and considering the importance of praise over scrutiny. In this context, we recommend referee unions to allow referees a continued training process following the IFAB certification. This would entail the development of novel training modalities, in addition to existing VAR simulators, so referees can continue independent practice. New online options, such as Zoom-based training (see Samuel et al., 2020c), can be exploited for this purpose. Analyzing the refereeing team communication in critical match incidents is of high importance. Relying on a clear statistical dataset to assess individual and collective progress of VAR-related skills is also important. Psychological support must focus on accepting the change and recognizing its impact both on- and off-field. Then, making a conscious decision to change and applying it in the ordinary refereeing routine. Finally, as VAR is not only a professional system but also a social–ethical one, much consideration should be given to education of the soccer world concerning its application. The clubs, fans, and media must be more knowledgeable of how and when the system is used to produce more correct, just, and fair decisions.

DATA AVAILABILITY STATEMENT

The datasets for this article are not publicly available due to the high profile of participants and confidentially of data. Requests to access the datasets should be directed to RS, roydsamuel@gmail.com.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RS wrote the literature review. RS and EF wrote the methods and findings. RS, YG, EF, and GT wrote the conclusion and discussion. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors would like to thank the Referee Union in the Israel Football Association for their support of data collection. A special gratitude to the Israeli VAR project manager, Mr. Shmuel Shteif, for his assistance in developing the manuscript.

REFERENCES

APA (2017). *Ethical Principles of Psychologists and Code of Conduct*. Washington, DC: American Psychological Association.

Aragão e Pina, J. A., Passos, A.-M., Carvalho, H., and Maynard, M. T. (2018). To be or notto be an excellent football referee: different experts' viewpoints. *J. Sports Sci.* 37, 692–700. doi: 10.1080/02640414.2018.1522940

Arias, J. L., Argudo, F. M., and Alonso, J. I. (2011). Review of rule modification in sport. *J. Sports Sci. Med.* 10, 1–8.

Armenteros, M., Benitez, A. J., and Betancor, M. (2020). *The use of Video Technologies in Refereeing and Football and Other Sports*. New York, NY: Routledge. doi: 10.4324/9780429455551

Armenteros, M., and Webb, T. (2020). "Educating international football referees: the importance of uniformity," in *The use of Video Technologies in Refereeing Football and Other Sports*, eds M. Armenteros, A. J. Benítez, and A. M. Betancor (New York, NY: Routledge), 301–327. doi: 10.4324/9780429455551-16

Bacigalupe, C. (2020). "The video assistant referee (VAR) protocol," in *The use of Video Technologies in Refereeing Football and Other Sports*, eds M. Armenteros, A. J. Benítez, and A. M. Betancor (New York, NY: Routledge), 183–207. doi: 10.4324/9780429455551-10

Baldwin, C. F. (2014). Video killed the sports referee: the role of video replay technology in creating anxiety and stress amongst rugby league referees. *Int. J. Arts Sci.* 7, 305–312.

Bradbury, S. (2020). *'Oh s**t' - Howard Webb Admits Mistake in Man United Game That Cost Liverpool. Echo*. Available online at: https://www.liverpoolecho.co.uk/sport/football/football-news/oh-st-howard-webb-admits-18202388 (accessed May 20, 2020).

Cunningham, I., Simmons, P., Mascarenhas, D., and Redhead, S. (2014). Skilled interaction: concepts of communication and player management in the development of sport officials. *Int. J. Spor. Comm.* 7, 166–187. doi: 10.1123/IJSC.2013-0098

de Dios Crespo, J. (2020). "The contributions of VARs to fairness in sport," in *The use of Video Technologies in Refereeing Football and Other Sports*, eds M. Armenteros, A. J. Benítez, and A. M. Betancor (New York, NY: Routledge), 23–35. doi: 10.4324/9780429455551-2

de la Vega, R., and Fuentes, J. P. (2020). "Psychological response of the referee to the introduction of VAR," in *The use of Video Technologies in Refereeing Football and Other Sports*, eds M. Armenteros, A. J. Benítez, and A. M. Betancor (New York, NY: Routledge), 339–350. doi: 10.4324/9780429455551-18

Fernández Ruiz, M., Betancor, M. Á, and Armenteros, M. (2020). "Didactic teaching strategies for the VAR: an approach to gamification and 21st-Century skills," in *The use of Video Technologies in Refereeing Football and Other Sports*, eds M. Armenteros, A. J. Benítez, and A. M. Betancor (New York, NY: Routledge), 328–338. doi: 10.4324/9780429455551-17

Folkesson, P., Nyberg, C., Archer, T., and Norlander, T. (2002). Soccer referees' experience of threat and aggression: effects of age, experience, and life orientation on outcome of coping strategy. *Aggress. Behav.* 28, 317–327. doi: 10.1002/ab.90028

Hodge, K., and Sharp, L. (2016). "Case studies," in *Routledge Handbook of Qualitative Research in Sport and Exercise*, eds B. Smith and A. C. Sparkes (London: Routledge), 62–74.

Holt, N., and Sparkes, A. (2001). An ethnographic study of cohesiveness in a college soccer team over a season. *Sport Psychol.* 15, 237–259. doi: 10.1123/tsp.15.3.237

International Football Association Board [IFAB] (2018a). *Historic Step for Greater Fairness in Football*. Available online at: http://www.theifab.com/news/historic-step-for-greater-fairness-in-football (accessed May 20, 2020).

International Football Association Board [IFAB] (2018b). *Information on the Video Assistant Referee (VAR) Experiment, Incl. Provisional Results*. Available online at: https://static-3eb8.kxcdn.com/documents/639/165902_220118_IFAB_Media_Package_ABM2017_all_media_FINAL.pdf (accessed May 20, 2020).

International Football Association Board [IFAB] (2017). *Video Assistant Referees (VARs) Experiment – Protocol (Summary)*. Available online at: http://static-3eb8.kxcdn.com/librarydocuments/12/211001_120517_VAR_EN_lang_Spreads.pdf (accessed May 20, 2020).

Janssen, J., Stoyanov, S., Ferrari, A., Punie, Y., Pannekeet, K., and Sloep, P. (2013). Experts'views on digital competence: commonalities and differences. *Comput Educ.* 68, 473–481. doi: 10.1016/j.compedu.2013.06.008

Kolbinger, O. (2020). "VAR experiments in the Bundesliga," in *The use of Video Technologies in Refereeing Football and Other Sports*, eds M. Armenteros, A. J. Benítez, and A. M. Betancor (New York, NY: Routledge), 228–245. doi: 10.4324/9780429455551-12

Kolbinger, O., and Lames, M. (2017). Scientific approaches to technological officiating aids in game sports. *Curr. Issues Sport Sci.* 2:001. doi: 10.15203/CISS_2017.001

Lago-Peñas, C., Rey, E., and Kalén, A. (2019). How does video assistant Referee (VAR) modify the game in elite soccer? *Int. J. Perform. Anal. Sport* 19, 646–653. doi: 10.1080/24748668.2019.1646521

Lipkin, G. (2020). *A crisis: Lack of Trust Between the Referees and the VARs [Hebrew].One*. Available online at: https://www.one.co.il/Article/19-20/1,1,1,0/353244.html (accessed May 20, 2020).

Patton, M. (1999). Enhancing the quality and credibility of qualitative analysis. *HSR Health Serv. Res.*34, 1189–1208.

Patton, M. (2002). *Qualitative Research & Evaluation Methods*, 3rd Edn. San Francisco, CA: Sage Publications.

Samuel, R. D. (2015). A psychological preparation framework for elite soccer referees: a practitioner's perspective. *J.Sport Psychol. Act.* 6, 170–187. doi: 10.1080/21520704.2015.1065938

Samuel, R. D. (2019). Soccer referees' transition to the premier league: a case study reflecting individual experiences and consultancy. *Case Stud. Sport Exerc. Psychol.* 3, 1–10. doi: 10.1123/cssep.2018-0017

Samuel, R. D., Basevitch, I., Wildikan, L., Prosoli, R., and McDonald, K. (2020a). Please stop changing the rules! The modifications of judo regulations as a change-event in judokas' and coaches' careers. *Sport Soc.* 23, 774–794. doi: 10.1080/17430437.2019.1669911

Samuel, R. D., Tenenbaum, G., and Galily, Y. (2020b). An integrated conceptual framework of decision-making in soccer refereeing. *Int. J. Sport Exerc. Psycholo.* doi: 10.1080/1612197X.2020.1766539. [Epub ahead of print].

Samuel, R. D., Tenenbaum, G., and Galily, Y. (2020c). The 2020 Coronavirus pandemic as a change-event in sport performers' careers: conceptual and applied practice considerations. *Front. Psychol.* 11:567966. doi: 10.3389/fpsyg.2020.567966

Samuel, R. D., Engelrt, C., Zhang, Q., and Basevitch, I. (2018). Hi ref, are you in control? Self-control, ego-depletion, and performance in soccer referees. *Psychol. Sport Exerc.* 38, 167–175. doi: 10.1016/j.psychsport.2018.06.009

Samuel, R. D., Galily, Y., and Tenenbaum, G. (2017). Who are you, ref? Defining the soccer referee's career using a change-based perspective. *Int. J. Sport Exerc. Psychol.* 15, 118–130. doi: 10.1080/1612197X.2015.1079792

Samuel, R. D., and Tenenbaum, G. (2011a). The role of change in athletes' careers: a scheme of change for sport psychology practice. *Sport Psychol.* 25, 233–252. doi: 10.1123/tsp.25.2.233

Samuel, R. D., and Tenenbaum, G. (2011b). How do athletes perceive and respond to change-events: an exploratory measurement tool. *Psychol. Sport Exerc.* 12, 392–406. doi: 10.1016/j.psychsport.2011.03.002

Samuel, R. D., Tenenbaum, G., Mangel, E., Virshuvski, R., Chen, T., and Badir, A. (2015). Athletes' experiences of severe injuries as a career change-event. *J. Sport Psychol. Action.* 6, 99–120. doi: 10.1080/21520704.2015.1012249

Sánchez Cid, M., and García García, A. L. (2020). "The interaction between audio and the video assistant referee in football," in *The use of Video Technologies in Refereeing Football and Other Sports*, eds M. Armenteros, A. J. Benítez, and A. M. Betancor (New York, NY: Routledge), 70–85. doi: 10.4324/9780429455551-4

Schweizer, G., and Furley, P. (2016). Reproducible research in sport and exercise psychology: the role of sample sizes. *Psychol Sport Exerc.* 23, 114–122. doi: 10.1016/j.psychsport.2015.11.005

Simón, J. A. (2020). "Between scepticism and the long-awaited quest for "football justice," in *The use of Video Technologies in Refereeing Football and Other Sports*, eds M. Armenteros, A. J. Benítez, and A. M. Betancor (New York, NY: Routledge), 3–22. doi: 10.4324/9780429455551-1

Slack, L. A., Maynard, I. W., Butt, J., and Olusoga, P. (2013). Factors underpinning football officiating excellence: perceptions of English Premier League referees. *J. Appl. Sport Psychol.* 25, 298–315. doi: 10.1080/10413200.2012.726935

Smith, B. (2018). Generalizability in qualitative research: misunderstandings, opportunities and recommendations for the sport and exercise sciences. *Qual. Res. Sport Exerc. Health* 10, 137–149. doi: 10.1080/2159676X.2017.1393221

Smith, B., and McGannon, K. R. (2018). Developing rigor in qualitative research: problems and opportunities within sport and exercise psychology. *Int. Rev. Sport Exerc. Psychol.* 11, 101–121. doi: 10.1080/1750984x.2017.1317357

Sparkes, A. C. (2015). Developing mixed methods research in sport and exercise psychology: critical reflections on five points of controversy. *Psychol. Sport Exerc.* 16, 49–59. doi: 10.1016/j.psychsport.2014.08.014

Sparkes, A. C., and Smith, B. (2014). *Qualitative Research Methods in Sport, Exercise and Health: From Process to Product*. London: Routledge. doi: 10.4324/9780203852187

Stambulova, N., and Samuel, R. D. (2020). "Career transitions," in *The Routledge International Encyclopedia of Sport and Exercise Psychology*, Vol. 2, eds D. Hackfort and R. Schinke (New York, NY: Routledge).

Stambulova, N. B., Ryba, T. V., and Henriksen, K. (2020). Career development and transitions of athletes: the international society of sport psychology position stand revisited. *Int. J. Sport Exerc. Psychol.* doi: 10.1080/1612197X.2020.1737836. [Epub ahead of print].

Tracy, S. J. (2014). Qualitative quality: eight "big-tent" criteria for excellent qualitativeresearch. *Q. Inq.* 16, 837–851. doi: 10.1177/1077800410383121

Voight, M. (2009). Sources of stress and coping strategies of U.S. soccer officials. *Stress Health* 25, 91–101. doi: 10.1002/smi.1231

Yin, R. K. (2014). *Case Study Research: Design and Methods*, 5th Edn. Thousand Oaks, CA: Sage.

Yun, Y.-K., and Jeon, J.-Y. (2016). Psychological experience and coping strategies for K-League referees just after the moment of wrong judgment. *Korean J. Sport Sci.* 27, 957–969. doi: 10.24985/kjss.2016.27.4.957

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Sports and Active Living

ORIGINAL RESEARCH
published: 08 January 2021
doi: 10.3389/fspor.2020.595703

A Preliminary Analysis of the Importance of Distance, Angle, and Insight When Soccer Referees Make Penalty Decisions

Bjørn Tore Johansen and Martin Kjeøen Erikstad*

Department of Sport Science and Physical Education, Faculty of Health and Sport Sciences, University of Agder, Kristiansand, Norway

Soccer referees move freely on the pitch to place themselves in the best location for making decisions. While Football Association UK (FA) highlights that a referee should never be more than 20 m away from the playing situation, previous studies have been inconsistent in indicating appropriate distance to a situation for increasing the likelihood of a correct decision. Further, appropriate angle and insight are also likely to influence the correctness of referees' decisions. The aim of this study was to provide an initial investigation of elite referees' positioning in the field (distance, angle, and insight) when making correct and erroneous decisions in potential penalty situations. An expert panel (EP) consisting of two active referees with relevant academic background analyzed referees positioning when making correct or erroneous decisions regarding penalties. The EP were asked to qualitatively analyze referees positioning in selected video clips by using recommended technical refereeing criteria and practical guidelines (i.e., the referee's distance from, angle to, and insight into the penalty situations). Of the 42 situations evaluated, the results revealed that the EP termed the referees positioning as good in terms of angle and insight in 25 and 21 situations, respectively. The angle was average in seven situations and poor in 10 situations, and the insight were average in 10 situations and poor in 11 situations. The match referee was <10 meters away in 12 situations, 10–20 m away in 22 situations, and >20 meters away in eight situations. Results revealed that referees' positioning that resulted in the highest rate of correct decisions were when the distance were under 10 meters (83% correct decisions), good angle (88%), and good insight (86%). In contrast, referees were poorly positioned in terms of angle and/or insight in nine of the 15 erroneous decisions made. Although the present study was a preliminary qualitative investigation containing a limited number of potential penalty situations, the findings indicated that soccer referees are more likely to produce a correct decision in potential penalty situations when the distance to the situation is under 10 meters, when the insight to the situation is good and the angle to the incident is good. In contrast, the match referees generally had a poor starting position to assess the penalty situations where they landed on a wrong decision. While previous studies have been somewhat inconsistent in indicating optimal referee positioning in soccer, the present study highlights the potential value of a more qualitative approach to understand referees' positioning and subsequent decision-making accuracy.

Keywords: penalty situations, distance, angle, insight, decision-making skills

OPEN ACCESS

Edited by:
Roy David Samuel, Interdisciplinary Center Herzliya, Israel

Reviewed by:
Alexandra Pizzera, German Sport University Cologne, Germany
Elia Morgulev, Ben-Gurion University of the Negev, Israel

***Correspondence:**
Bjørn Tore Johansen
bjorn.t.johansen@uia.no

Specialty section:
This article was submitted to Movement Science and Sport Psychology, a section of the journal Frontiers in Sports and Active Living

Received: *17 August 2020*
Accepted: *09 December 2020*
Published: *08 January 2021*

Citation:
Johansen BT and Erikstad MK (2021) A Preliminary Analysis of the Importance of Distance, Angle, and Insight When Soccer Referees Make Penalty Decisions. Front. Sports Act. Living 2:595703. doi: 10.3389/fspor.2020.595703

INTRODUCTION

Soccer referees move freely on the pitch to be appropriately positioned when enforcing the Laws of the Game (International Football Association Board IFAB, 2019). By analyzing 31 matches in the EURO 2000 finals, Helsen and Bultynck (2004) found that soccer referees typically made 137 observable decisions out of approximately 200 observable and non-observable decisions during a match. At elite levels, it is expected that referees should be in control of the game, make impartial decisions, and perform without making erroneous decisions (Samuel, 2015; Samuel et al., 2018, 2020). However, referees can and do make errors (Mascarenhas et al., 2009) that may influence the match result. Indeed, the error percentage of top-class referees was found to be approximately 14% (Mallo et al., 2012). To increase the fairness of soccer matches, it should be interesting to investigate factors that make correct decisions by referees more likely, and erroneous decisions less likely.

As the probability to score on a penalty kick in soccer is about 80% (Bar-Eli et al., 2007), potential penalty decisions are arguably some of the most important decisions made by soccer referees. The importance of penalty kicks is underpinned by the fact that the average number of goals in professional soccer is about 2.5–2.7 (Bar-Eli et al., 2007; Premier League, 2019); therefore, a decision regarding whether to award a penalty or not has a high chance of influencing the outcome of the game. One may argue that such decisions should be easy to make because the rules and guidelines regarding penalties are quite clear: "a penalty kick is awarded if a player commits a direct free kick offense inside their penalty area or off the field as part of play as outlined in Laws 12 and 13" (International Football Association Board IFAB, 2019, p. 123). Further, soccer referees must document their knowledge of the regulations and that they meet the physical requirements set for speed and endurance (UEFA, 2018).

Consequently, managing the physical requirements is a prerequisite for a referee to be adequately correct positioned in the different situations that they must consider to be able to make correct decisions (Helsen and Bultynck, 2004; Mascarenhas et al., 2006; Slack et al., 2013; Weston, 2015; Aragãoe Pina et al., 2019; Joo and Jee, 2019). In terms of distance, the Football Association UK (FA) standard says that a referee should never be more than 20 m away from the playing situation (Football Association (FA) UK, 2017) because a greater distance will increase the risk of missing important information that should be used as the basis for making a correct decision. Indeed, when investigating stress in soccer referees, Voight (2009) found that referees experience that both being misplaced when making important decisions, and making erroneous decisions in potentially match-decisive situations (e.g., penalty situations) are some of their key sources of stress. While Plessner and Betsch (2001) have argued that referees base their decisions primarily on intuition Samuel et al. (2020) have recently suggested a sequential decision-making model for understanding referees' decisions. They argue that referees' decision-making sequence consists of a series of decisions (e.g., where to run, what to anticipate, what to call), beginning with visual and attention selection and ending with evaluation of potential actions. Further, the model highlights that the decision-making sequence is influenced by multiple factors (e.g., experience, knowledge of law criteria, and referee's mental state), and that referees must encode relevant environmental cues to make informed decisions. Indeed, an analysis of various experts' opinions has highlighted the complex role of soccer referees, as Aragãoe Pina et al. (2019) found football referee excellence to be shaped by individual preparation, game preparation and game management. However, many situations will have both cues that indicate that the referee should award a penalty (i.e., a harsh tackle) and cues that indicate that the referee should not award a penalty (i.e., the tackler hitting the ball first). Williams et al. (1999) claim that skilled referees should know how to keep their attention on numerous stimuli and be able to differentiate between essential and less-important cues. However, when referees make important decisions with limited time, under pressure, and often with limited relevant input, it can be difficult to evaluate these cues appropriately (Wolfson and Neave, 2007; Plessner et al., 2009). Referees should therefore aim to be positioned in a way that allows them to obtain relevant cues for making correct decisions.

Hence, applicable physical fitness is required for referees to be able to keep up with the play and get an unobstructed view of potential foul play (Riiser et al., 2019). Joo and Jee (2019) highlighted in their study of Korean elite referees that both referees' physical fitness and positioning skills should be emphasized to reduce the number of referee errors throughout the match. However, although it is highlighted that soccer referees will benefit from being in the right place at the right time (e.g., Plessner and Betsch, 2001; Mascarenhas et al., 2002), the scientific evidences regarding the relationship between decision-making accuracy and referee positioning are somewhat inconclusive. For instance, de Oliveria et al. (2011) found in their study of the Brazilian soccer referees' performances that there was no association between correct calls and the referee's distance from the play. In contrast, Mallo et al. (2012) found that the distance from the referee to the event itself affected the quality of the referee's decision, with excessive distance increasing the danger of missing out on important information, while being too close to the incident increased the risk of losing track of the situation. Specifically, Mallo et al. (2012) demonstrated that an appropriate distance (11–15 m) for referees in the central zone of the playing field gave the lowest error rate in referees' decision-making, whereas the risk of incurring errors increased when referees were more distant from foul play situations. Further, Hossner et al. (2019) analyzed both distance and angle of the match referees' position relative to foul-play match infringements in all 64 matches of the 2014 FIFA World Cup. They found that referees error rates were highest when the distance to the incident were between 10 and 15 m for whistle errors and 0–5 m for non-whistle errors. IFAB (International Football Association Board IFAB, 2019) generally recommend that viewing angle of 90 degrees to the situation (i.e., from the referee's perspective, a player is attacked by the offender perpendicularly from either the left or the right side) is appropriate to gain optimal insight. However, by calculating referees' viewing angles (between 0 and 180°), Hossner et al. (2019) identified no significant effect for viewing angles on decision-making accuracy. Nevertheless, as noted by Samuel et al. (2020), when officiating soccer, the referees must use their expertise to be in the position that allows them to make a correct decision (International Football Association Board IFAB, 2019). For instance, what can be appropriate angle for detecting a tackle may be different than for detecting a handball. As an appropriate angle to the situation may vary depending on situational factors (e.g., type of incident), a more qualitative approach to analyzing referees positioning is warranted.

In summary, referees positioning skills appear to be relevant when aiming to increase accurate decision-making (Mallo et al., 2012; Football Association (FA) UK, 2017), although the empirical evidence is somewhat inconsistent (see, Hossner et al., 2019). As some of the most crucial decisions in soccer concern whether to award a penalty because of the high probability of scoring on a penalty kick (≈ 80%; Bar-Eli et al., 2007), it can be considered interesting to investigate factors that may increase the likelihood of making a correct decision in potential penalty situations in soccer (i.e., distance, angle, and insight). Thus, the aim of this study was to conduct a preliminary investigation of elite referees' positioning in the field when making a correct or an erroneous decision in penalty situations.

METHODS

Participants

The present study used an expert panel (EP) consisting of two referees with relevant academic background to analyze referees positioning when making correct or erroneous decisions regarding penalties. The two referees in the EP were licensed and active referees and had experience at professional level in Norway. In addition, both referees had relevant academic background (i.e., degree in sport science), with one referee holding a relevant PhD. After being informed about the purpose of the study and agreeing to participate, the EP were asked to qualitatively analyze referees positioning in selected video clips by using recommended technical refereeing criteria and guidelines (i.e., the referee's distance from, angle to, and insight into the penalty situations) set by IFAB (International Football Association Board IFAB, 2019).

Procedures

The Norwegian Social Science Data Services (NSD) provided ethical approval of the study, and the procedures were in accordance with the ethical standards of the relevant University. The EP was asked to qualitatively analyze the match referees' positions in 42 potential penalty situations. Selection of situations were based on previous research, where Erikstad and Johansen (2020) used four elite Norwegian r eferees to assess all potential penalty situations from one season of the Norwegian premier league (NPL). Specifically, Erikstad and Johansen (2020) identified 98 potential penalty situations by examining two independent objective match reports from all matches from one NPL season ($N = 240$ matches). Video clips from the potential penalty situations were collected and edited using the Camtasia Studio (Tech Smith) software to present the situations in accordance with the video assistant referee system (VAR; FIFA, 2018). Like Plessner and Betsch (2001), the clips were stopped before it was possible to identify the decision made by the match referee. Further, they muted the sound, the time and result were hidden, and the situations were shown from different angles in both fast and slow motion, as well as with zoom. After viewing a situation, the four referees were told to put a mark in the box(es) that matched their judgment of the situation (e.g., no foul, free kick to defensive team, penalty, yellow card, red card). While Erikstad and Johansen (2020) used the 98 potential penalty situations to determine potential referee biases, the present study re-analyzed video clips of the situations in which all four referees in Erikstad and Johansen's study considered that the match referee made a correct decision ($N = 28$) or an erroneous decision ($N = 14$). The 14 situations where the match referee made an erroneous decision are previous studied (see Johansen and Erikstad, 2018) but are included in present paper for comparison of the analysis of the situations where the match referee made a correct decision.

Analysis

A video-based analysis of 42 selected penalty situations was conducted with the background of national and international officiating technical variables related to the referee's placement and movement (Football Association (FA) UK, 2017; UEFA, 2018). Based on findings from previous studies (e.g., Mallo et al., 2012), the distance between the referee and the penalty situation was estimated from the referee's position in relation to various markings on the field and further thematically categorized as short (<10 m), average (10–20 m), or long (>20 m). Further, the referees' guidelines (International Football Association Board IFAB, 2019) highlight that referees must use their expertise to be in the position that allows them to make a correct decision. The EP further qualitatively analyzed the referees' position in terms of angle and insight into the various situations. While insight was defined as whether one or more players blocked the referee's view, the expert panel determined the quality of the angle into the situation based on their experience and expert knowledge, as appropriate angle to the situation may vary depending on situational factors (e.g., type of incident). Based on their evaluation, good angle was obtained when the referee were in a position where they clearly could see a gap between the involved players' bodies and the relevant actions, allowing the referee to observe potential contact between the defender and attacker. Although good angle depends on situational factors, the EP noted that a good angle most often is obtained when the viewing angle is between 45 and 135° to the infringement (see Hüttermann et al., 2017; Hossner et al., 2019 for more information about how viewing angles can be calculated). Poor angle represented situations where details of the situation was hidden due to no gap between players and the related action. Poor angle typically occurred when the attacker was between the defender and referee or the defender was between the attacker and referee (i.e., close to either 0 or 180°). Average angle represented situations where details of the situations were only partly visible due to small gap between involved players' bodies. The results were thematically divided into the categories good, average, or poor. While the EP were shown the situations from three or more angles and in normal and slow-motion, they were like Erikstad and Johansen (2020) told to skip that situation if they felt the video did not expose the situation sufficiently. However, the EP reported that the quality of the clips was adequately presented for all 42 situations.

The organization and categorization in the analysis of the various video clips was carried out in accordance with general guidelines for analyzing qualitative research inspired by a thematic analysis (Braun and Clarke, 2006). An overall aim in thematic analysis is that the themes operationalized in the different categories are strongly linked to the data (Braun et al., 2016), which were the various video clips in this case. The analysis of the video clips was performed by two independent persons, as recommended for increasing validity in studies using subjective analysis (Thomas et al., 2015). Both were authorized soccer referees with professional experience in Norway, and one had relevant research experience. The correspondence between the two experts in the thematic analysis of the different situations was 98%, which is very high (Pearce et al., 2010). The one situation containing disagreement were subsequently discussed by the EP, and a mutual analysis was produced.

Due to the limited number of situations, results are descriptive comparisons of the EP's evaluations of distance, angle, and insight in the situations where the actual match referee made a correct and erroneous decision.

TABLE 1 | Characteristics of distance, angle, and insight in penalty situations where the match referee had ruled correctly.

Video clip	Distance	Angle	Insight	Decision	Infringement
4	Short	Good	Good	Penalty	Leg tripping
8	Short	Good	Good	Penalty	Leg tripping
13	Short	Good	Good	Penalty	Leg tripping
15	Average	Good	Average	Penalty	Leg tripping
17	Short	Good	Good	Penalty	Leg tripping
24	Average	Poor	Average	Penalty	Leg tripping
25	Average	Good	Average	Penalty	Leg tripping
28	Average	Good	Good	Penalty	Leg tripping
29	Average	Good	Average	Penalty	Leg tripping
30	Average	Poor	Average	Penalty	Leg tripping
31	Short	Good	Good	Penalty	Leg tripping
34	Long	Poor	Poor	Penalty	Leg tripping
35	Average	Good	Good	Penalty	Holding
36	Long	Poor	Poor	Penalty	Leg tripping
1	Average	Good	Good	No penalty	Pushing
2	Average	Good	Good	No penalty	Pushing
11	Short	Good	Good	No penalty	Leg tripping
15	Short	Good	Good	No penalty	Handball
16	Short	Good	Good	No penalty	Leg tripping
18	Long	Good	Good	No penalty	Leg tripping
21	Short	Good	Good	No penalty	Leg tripping
22	Long	Good	Average	No penalty	Leg tripping
23	Average	Good	Good	No penalty	Pushing
26	Average	Average	Average	No penalty	Holding
27	Short	Average	Poor	No penalty	Leg tripping
33	Average	Good	Good	No penalty	Pushing
38	Average	Good	Good	No penalty	Pushing
42	Average	Good	Good	No penalty	Handball

RESULTS

Overall, the EP evaluated 42 situations, of which the actual match referee made a correct decision in 28 situations (67%) and the remaining 14 situations were incorrect (33%). Regarding distance, the results revealed that the referee was <10 meters away in 12 situations (29%), 10–20 m away in 22 situations (52%), and >20 meters away in eight situations (19%). For angle, the EP considered that the match referees had a good angle in 25 situations (60%), average angle in seven situations (17%), and poor angle in 10 situations (24%). Finally, regarding insight, the EP considered the match referee to have a good insight in 21 situations (50%), average insight in 10 situations (24%), and poor insight in 11 situations (26%). Of the 42 situations, 29 (69%) were related to tripping infringements, and the remaining situations were related to pushing (eight situations; 19%), handball (three situations; 7%) and shirt holding (two situations; 5%) infringements. Distance, angle, and insight in relation to type of infringements and correct and erroneous decisions are presented below (see **Tables 1**, **2**).

Regarding distance, the match referee was positioned within 10 m from the incident in 12 situations, of which 10 (83%) were evaluated correctly (see **Table 1**). For distances between 10 and 20 meters ($N = 22$), referees made a correct decision in 14 situations (64%) and incorrect in eight situations (36%). When the match referee was positioned over 20 meters away from the situations ($N = 8$), four of eight situations (50%) were correctly refereed. Regarding angle, the referee made a correct decision in 22 of the 25 situations (88%) where the EP evaluated the positioning as good. When the match referee had an average angle to the situation ($N = 7$), the match referee was correct in two situations (29%).When having a poor angle ($N = 10$), referees made a correct decision in four situations (40%). Finally, regarding insight, the match referees made a correct decision in 18 of the 21 situations (86%) when the EP evaluated that their position allowed them to have good insight into the situation. For average positioning regarding insight, the match referee made a correct decision in seven of 10 situations (70%). When being poorly positioned in terms of insight, the match referee made a correct decision in three of 11 situations (27%). The positioning that were associated with the highest rate of correct decisions were thereby distances of under 10 meters (83% correct decisions), good angle (88%), and good insight (86%). Furthermore, no erroneous decisions were made in the nine situations where referees were under 10 m away from the situation, and where the angle and insight were good. In contrast, referees made an erroneous decision in nine of the 14 situations (64%) where either the angle or the insight were poor (See **Table 2**).

DISCUSSION

The present study was a preliminary investigation of referees' positioning in potential penalty situations where they made a

TABLE 2 | Characteristics of distance, angle, and insight in penalty situations where the match referee had ruled incorrectly.

Video clip	Distance	Angle	Insight	Decision	Infringement
3	Average	Poor	Average	No penalty	Pushing
6	Average	Average	Average	No penalty	Pushing
7	Short	Average	Average	No penalty	Leg tripping
9	Long	Poor	Poor	No penalty	Leg tripping
10	Long	Poor	Poor	No penalty	Leg tripping
12	Average	Good	Good	No penalty	Leg tripping
14	Long	Poor	Poor	No penalty	Leg tripping
32	Short	Good	Poor	No Penalty	Leg tripping
37	Average	Average	Poor	No penalty	Leg tripping
39	Long	Poor	Poor	No penalty	Leg tripping
41	Average	Average	Good	No penalty	Leg tripping
5	Average	Poor	Poor	Penalty	Leg tripping
20	Average	Good	Good	Penalty	Handball
40	Average	Average	Poor	Penalty	Pushing

correct decision and where they made an erroneous decision. Based on a qualitative analysis of two expert referees, the results revealed that the highest rate of correct decisions by the match referees occurred when the distance to the incident were under 10 m (83% correct decisions), and when the EP determined that the match referee had good angle (88%) and good insight (86%) into the situation. Erroneous decisions were more likely to occur with extended distance to the situation (36% incorrect with distances 10–20 m, and 50% with distances exceeding 20 meters). The referee made an erroneous decision in nine of the 14 situations (64%) where the EP considered that either the angle or the insight were poor. The results may indicate that appropriate positioning (i.e., short distance, good angle and good insight to the situation) allows referees to make decisions based on solid cues, and that inappropriate positioning may lead to limited visual input and increased risk of erroneous decisions.

The findings of current study indicated that good quality of referees' positioning appear to increase the likelihood of making a correct decision in penalty situations. The placement of the referee at an optimal angle and with good insight into the match play appeared to be relevant for a correct decision on whether to award or not award a penalty. These findings are in correspondence with the refereeing criteria and practical guidelines set IFAB (International Football Association Board IFAB, 2019) and seemed to provide the referee with views of the situations that enable them to attend to important and essential cues that indicate whether a penalty kick should be awarded or not. Further, the present findings are in line with the arguments provided by Samuel et al. (2020), which note that appropriate positioning by soccer referees will influence the likelihood of making correct decisions. Indeed, as Williams et al. (1999) suggested, skilled referees should know how to keep their attention on numerous stimuli and be able to differentiate between essential and less-important cues. The present findings may thereby indicate that short distance and good angle and insight allows referees to base their decisions on solid cues. This may be particularly important to detect tripping offenses, which were the most typical infringement in the present study (69% of the situations). Indeed, being close and with a clear view and angle may be crucial to identify a contact between players legs, and thus making an informed decision.

While the findings of present study have indicated that appropriate positioning of the match referee appear beneficial for making a correct decision in a potential penalty situation, appropriate positioning is not always the case. Regarding distance, the present study identified that the referee had a long distance (> 20 m) in eight situations, of which four was incorrectly refereed. Previous studies have been somewhat inconsistent in identifying appropriate distance to foul situations (e.g., de Oliveira et al., 2011; Mallo et al., 2012). However, of the misjudged 14 situations, the referee had extended distance in four situations, poor insight in eight situations and poor angle in six situations. The findings may thereby indicate that appropriate angle and insight into the situation is important for identifying a potential penalty kick. This appear particularly relevant to avoid making a non-whistling error, as 11 of 14 misjudged penalty situations where penalty kicks were not given. This is in line with previous evidence that referees error rates are lower in situations where they whistle compared to situations where they avoid whistling (Hossner et al., 2019). Furthermore, the findings of present study demonstrated that when the match referees made a correct decision by not awarding a penalty kick both the angel and insight to situation were good or average in all (14) but one situation (see **Table 1**). In the situations where a correct decision was to award a penalty kick the angle and insight were good or average in 10 of the 14 situations (see **Table 1**). The results thereby indicate that good angle and insight for the match referee may contribute to errorless decision-making in potential penalty situations. However, the findings regarding viewing angles differs from Hossner et al. (2019), which identified no significant effect for calculated viewing angles on decision-making accuracy. In combination, the findings may thereby indicate that appropriate viewing angles depends on situational factors, but generally is obtained when the referee is able to see a gap between the involved players' bodies and the relevant actions (i.e., most often between 45 and 135° to the infringement).

Potential penalty situations are potentially match-decisive decisions (Bar-Eli et al., 2007). The misplacement of the referees of present study may have forced them to make these important decisions in a limited time and it might be possible that their judgement of the importance of the cues has been affected. The referees' mispositioning in the field of play might have created uncertainty, and Johansen and Erikstad (2018) used error management theory (Haselton and Nettle, 2006), which says that decisions made under uncertainty will be biased toward the least costly error, to explain their findings. Specifically, it may be that a wrongful given penalty is perceived as a bigger error than a wrong dismissed penalty. However, it is important to emphasize that the soccer referee is human and makes quick decisions based on a subjective assessment of various play situations (Poolton et al., 2011). Even if he or she can move freely on the playing field to access the best possible distance, angle, and insight, the referee does not always have optimal insight into a situation and must decide based on their intuition and the environmental cues obtained (Plessner et al., 2009; Samuel et al., 2020). Nevertheless, as research has indicated that referees may be influenced by social pressure (Sutter and Kocher, 2004; Erikstad and Johansen, 2020), and that errors thereby not necessarily are equally distributed across teams, appropriate positioning in penalty situations as demonstrated in our study may also contribute to reduce the risk of (unintentionally) biased decisions.

Strengths and Limitations

Selection of video clips was based on previous empirical work where referees unanimously agreeing that the referee had either made a correct or an erroneous decision (see Erikstad and Johansen, 2020). Therefore, there is a basis for claiming that these situations represent situations where the judgments are most likely to have been correct. Further, the qualitative video-based analysis was performed by two persons with experience from refereeing at professional level, and relevant academic background. Weston (2015) claims that that the use of such EP is a suitable method when examining the degree of accuracy of a judicial decision. The correspondence between the two independent experts in their characteristics of distance, angle, and insight into the different situations was 98%, which is considered very high (Pearce et al., 2010).

Nevertheless, the study had limitations that must be considered. First, the study included a limited number of situations, and statistical comparisons were therefore not considered appropriate. Consequently, the present study is considered a preliminary investigation, and the findings may not be generalized. Further, while the EP could refrain from evaluate a situation if they felt the video clip did not provide sufficient information, the lack of complete insight into the video clips and the ability to assess what the referee could have seen makes the categorizations done somewhat uncertain. The categorizations should therefore be viewed as indications and future research could include eye-tracking to better understand referees' visual input and subsequent decisions. Also, while the experimental design allowed the EP to categorize situations based on their experience and expertise, the lack of objective measures (e.g., appropriate angle) is highlighted as a potential limitation. Moreover, any communication between referees and assistant referees was unknown. Messages may have been exchanged between the referees on the internal communication network and might have influenced the decisions that were made. Knowledge of the contents of the internal communication between the referees would have provided valuable information about the match referee's possible uncertainty and doubts in the various penalty situations.

CONCLUSION AND PRACTICAL IMPLICATIONS

In conclusion, the present preliminary investigation of referees positioning in potential penalty situations indicates that referees are most likely to make a correct decision when the distance to the incident are under 10 m, and when the angle and insight to the situation is good. Thus, referees' positioning skills may therefore be highlighted in both referees' training programs, and when evaluating their performances. Indeed, Samuel et al. (2020) notes that the first decision for the referee to make is where to run on the field, and often how fast he or she should run. Based on the findings of the present preliminary investigation, referees may benefit from being aware that most potential penalty incidents relates to tripping infringements, thus highlighting the need have a clear view of the players legs through short distance and clear view and angle (i.e., a viewing angle that allows the referee see a gap between the involved players' bodies and the relevant actions). The referee's optimal placement and subsequent correct decision-making in various penalty situations uncovered in this study also indicates the importance of referees being physically and mentally prepared. Specifically, as the highest rate of correct decisions was achieved when being placed < 10 m from the incident, the present study adds to the literature highlighting the importance of referees physical capabilities (Joo and Jee, 2019; Riiser et al., 2019). However, physical fitness should be combined with referees positioning skills, as the present findings indicates that good angle and insight is beneficial to make a correct decision. Referees may therefore benefit from including positioning skills in their training, for instance through observations and video analysis. It would also be interesting to expose the match referees to these situations to investigate what they perceived and thought at the time, and how they consider and assess these situations in retrospect. Such knowledge may lead to the use of individual video clips and to personalize decision-making training of referees. Especially fruitful may such training be if the video game clips, as suggested by Raab et al. (2020), include the context of the game to show sequences of decisions including the sequentially previously made choices (e.g., the referees' preferred positioning in potential penalty situations for making a decision).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary materials, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Research Ethics Committee, Faculty of Health and Sports Sciences, University of Agder. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

BJ developed the study rationale and design, interpreted the data, drafted, and revised the article. ME contributed to the development of the study rationale and design, collected, interpretation of data, drafted, and revised the article.

ACKNOWLEDGMENTS

The authors would like to thank the referees who took part in the study for their participation.

REFERENCES

Aragãoe Pina, J., Passos, A. M., Carvalho, H., and Travis Maynard, M. (2019). To be or not to be an excellent football referee: different experts' viewpoints. *J. Sports Sci.* 37, 692–700. doi: 10.1080/02640414.2018.1522940

Bar-Eli, M., Azar, O. H., Ritov, I., Keidar-Levin, Y., and Schein, G. (2007). Action bias among elite soccer goalkeepers: the case of penalty kicks. *J. Econ. Psychol.* 28, 606–621. doi: 10.1016/j.joep.2006.12.001

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Braun, V., Clarke, V., and Weate, P. (2016). "Using thematic analysis in sport and exercise research," in *Routledge Handbook of Qualitative Research in Sport and Exercise,* eds B. Smith, and A. C. Sparkes (London: Routledge), 191–205.

de Oliveira, M. C., Orbetelli, R., and Neto, T. L. D. B. (2011). Call accuracy and distance from the play: a study with Brazilian soccer referees. *Int. J. Exerc. Sci.* 4, 30–38.

Erikstad, M. K., and Johansen, B. T. (2020). Referee bias in professional football: favoritism toward successful teams in potential penalty situations. *Front. Sport. Active Living* 2:19. doi: 10.3389/fspor.2020.00019

FIFA (2018). *VAR at the 2018 FIFA World Cup.* Available online at: https://football-technology.fifa.com/en/innovations/var-at-the-world-cup/ (accessed December 17, 2020).

Football Association (FA) UK (2017). *Laws of the Game & FA Rules*. Available online at: http://www.thefa.com/soccer-rules-governance/lawsandrules.

Haselton, M. G., and Nettle, D. (2006). The paranoid optimist: an integrative evolutionary model of cognitive biases. *Pers. Soc. Psychol. Rev.* 10, 47–66. doi: 10.1207/s15327957pspr1001_3

Helsen, W., and Bultynck, J.-B. (2004). Physical and perceptual-cognitive demands of top-class refereeing in association soccer. *J Sports Sci.* 22, 179–189. doi: 10.1080/02640410310001641502

Hossner, E.-J., Schnyder, U., Schmid, J., and Kredel, R. (2019). The role of viewing distance and viewing angle on referees' decision-making performance during the FIFA world cup 2014. *J. Sports Sci. 37*, 1481–1489. doi: 10.1080/02640414.2019.1570898

Hüttermann, S., Noël, B., and Memmert, D. (2017). Evaluating erroneous offside calls in soccer. *PLoS ONE* 12:e0174358. doi: 10.1371/journal.pone.0174358

International Football Association Board IFAB. (2019). *Laws of the Game 2018/19.* Zurich: IFAB.

Johansen, B. T., and Erikstad, M. K. (2018). "Når dommeren i fotball dømmer feil … en videobasert analyse av straffesparksituasjoner i tippeligaen der norske toppdommere og toppspillere har vurdert at dommerens avgjørelse er feil," in *Trender i idrettspsykologisk forskning i Skandinavia,* eds T. Haugen, and R. Høigaard (Kap. 10, s.) (Oslo: Cappelen Damm Akademisk), 205–224. doi: 10.23865/noasp.39.ch10

Joo, C. H., and Jee, H. (2019). Activity profiles of top-class players and referees and accuracy in foul decision-making during korean National league soccer games. *J. Strength Cond. Res.* 33, 2530–2254. doi: 10.1519/JSC.0000000000003083

Mallo, J., Gonzalez Frutos, P., Juarez, D., and Navarro, E. (2012). Effect of positioning on the accuracy of decision-making of association soccer top-class referees and assistant referees during competitive matches. *J. Sports Sci.* 30, 1437–1445. doi: 10.1080/02640414.2012.711485

Mascarenhas, D. R. D., Button, C., O'Hare, D., and Dicks, M. (2009). Physical performance and decision-making in association soccer referees: a naturalistic study. *Open Sport Sci. J.* 2, 1–9. doi: 10.2174/1875399X00902010001

Mascarenhas, D. R. D., Collins, D., and Mortimer, P. (2002). 'The art of reason versus the exactness of science in elite refereeing: comments on plessner and betsch (2001)'. *J Sport Exerc. Psychol.* 24, 328–333. doi: 10.1123/jsep.24.3.328

Mascarenhas, D. R. D., O'Hare, D., and Plessner, H. (2006). The psychological and performance demands of association soccer refereeing. *Int. J. Sport Psychol.* 37, 99–120.

Pearce, C., Arnold, M., Phillips, C., and Dwan, K. (2010). Methodological considerations of digital video observation: beyond conversation analysis. *Int. J. Multiple Res. Approaches* 4, 90–99. doi: 10.5172/mra.2010.4.2.090

Plessner, H., and Betsch, T. (2001). Sequential effects in important referee decisions: the case of penalties in soccer. *J. Sport Exerc. Psychol.* 23, 254–259. doi: 10.1123/jsep.23.3.254

Plessner, H., Schweizer, G., Brand, R., and O'Hare, D. (2009). A multiple-cue learning approach as the basis for understanding and improving soccer referees' decision making. *Prog. Brain Res.* 171, 151–158. doi: 10.1016/S0079-6123(09)01313-2

Poolton, J., Siu, C. M., and Masters, R. (2011). The home team advantage gives soccer referees something to ruminate about. *Int. J. Sports Sci. Coaching* 6, 545–552. doi: 10.1260/1747-9541.6.4.545

Premier League. (2019). *Premier League Club Stats*. Available online at: https://www.premierleague.com/stats/top/clubs/goals?se=210 (accessed August 10, 2020).

Raab, M., Avugos, S., Bar-Eli, M., and MacMahon, C. (2020). The referee's challenge: a threshold process model for decision making in sport games. *Int. Rev. Sport Exerc. Psychol.* 1–20. doi: 10.1080/1750984X.2020.1783696

Riiser, A., Andersen, V., Sæterbakken, A., Ylvisaker, E., and Moe, V. F. (2019). Running performance and position is not related to decision-making accuracy in referees. *Sport. Med. Int. Open* 3, E66–E71. doi: 10.1055/a-0958-8608

Samuel, R. D. (2015). A psychological preparation framework for elite soccer referees: a practitioner's perspective. *J. Sport Psychol. Action* 6, 170–187. doi: 10.1080/21520704.2015.1065938

Samuel, R. D., Englert, C., Zhang, Q., and Basevitch, I. (2018). Hi ref, are you in control? Self-control, ego-depletion, and performance in soccer referees. *Psychol. Sport Exerc.* 38, 167–175. doi: 10.1016/j.psychsport.2018.06.009

Samuel, R. D., Tenenbaum, G., and Galily, Y. (2020). An integrated conceptual framework of decision-making in soccer refereeing. *Int. J. Sport Exerc. Psychol.* 1–23. doi: 10.1080/1612197X.2020.1766539

Slack, L. A., Maynard, I. W., Butt, J., and Olusoga, P. (2013). Factors underpinning soccer officiating excellence: perceptions of English premier league referees. *J. Appl. Sport Psychol.* 25, 298–315. doi: 10.1080/10413200.2012.726935

Sutter, M., and Kocher, M. G. (2004). Favouritism of agents - the case of referees' home bias. *J. Econ. Psychol.* 25, 461–469. doi: 10.1016/S0167-4870(03)00013-8

Thomas, J. R., Silverman, S., and Nelson, J. (2015). *Research Methods in Physical Activity, 7th Edn.* Champaign, IL: Human Kinetics.

UEFA (2018). *Refereeing.* Available online at: https://www.uefa.com/insideuefa/protecting-the-game/refereeing/ (accessed December 17, 2020).

Voight, M. (2009). Sources of stress and coping strategies of US soccer officials. *Stress Health* 25, 91–101. doi: 10.1002/smi.1231

Weston, M. (2015). Match performances of soccer referees: the role of sports science. *Movement Sport Sci.* 87, 113–117. doi: 10.3917/sm.087.0113

Williams, A. M., Davids, K., and Williams, J. G. (1999). *Visual Perception and Action in Sport.* London: E & F. N. Spon.

Wolfson, S., and Neave, N. (2007). Coping under pressure: cognitive strategies for maintaining confidence among soccer referees. *J. Sport Behav.* 30, 232–247.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINAL RESEARCH
published: 02 October 2020
doi: 10.3389/fpsyg.2020.579847

Influence of Video Speeds on Visual Behavior and Decision-Making of Amateur Assistant Referees Judging Offside Events

Vicente Luis Del Campo and Jesús Morenas Martín*

Laboratory of Motor Control and Learning, Faculty of Sport Sciences, University of Extremadura, Badajoz, Spain

The aim of the study was to assess the effects of manipulating video speeds on visual behavior and decision accuracy of 10 amateur football assistant referees (ARs) when perceived video sequences of 24 possible offside actions on a large screen. An eye tracker was used to analyze participants' visual behaviors. Signal detection analysis provided further detail of participants' decision-level accuracy. Participants were required to decide when they perceived a player to be offside during observed sequences with different video speed manipulations (*Normal speed*, *2 speed*, and *3 speed*). Results revealed that the manipulation of video speed did not attune emergent gaze patterns differently because participants displayed similar visual behaviors, regardless of speed. However, the *normal speed* resulted in a higher percentage of correct decisions than the *3 speed*. Participants tended toward non-flagging decision bias errors when judging offsides with the *3 speed* because they made more misses, than false alarms.

Keywords: gaze pattern, decision accuracy, video speed, offside, football

OPEN ACCESS

Edited by:
Yair Galily,
Interdisciplinary Center Herzliya, Israel

Reviewed by:
Heiko Lex,
University of Rostock, Germany
Asaf Blatt,
Florida State University, United States

****Correspondence:***
Vicente Luis Del Campo
viluca@unex.es
orcid.org/0000-0003-3749-3294

Specialty section:
This article was submitted to Movement Science and Sport Psychology, a section of the journal Frontiers in Psychology

Received: *03 July 2020*
Accepted: *01 September 2020*
Published: *02 October 2020*

Citation:
Luis Del Campo V and Morenas Martín J (2020) Influence of Video Speeds on Visual Behavior and Decision-Making of Amateur Assistant Referees Judging Offside Events. Front. Psychol. 11:579847. doi: 10.3389/fpsyg.2020.579847

INTRODUCTION

Sport officials must interpret and correctly enforce the rules of each sport to maintain fairness and players' safety, but also to achieve high performance in judging and making decisions about ambiguous performance situations (Bar-Eli et al., 2011). Expert judgments in sports need effective perceptual strategies to achieve improvements in the process of decision-making and anticipation (Williams and Ward, 2007). Similarly, skilled decision-making is related to the perception of relevant cues from the environment and the selection of an appropriate response (Baker et al., 2003). Previous research has demonstrated that gaze behaviors can be used as a process tracing measure to provide insights on decision-making (Hancock and Ste-Marie, 2013).

In football, Williams (2000) argued that the players displayed different perceptual strategies in 11 × 11, 2 × 2, and 3 × 3 situations because of the task nature constrained the visual strategy used. Similarly, Vaeyens et al. (2007) concluded that the number of players playing the reduced game situation influenced the visual behavior and decisions. They also showed that the experts made better decisions and a visual search strategy more adapted to the task constraints than the novels.

In judging offside situations in association football, for example, Catteeuw et al. (2009b) concluded that international assistant referees (ARs) were more accurate in detecting offside decisions and displayed longer fixations on video-projections of match play, than national-level referees. Similarly, Catteeuw et al. (2009a) found that the higher level referees made longer fixations on the offside line, and fewer flag-lag errors (i.e., to raise the flag when a player is not really offside) than national ones because they had learned to compensate for the perceptual illusion of flash-lag effect (FLE) (Gilis et al., 2009). This perceptual illusion appears when the flashed (stationary) object is perceived behind the moving target (Nijhawan, 1994). Specifically, the last defender appears to be as spatially behind the attacker receiving the ball, resulting in more cases of false alarms than misses. For a better understanding of the compensation of the FLE, some measures of verbal reports or gaze behaviors studies should be addressed in future studies (Put et al., 2013).

To facilitate acquisition of perceptual-cognitive expertise, many years of specific goal-oriented practice activities and instructions are necessary (Williams and Ericsson, 2005). In this vein, extensive experience enhances the performance of professional referees (MacMahon et al., 2007) and more hours of practice in officiating have been shown to increase the accuracy of their judgments (Pizzera and Raab, 2012). For instance, an estimation of ARs' errors in judging offside situations is 25% (Helsen et al., 2006). To reduce the level of these error rates, specific perceptual training programs using different tools (e.g., the video and/or feedback) have been investigated to enhance assistant referees' decision-making in the perception of offside events in football (Catteeuw et al., 2010a,c; Put et al., 2013).

A promising new paradigm to enhance the decision-making in sport is the use of "above real-time training" (ARTT). This research paradigm consists of using speeded-up video images as a realistic method to improve decision-making in elite athletes (Lorains et al., 2013a), providing greater task fidelity and representative design than computer animations (Catteeuw et al., 2010a; Put et al., 2016a). In football, studies by Lorains and colleagues have investigated the specific effects of video speed manipulations on decision-making of expert football players. For example, Lorains et al. (2013a) demonstrated that expert footballers achieved better accuracy in making decisions than sub-elite and novice groups in this off-field test with different video speed manipulations (1.25, 1.5, 1.75, and 2.0 times normal speed). The authors concluded that experts showed better performance in ARRT situations supported by faster processing and automaticity. Interestingly, the elite and sub-elite footballers reported 1.25 and 1.5 speeds as most "game-like." Similarly, Lorains et al. (2013b) showed that expert footballers obtained better decision accuracy in a video-based choice task using ARRT than with a normal speed during the training and transfer tests.

Similarly, Farahani et al. (2017) found that training based on ARRT improved the accuracy and response times in elite footballers' decision-making. However, these effects were limited in time because they did not last more than 2 weeks after the end of the training period. Lorains et al. (2014) also found some effects of the ARTT on visual patterns of expert football players. Specifically, the group performing ARTT made longer visual fixations on the best option to take during performance, than the normal speed and control groups after the retention tests of an intervention, where video speed was manipulated.

In football refereeing, Put et al. (2016a) reported that normal and faster speeds enhanced decisions of international assistant referees in judging offside situations. These authors concluded that training interventions for expert officials should decrease the video speed to improve response accuracy rather than increasing or varying the video speed manipulations. More recently, Spitz et al. (2018) concluded that elite referees judged ambiguous foul-play situations more severely in slow-motion replays than in real-time.

In general, these previous findings seem to reveal that the ARTT could be an effective strategy to train decision-making in football refereeing, specifically for referees at higher skilled levels. Regardless, little evidence exists about the impact of video speed manipulations in enhancing the performance of novice athletes and sport officials. In an exception, Lorains and MacMahon (2009) addressed the impact of two video speed manipulations (normal and 1.5 speed) on decision-making of footballers varying in skill level (elite, sub-elite, and novice participants). Results revealed that the high-skill group augmented differences in performance, compared to low-skill groups, as video speed was increased. They proposed that the skilled athletes seemed to need less time to process the information for decision-making because of their higher expertise levels. In this vein, Gilis et al. (2009) found that non-expert ARs achieved a lower level of performance judging offside situations when computer animations were observed at a faster speed compared to a slower speed.

However, the expert ARs learned to adopt more conservative response criteria when judged offside sequences (i.e., "not raise the flag in case of doubt"; see Put et al., 2013). Interestingly, this biased response has not been observed in studies of less experienced national ARs. For example, Luis et al. (2018) concluded that amateur ARs, but not football players, compensated for the FLE due to their embodied specific refereeing experiences. This finding could have practical implications for testing and training of football officials at different levels of experience, highlighting the need for differentiated training programs for ARs of different skill levels (Put et al., 2016a). Consequently, the research issue examined in this study concerned whether the manipulation (i.e., increasing) of video speed would reveal a compensation for the FLE in less experienced ARs when judging filmed offside events.

There is also no evidence about the contribution of manipulating video speeds in offside decisions while assessing visual behaviors of amateur ARs. Therefore, the overarching objective of this study was to address whether fast video speed manipulations (two times normal speed and three times normal speed), compared to a normal video speed condition, would influence the visual behaviors of amateur ARs and their decision accuracy in judging offside situations, perceived from a specific AR perspective. We decided to investigate increasing speed conditions due to the lack of studies testing their effects on performance of referees of lower-skill levels. Specifically, we selected the three times normal speed for the first time in studies of video speed manipulation to address whether this high video speed condition would be accompanied by decreases in amateur ARs' decision-making performance, compared to other slower video speeds manipulated in this study.

Based on previous research, we hypothesized that perception of offside events at a normal speed would increase the efficiency of visual behaviors of these amateur ARs (e.g., performing a longer fixation on the last defender during the perception of the offsides; see Catteeuw et al., 2009a), compared to the other increasing speed manipulations. Additionally, the perception of real-time condition might lead to the emergence of better decisions in judging offside events, compared to increasing video speed manipulations because the amateur ARs might benefit from more time to process relevant stimuli from this complex situation in football (Gilis et al., 2009; Lorains and MacMahon, 2009). In accordance with this assumption, the normal video speed was expected to lead to more correct decisions and fewer incorrect decisions compared to the increasing speed manipulations, allowing compensation for the FLE.

MATERIALS AND METHODS

Participants

A total of 10 male assistant referees from the Spanish Football Association took part in the study (M_{age} = 28.7; *SD* = 5.9). All participants had accumulated more than 10 years performing as ARs at the Third National Football League and/or the First Regional League. These leagues are the fourth and fifth levels of competition in the Spanish male football, respectively. Therefore, the ARs were defined as amateurs because although they had experience in a national league, and only officiated at regional-level competitions without professional experience (Swann et al., 2015). They neither reported no vision impairments, nor prior participation in talent-development programs for football officials to improve their decision-making skills in refereeing.

Ethics

Participants voluntarily took part in the study and written informed consent, to a procedure that conformed to the Declaration of Helsinki, was obtained from the individuals for the publication of any potentially identifiable images or data included in this article. The study involving human participants were reviewed and approved by The Bioethics and Biosecurity Committee of Extremadura University (approval number 33/2018). Participants received general information about the research contexts, but were naïve to the specific objectives and hypotheses.

Apparatus

Visual Behavior Assessment

An Applied Sciences Laboratories Eye Tracking ASL SE5000 recorded the visual fixations made by the participants as they observed the video speed manipulations. This device is a head-mounted, monocular eye-tracking system using corneal reflection to measure eye-line-of-gaze with respect to the field of view with an accuracy and precision of ±0.5 visual angle. The assistant referees' gaze data were stored on a digital recording device (Panasonic NV-HS1000ECP).

Video Test

A total of 240 offside events were recorded from games during several training sessions of experienced football players from the Third Division of the Spanish National Football League. These events included judgments to be made whether a player had moved offside or not during the games performed in the training sessions.

The offside judgments were recorded with a digital camera (Sony DCR-SR30), and the location on the field was chosen to simulate the viewing perspective of an assistant referee (i.e., in line with the last defender, in accordance with the Rule 11 of the International Football Federation, FIFA). Specifically, the camera was placed 25 m from the goal line, 1.20 m beyond the sideline, and 1.70 m above ground level. The Kinovea software (version 8.27) was used to edit the play sequence.

The design used for the video test was based in previous studies to ensure a realistic offside situation for analysis of ARs' perceptual and decision-making behaviors. To exemplify: (i) the interactions between the attacker receiving the ball and the last defender took place in front of the camera to simulate the correct position of the AR during the perception of these events. This methodological decision eliminated possible optical error effects as the camera angle for viewing offside events occurred at the exact place that the assistant referee should be located to observe play on the field (i.e., to eliminate an incorrect angle of view; see Catteeuw et al., 2010b); (ii) the offsides involved interactions of a small sub-group of attackers and defenders (e.g., four vs. four players visible at all times; see Catteeuw et al., 2009a); (iii) the projected game sequence contained trials with variations in viewing features, including not only small and big viewing angles but also near-medium-far distance values (Luis et al., 2015); and (iv) the sequences included distinct levels of trial difficulty, according to the spatial position of the attacker receiving the ball and the last defender (Put et al., 2014).

Variables

The independent variable was the video speed manipulation (Level 1: real-time speed or *normal speed*, Level 2: two times normal speed or *2 speed*, and Level 3: three times normal speed or *3 speed*).

The dependent variables for visual behavior were the mean number and duration of visual fixations made by the participants to different locations of the filmed sequence of play. The visual locations of interest were the ball carrier of the attacking team and last defender who defined the offside line, and fixations made on areas of no relevant interest (the attacker receiving the ball, the ball, defensive line, and offensive line). These specific locations during the perception of offside events have been previously used in other studies (e.g., see Catteeuw et al., 2010c; Luis et al., 2018).

A visual fixation was coded when the gaze remained within one degree of visual angle of a location for a minimum duration of at least 100 ms (Piras and Vickers, 2011). Since the video speed manipulation elicited speed differing in temporal durations, the number and durations of fixations in each trial were normalized relative to the mean duration of all trials presenting offside events. The percentage viewing time for these offside locations was also calculated for dividing the time of visual fixation on each visual location by the total fixation time (of all visual locations) in each trial.

The dependent variable for decision-making performance was the response accuracy. Participants made correct decisions judging an offside action if they raised a flag when they perceived an offside situation to have occurred (i.e., termed a *hit*) or when they did not raise the flag if an attacking player was not offside (i.e., termed a *correct rejection*). In contrast, participants made incorrect decisions when they raised the flag when a player was not offside (i.e., termed a *false alarm*) or they did not raise the flag when offside occurred (i.e., termed a *miss*).

Procedure

A selection of high quality sequences (n = 24), containing offside judgment decisions was video-projected onto a large screen (5 × 3 m; Hitachi CP-S310W), from the 240 initial offside events recorded during training sessions of football players. The ratio of onside and offside situations in the randomized video sequences displayed was exactly 50%. The same rate of onside vs. offside decisions was used to minimize the potential influence of any pre-conceived ideas that the ARs might have had about the likelihood of an offside taking place.

Each AR observed the same sequence of offside events at each of the three video speed manipulations. The speeds selected (normal speed and two times normal speed) were based on previous studies (Lorains et al., 2013a), and the three times normal speed was novelty used to test whether this high video speed manipulation would hinder or not AR's decision-making performance, compared to the other speed conditions. Their order of appearance in the sequence was randomized being the same for all participants.

The procedure used in this study was identical to that undertaken by Luis et al. (2018) investigating potential differences in visual search activity of ARs and players in football. For example, the distance of the ARs to the video-projection screen was 4 m, and the image size of the football players used in the sequences was calculated to provide a realistic view of the offside judgments. Participants were required to raise a flag and press a laser pointer only in those trials that they judged to include a valid offside decision while they observed the sequence wearing the eye tracker. The laser pointer was visible on the recorded film to analyze decision accuracy. This procedure ensured that decision-making information, together with the gaze location data, was recorded to a digital device for the total number of trials. These recordings allowed a further analysis of ARs' visual behavior and decision-making performance (see **Figure 1**).

The mean duration of the video sequences of offside events was about 7 s, lasting between 4 and 12 s depending on the video speed manipulation. For example, if a normal speed clip was 12 s in duration, then, the edited version of 3x that speed, was completed in 4 s. If it was edited to 2x the normal speed, then the film sequence was completed in 6 s. These temporal interval values were similar to those used by Spitz et al. (2018) in judgments of foul-play situations (e.g., 3.08 s in real-time and 12.32 s in slow-motion).

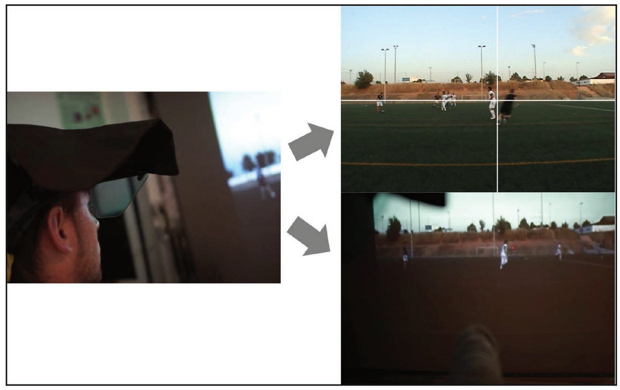

FIGURE 1 | An assistant referee observing and judging a video-projected offside event in laboratory.

The sequence contained a pause of 3 s between trials to avoid psychological fatigue in the participants give allow them to prepare to judge the next offside event. Before the observations of the offside events, ARs undertook two practice trials to familiarize them with the experimental procedures. No performance feedback was given during the test. ARs taking part in this study had never previously viewed the offside sequences observed during the test.

Statistical Analysis

Shapiro-Wilks and Levene analyses confirmed that the data of the dependent variables did not display a normal distribution, and then, nonparametric tests were used in this study. For the point-of-gaze data, descriptive statistics of means and SDs were used to explore the visual patterns of the participants with respect to number of visual fixations made, fixation duration times, and percentage viewing time spent on specific visual locations. A Kruskal-Wallis test was performed to determine differences between groups of video speed manipulation. The Rho Spearman correlation coefficient was calculated to address relations between visual behavior and decision-making in each video speed condition.

For the response accuracy, a Chi-Squared test was used to determine differences between video speeds in the percentages of *hits*, *correct rejections*, *false alarms*, and *misses*, and between correct (*hits* and *correct rejections*) and incorrect decisions (*false alarms* and *misses*). Additionally, The Kruskall-Wallis test was performed to determinate group differences in percentages of correct (*hits* vs. *correct rejections*) and incorrect (*false alarms* vs. *misses*) decisions made by the assistant referees. This last analysis of incorrect decisions provided information whether the groups of video speed compensated for the FLE. The Mann-Whitney test was used to determinate pairwise comparisons in these types of decisions.

The effect sizes (*ESs*), based on the correlation coefficient (r), were calculated to provide a better interpretation of the results. The value of the z distribution, obtained from performing the Mann-Whitney tests, was used to estimate the magnitude of *ES*. This statistic was reported for those non-parametric tests with significant differences between pairs of video speed conditions. Specifically, three categories of Cohen (1988) were used to interpret *ES* (small: $r = 0.10$ or $d = 0.20$, medium: $r = 0.24$ or $d = 0.50$, and large: $r = 0.37$ or $d = 0.80$). The CIs for *ES*s were calculated to provide a practical value of the study in real-world terms (Thompson, 2002). The Pearson's *r* was converted into Cohen's *d ES* to provide *CI* with the formula: 95% $CI = ES - 1.96se$ to $ES + 1.96se$ (Cumming, 2012). Finally, the statistical power was calculated with the G*Power software 3.1.9.2 (Faul et al., 2007) to test whether the statistically significant findings reflected true effects. A value of ≥80% power was set for analyses because it is an acceptable level to correctly reject the null hypothesis (Cohen, 1988).

A signal detection analysis (Macmillan and Creelman, 2005) was also used to analyze the response accuracy of each group in further detail. Specifically, we used d', as a sensitive index describing the assistants' ability to discriminate between "offside" and "not offside" for the three groups of video speeds above the statistical level of chance. When d' was zero, participants were not able to discriminate above the chance level between *hits* and *false alarms* in our study. If d' differed significantly from zero, the participants were able to make this distinction. Response bias or criterion *c* was also calculated to investigate the tendency of participants to make flag errors or non-flag errors. If *c* was zero, the *false alarm* and the *miss* rates were equal. When *c* was positive, the response bias indicated that the participants tended not to press the flag in our study. When *c* was negative, the response bias indicated that the participants tended to press the flag. An alpha level of <0.05 was set for all analyses. Statistical analysis was performed using the statistical package SPSS 25.0 (Statistical Package for the Social Sciences; © 2017 SPSS Inc.).

RESULTS

Gaze Behavior

Table 1 shows the mean data for fixations, fixation time, and percentage viewing time on different visual locations when participants judged the offside events at different video speeds. It is important to highlight that ARs displayed an increase of fixation time and percentage viewing time on the last defender at faster video speeds.

Specifically, the Kruskal-Wallis analyses did not reveal statistical differences between values of any visual variables as a function of video speed manipulations. No relations were also found between visual and decision-making variables in any of the video speed manipulations.

Decision-Making Performance

Table 2 shows the accuracy of perceptual judgments made by ARs during judgments of offside events, highlighting that the number of accurate decisions decreased with increasing video speeds. In the *3 speed*, the ARs achieved the highest number of *correct rejections* in correct decisions, compared to the other video speed manipulations. For incorrect decisions, the ARs showed the lowest percentage of *false alarms* and *misses* in the *normal speed*, compared to the speeded video conditions.

The Chi-Squared test showed no significant differences between video speed manipulations in the proportions of decisions made ($X^2 = 10.74$; $p = 0.09$), nor in the proportion of correct and incorrect decisions ($X^2 = 4.85$; $p = 0.08$). The Kruskal-Wallis test also showed no significant differences between video speed conditions in the percentages of decisions when participants correctly judged the offside events ($H = 5.58$; $p = 0.06$), neither when incorrectly judging the offside events ($H = 0.47$; $p = 0.78$).

The Mann-Whitney test revealed that some pairwise comparisons between the *normal speed* and *3 speed* achieved significant differences in the percentages of decisions made by participants. These differences were found in the proportions of *hits*, *correct rejections*, *false alarms*, and *misses* ($U = 2{,}186$; $p < 0.01$), and in the proportions of *hits* and *correct rejections* when correctly judged the offside events ($U = 1{,}599$; $p < 0.05$). Specifically, the *ES* was medium when the comparison included all types of decisions ($d = 0.51$; 95% *CI*: 0.19–0.83), and small for the correct decisions ($d = 0.39$; 95% *CI*: 0.04–0.74). The power of these tests was 57.12 and 56.31%, respectively.

TABLE 1 | Mean and SD (M ± DT) of visual fixations (in n°fij), fixation time (in ms), and percentage viewing time (in % respect to the 100%) for the assistant referees during the three video speed manipulations.

	Normal speed	2 speed	3 speed
Number of visual fixations			
Ball carrier	0.23 ± 0.37	0.30 ± 0.66	0.30 ± 0.70
Last defender	1.18 ± 0.95	1.50 ± 1.29	2 ± 2.06
Not on areas of interest[a]	3.71 ± 2.04	4.22 ± 2.16	4.60 ± 2.88
Fixation time			
Ball carrier	343.21 ± 736.23	235.27 ± 584.78	266.44 ± 797.31
Last defender	855.45 ± 1067.66	943.97 ± 1145.85	1102.22 ± 1363.43
Not on areas of interest[a]	2610.96 ± 1598.52	2415.47 ± 1274.62	2320.76 ± 1528.06
Percentage viewing time			
Ball carrier	8.33 ± 17.34	6.45 ± 14.58	6.16 ± 16.20
Last defender	22.94 ± 22.40	24.87 ± 24.61	28.86 ± 30.15
Not on areas of interest[a]	68.72 ± 24.30	68.66 ± 25.41	64.97 ± 30.90

[a]*Not on areas of interest: attacker receiving the ball, ball, defensive line, offensive line, and gap between offensive and defensive line.*

TABLE 2 | Type of decisions (in percentage respect to the 100%) made by assistant referees when judged offside events in football with three video speed manipulations.

	Normal speed (%)	2 speed (%)	3 speed (%)
Hits	45.6	39.6	23.7
Correct rejection	43	42.3	51.3
False alarm	6.3	9	10.5
Misses	5.1	9	14.5

The ARs discriminated between "offside" and "no offside" above chance level because the sensitivity index (d') was significantly different from zero for all video speed manipulations, being 2.41 for *normal speed*, 1.83 for *2 speed*, and 1.26 for *3 speed*. The ARs showed no decision bias toward making *false alarms* or *misses* because the criterion *c* showed no difference from zero, being −0.03 for *normal speed*, and 0.01 for *2 speed*. However, they reported a slight bias in response accuracy toward making non-flagging errors when they perceived offsides with the *3 speed* because the criterion *c* achieved a positive value of 0.23.

DISCUSSION

In this study, we investigated the impact of different video speeds on the visual behavior and decision accuracy of amateur ARs while judged a video sequence of offside events in football. More specifically, we examined whether the manipulation of the video speeds in an off-field laboratory test would lead to changes in visual behavior and decision-making of amateur assistant referees.

Interestingly, the video speed manipulation had no impact on ARs' gaze patterns because they did not modify their visual behavior despite the task constraints of the research in the form of the video speed manipulations. Consistent with these data, the hypothesis is rejected. We expected to find more gaze behavior efficiency in ARs when they observed the offside sequences at a *normal speed* compared to the speeded video conditions (i.e., to fixate the gaze longer on areas of interest: ball carrier and last defender). In contrast, participants displayed a similar gaze pattern between the different speed manipulations. Previously, Roca et al. (2013) found no differences in the visual behavior of less skilled football players when observed life-size video sequences of dynamic 11 vs. 11 situations, containing far and near task conditions, and from the central defender's viewing perspective. Similarly, Put et al. (2016b) found that an intensive, off-field training protocol enhanced ARs' response accuracy through the strategy of compensating for flag errors (i.e., participants learned to compensate for the flash-lag illusion), but without changes in visual perception.

We reasoned that the increase of the video speeds during the perception of offsides may have created a "juggling effect" for the collapse of ARs' object tracking ability (Faubert and Sidebotton, 2012). According to these authors, speeds up to the player's boundary seem to be perceived faster and more complex than they actually are. In this vein, it is possible that the speeded videos choked this ability to track multiple stimuli, masking the possibility of amateur ARs to elaborate different perceptual strategies from the observation of offsides.

Previous studies have reported higher levels of decision accuracy when assistants fixated on the offside line for longer (Catteeuw et al., 2009a,b). However, in this study, no correlations between visual behaviors and decision accuracy were found at each video speed manipulation. It is worth noting, a tendency for higher fixation times and percentage viewing times of ARs on the last defender when video speed was increased. We argue that the increase of the video speed involved the perception and judgment of offside events that were highly temporally-constrained. This temporal limitation may have generated a hard tracking of offside stimuli, driving participants to focus their gaze for longer on this visual location of interest. Therefore, the increase of the video speed conditions during the viewing of offsides ensured longer fixations on one key stimulus of these events (i.e., the last defender who defined the offside line and the offline line). In light of these findings, we suggest that the ARs' visual behaviors could not be taken to explain any differences in response accuracy.

For the decision-making performance, some specific effects in the ARs' response accuracy were found between video speed conditions. To exemplify, they performed different percentages of decisions or different proportions of correct decisions when they judged the offsides at the *normal speed* compared to the *3 speed*. The *CI*s for *ES*s did not include zero or a negative number, and then, there was a 95% likelihood that a true population effect was found between the lower and the upper scores. However, this finding did not exist in the real-world because the values achieved in the power analyses were lower than 0.80 (i.e., the likelihood that the finding reflected a null effect was higher than 20%). With this low statistical power, we are aware about our limitation to state that the video speed manipulation caused true effects in ARs' response accuracy.

In light of these statistical data, our initial hypothesis that the amateur ARs would outperform their decisions with the *normal speed* compared to the speeded video conditions is not fully accomplished. We reasoned that an underlying effect would exist in response accuracy as a result of manipulating video speed conditions because ARs showed a clear tendency to make less correct decisions and more incorrect decisions when video speeds were increased. A possible explanation to this decrease in decision-making performance during the speeded video conditions was the more distorted relative motion information between players, masking the relative position between the attacker and the last defender at the moment of the pass. We argue that the increase of the video speeds could have prevented the ARs from correcting their perception of the positions of ball carrier, attacker receiving the ball, and last defender just before the last pass; constraining a precise recall of these spatial locations at the moment of the pass (Catteeuw et al., 2010b; Put et al., 2013). Specifically, the increasing video speed inhibited a precise spatial location of the co-positioning of these three players on field at the moment to the through pass was played by the ball carrier. From this viewpoint, the fastest speed could have exceeded the capacity of amateur ARs to process the relevant stimuli of the offside events because of their low level of experience and skill in officiating (Catteeuw et al., 2009a).

The findings on perceptual sensitivity (d′) revealed that the assistants discriminated between "offsides" and "not offsides" in each video speed manipulation, even in the *3 speed*, because their sensitive index (d′) differed significantly from 0. Results suggest that the ARs had the perceptual sensitivity to discriminate these ambiguous situations in football officiating because they had accumulated enough visual and motor experiences for more than 10 years whistling as assistant referees in amateur categories. These previous observations and executions of the offside events could have helped assistants to save not biased decisions in judging offsides (Cañal-Bruland et al., 2010), showing similar rate of *hits* and *false alarms*.

Intriguingly, no differences were found between video speed manipulations for the number of incorrect decisions made. However, the assistants showed a decision bias toward errors in perceiving offsides in the *3 speed* because the criterion *c* showed a positive value different from zero (i.e., they tended to make more non-flag errors than flag errors). In elite assistant referees, this decision-level for the FLE was compensated eliminating a forward memory shift induced by this spatiotemporal illusion (Catteeuw et al., 2010a,c). However, in this study, the highest speed manipulation led to a conservative flagging strategy for the compensation of the referred perceptual illusion, when the usual decision tendency in these low-skill levels of refereeing is to make more *false alarms* than *misses* (Baldo et al., 2002; Catteeuw et al., 2010b). We suggest that the *3 speed* provoked perceptively a high task difficulty in judging offsides, with images presented too quickly on the screen (e.g., the position of the attacker receiving the ball relative to the last defender, at the moment of the pass). As the assistants had more difficulties to perceive exactly that information in the fastest videos, they gave the attacker the benefit of the doubt, showing a decisional behavior biased to not raise the flag.

Taken together, these findings posited that the video speed manipulation caused some underpowered effects in the decision accuracy (at a cognitive level) but not on the visual behavior (at a perceptual level) of amateur ARs when observed offside sequences. Specifically, the speeded video manipulations did not lead to further improvements in response accuracy. Conversely, the ARs decreased significantly the percentage of correct decisions in the *3 speed* compared to the *normal speed*, and reported a FLE in this *3 speed* toward the misses.

Strengths and Limitations

This research study had two main advantages. First, we investigated effects of video speed manipulations on visual behaviors and types of decisions of amateur ARs (e.g., the compensation for the FLE), using signal detection analysis methodology. Second, the design used for the video test of offside decisions, based on previous studies (Catteeuw et al., 2009a, 2010b; Put et al., 2014) and perceived from the AR perspective, created fidelity and representative design for analyses of visual patterns and decision-making processes.

The study could be improved by increasing the limited amount of cases analyzed, to avoid masking observations of statistically significant differences in response accuracy levels between video speeds, which in this study remained at the level of statistical trends in the data. For example, questions exist whether there would be differences between *normal speed* and *2 speed* or between *2 speed* and *3 speed* if more assistant referees were recruited to participate in the study. This small number of assistant referees investigated could explain the low power estimation observed in this study. According to the G*Power software 3.1.9.2 (Faul et al., 2007), a total sample of 21 ARs per group should be tested to achieve the threshold of 80% power, an alpha level of 0.05, and a large *ES* ($d' = 0.8$). Similarly, it would be interesting to address whether this tendency found of fixating longer on the offline line achieves significant levels with larger samples of participants.

Another limitation was that the experienced football players involved in the recording of the offside sequences did not receive indications about what speed execution should be performed during these specific events in football. They had freedom to perform, in a natural manner, these specific situations to accomplish with the requirements of the research team. The absence of control in this variable (e.g., to execute offsides with high, medium, and low speeds) could mask a possible effect of perceptual sensitivity in judging offsides, intrinsically linked to the speed in which the offside events were performed and recorded.

Training Perspectives

The video speed manipulation for the refereeing performance should be introduced carefully in the perceptual trainings of the referees, according to their sport skill level. To do this, the learning task designs should state clear previously what video speed thresholds are more adequate to enhance decision accuracy for expert and amateur assistant referees. This study represents another step to elucidate this question, trying to add new evidence for a better design of perceptual training programs in development of skills in assistant referees because the influence of video speed manipulations in referees' decision-making has received sparse interest in the literature (Spitz et al., 2018).

With these data, the speeded video conditions seem not to be a reasonable strategy for the improvement of decision performance in amateur assistant referees, at least when they perceived offsides with the *3 speed*. In contrast, the use of normal speed video clips for observations of offside sequences may prevent ARs from making incorrect decisions. It is also possible that low skilled ARs should also be exposed to different emotional and contextual variables of real-game scenarios (e.g., stress, anxiety, home advantage, and external crowd pressure) to learn how to maintain their decision-making skills under pressure. Following this relationship between emotions and decision-making, the off-field training programs undertaken in controlled-laboratory settings would gain fidelity and representativeness by including cognitive and emotional constraints that are embedded in football matches.

To do this, emotional intelligence training (e.g., the emotional regulation and pre-competitive routines; see Campo et al., 2015) or virtual-reality technology could be promising strategies to achieve transfer from laboratory to on-field scenarios when judged offside events. In this line, teaching strategies in offside situations could include, as in foul-play situations, a combination of immediate feedback about the correctness of decisions (Schweizer et al., 2011) together with a scheduling of decreasing video speed sequences (Put et al., 2016a) to train amateur ARs to better deal with the FLE.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The study involving human participants were reviewed and approved by The Bioethics and Biosecurity Committee of Extremadura University (approval number 33/2018). The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individuals for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

VL and JM equally contributed to the conception and design of the study. JM performed the measurements of the experiment. VL performed the experimental data and the statistical analyses. VL and JM wrote the first draft of the manuscript. Both authors contributed to manuscript revision, and approved the final version of the manuscript and agreed with the order of presentation of the authors.

FUNDING

This study was carried out thanks to the contribution of the Ministry of Economy and Infrastructure of the Board of Extremadura through the European Regional Development Fund (GR18129).

ACKNOWLEDGMENTS

The authors acknowledge Regional Football Referee Committee of Extremadura and all participants who voluntary participated in this study.

REFERENCES

Baker, J., Côté, J., and Abernethy, B. (2003). Sport-specific practice and the development of expert decision-making in team ball sports. *J. Appl. Sport Psychol.* 15, 12–25. doi: 10.1080/10413200305400

Baldo, M. V. C., Ranvaud, R. D., and Morya, E. (2002). Flag errors in soccer games: the flash-lag effect brought to real life. *Perception* 31, 1205–1210. doi: 10.1068/p3422

Bar-Eli, M., Plessner, H., and Raab, M. (2011). *Judgement, decision making and success in sport*. Oxford: Wiley-Blackwell.

Campo, M., Laborde, S., and Weckemann, S. (2015). "Emotional intelligence training: implications for performance and health" in *Advances in psychology research*. ed. A. M. Colombus (New York: Nova Publishers), 75–92.

Cañal-Bruland, R., van der Kamp, J., and van Kersteren, J. (2010). An examination of motor and perceptual contributions to the recognition of deception from others' actions. *Hum. Mov. Sci.* 29, 94–102. doi: 10.1016/j.humov.2009.10.001

Catteeuw, P., Gilis, B., Jaspers, A., Wagemans, J., and Helsen, W. (2010a). Training of perceptual-cognitive skills in offside decision making. *J. Sport Exerc. Psychol.* 32, 845–861. doi: 10.1123/jsep.32.6.845

Catteeuw, P., Gilis, B., Wagemans, J., and Helsen, W. (2010b). Offside decision making of assistant referees in the English premier league: impact of physical and perceptual-cognitive factors on match performance. *J. Sports Sci.* 28, 471–481. doi: 10.1080/02640410903518184

Catteeuw, P., Gilis, B., Wagemans, J., and Helsen, W. (2010c). Perceptual cognitive-skills in offside decision making: expertise and training effects. *J. Sport Exerc. Psychol.* 32, 828–844. doi: 10.1123/jsep.32.6.828

Catteeuw, P., Helsen, W., Gilis, B., Van Roie, E., and Wagemans, J. (2009a). Visual scan patterns and decision-making skills of expert assistant referees in offside situations. *J. Sport Exerc. Psychol.* 31, 786–797. doi: 10.1123/jsep.31.6.786

Catteeuw, P., Helsen, W., Gilis, B., and Wagemans, J. (2009b). Decision-making skills, role specificity, and deliberate practice in association football refereeing. *J. Sports Sci.* 27, 1125–1136. doi: 10.1080/02640410903079179

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences. 2nd Edn*. Hillsdale, NJ: Erlbaum.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.

Farahani, J. J., Javadi, A. H., O'Neill, B. V., and Walsh, V. (2017). Effectiveness of above real-time training on decision-making in elite football: a dose—response investigation. *Prog. Brain Res.* 234, 101–116. doi: 10.1016/bs.pbr.2017.08.007

Faubert, J., and Sidebotton, L. (2012). Perceptual-cognitive training of athletes. *J. Clin. Sport Psychol.* 6, 85–102. doi: 10.1123/jcsp.6.1.85

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/bf03193146

Gilis, B., Helsen, W., Catteeuw, P., Van Roie, E., and Wagemans, J. (2009). Interpretation and application of the offside law by expert assistant referees: perception of spatial positions in complex dynamic events on and off the field. *J. Sports Sci.* 27, 551–563. doi: 10.1080/02640410802702178

Hancock, D. J., and Ste-Marie, D. (2013). Gaze behaviors and decision-making accuracy of higher-and lower-level ice hockey referees. *Psychol. Sport Exerc.* 14, 66–71. doi: 10.1016/j.psychsport.2012.08.002

Helsen, W., Gilis, B., and Weston, M. (2006). Errors in judging "offside" in association football: test of the optical error versus the perceptual flash-lag hypothesis. *J. Sports Sci.* 24, 521–528. doi: 10.1080/02640410500298065

Lorains, M., Ball, K., and MacMahon, C. (2013a). Expertise differences in a video decision-making task: speed influences on performance. *Psychol. Sport Exerc.* 14, 293–297. doi: 10.1016/j.psychsport.2012.11.004

Lorains, M., Ball, K., and MacMahon, C. (2013b). An above real time training intervention for sport decision making. *Psychol. Sport Exerc.* 14, 670–674. doi: 10.1016/j.psychsport.2013.05.005

Lorains, M., and MacMahon, C. (2009). "Adapting the functional overreaching principle to cognitive skills: expertise differences using speed manipulations in a video-based decision-making task" in *Proceedings of the 12th World Congress of Sport Psychology (ISSP)*; June 17–29, 2009; Marrakesh, Morocco.

Lorains, M., Panchuk, D., Ball, K., and MacMahon, C. (2014). The effect of an above real time decision-making intervention on visual search behaviour. *Int. J. Sports Sci. Coach.* 9, 1383–1392. doi: 10.1260/1747-9541.9.6.1383

Luis, V., Canelo, A., Domínguez, F. J., and Morenas, J. (2018). The influence of refereeing experiences judging offside actions in football. *Psychol. Sport Exerc.* 37, 139–145. doi: 10.1016/j.psychsport.2018.03.004

Luis, V., Canelo, A., Morenas, J., Gómez-Valadés, J. M., and Gómez, J. S. (2015). Referees' visual behaviour during offside situations in football. *Int. J. Med. Sci. Phys. Activ. Sport.* 15, 325–338. doi: 10.15366/rimcafd2015.58.007

MacMahon, C., Helsen, W. F., Starkes, J. L., and Weston, M. (2007). Decision-making skills and deliberate practice in elite association soccer referees. *J. Sports Sci.* 25, 65–78. doi: 10.1080/02640410600718640

Macmillan, N. A., and Creelman, C. D. (2005). *Detection theory: A user's guide*. Mahwah, NJ: Erlbaum.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* 370, 256–257. doi: 10.1038/370256b0

Piras, A., and Vickers, J. N. (2011). The effect of fixation transitions on quiet eye duration and performance in the soccer penalty kick: instep versus inside kicks. *Cogn. Process.* 12, 245–255. doi: 10.1007/s10339-011-0406-z

Pizzera, A., and Raab, M. (2012). Perceptual judgments of sports officials are influenced by their motor and visual experience. *J. Appl. Sport Psychol.* 24, 59–72. doi: 10.1080/10413200.2011.608412

Put, K., Wagemans, J., Jaspers, A., and Helsen, W. F. (2013). Web-based training improves on-field offside decision-making performance. *Psychol. Sport Exerc.* 14, 577–585. doi: 10.1016/j.psychsport.2013.03.005

Put, K., Wagemans, J., Pizzera, A., Williams, A. M., Spitz, J., Savelsbergh, G. J. P., et al. (2016a). Faster, slower or real time? Perceptual-cognitive skills training with variable video speeds. *Psychol. Sport Exerc.* 25, 27–35. doi: 10.1016/j.psychsport.2016.03.007

Put, K., Wagemans, J., Spitz, J., Armenteros Gallardo, M., Williams, A. M., and Helsen, W. F. (2014). The use of 2D and 3D video clips in a perceptual-cognitive judgement task. *J. Sports Sci.* 32, 1688–1697. doi: 10.1080/02640414.2014.912760

Put, K., Wagemans, J., Spitz, J., Williams, A. M., and Helsen, W. F. (2016b). Using web based training to enhance perceptual-cognitive skills in complex dynamic offside events. *J. Sports Sci.* 34, 181–189. doi: 10.1080/02640414.2015.1045926

Roca, A., Ford, P. R., McRobert, A. P., and Williams, A. M. (2013). Perceptual-cognitive skills and their interaction as a function of task constraints in soccer. *J. Sport Exerc. Psychol.* 35, 144–155. doi: 10.1123/jsep.35.2.144

Schweizer, G., Plessner, H., Kahlert, D., and Brand, R. (2011). A video-based training method for improving soccer referees' intuitive decision-making skills. *J. Appl. Sport Psychol.* 23, 429–442. doi: 10.1080/10413200.2011.555346

Spitz, J., Moors, P., Wagemans, J., Williams, A. M., and Helsen, W. F. (2018). The impact of video speed on the decision-making process of sports officials. *Cogn. Res. Princ. Implic.* 3:16. doi: 10.1186/s41235-018-0105-8

Swann, C., Moran, A., and Piggott, D. (2015). Defining elite athletes: issues in the study of expert performance in sport psychology. *Psychol. Sport Exerc.* 16, 3–14. doi: 10.1016/j.psychsport.2014.07.004

Thompson, B. (2002). What future quantitative social science research could look like: confidence intervals for effect sizes. *Educ. Res.* 31, 25–32. doi: 10.3102/0013189X031003025

Vaeyens, R., Lenoir, M., Williams, A. M., Mazyn, L., and Philippartes, R. M. (2007). The effect of task constraints on visual search behavior and decision making in youth soccer players. *J. Sport Exerc. Psychol.* 29, 147–169. doi: 10.1123/jsep.29.2.147

Williams, A. M. (2000). Perceptual skill in soccer: implications for talent identification and development. *J. Sports Sci.* 18, 737–750. doi: 10.1080/02640410050120113

Williams, A. M., and Ericsson, K. A. (2005). Perceptual-cognitive expertise in sport: some considerations when applying the expert performance approach. *Hum. Mov. Sci.* 24, 283–307. doi: 10.1016/j.humov.2005.06.002

Williams, A. M., and Ward, P. (2007). "Anticipation and decision making: exploring new horizons" in *Handbook of sport psychology*. *3rd Edn*. eds. G. Tenenbaum and R. C. Eklund (Hoboken, NJ, USA: John Wiley and Sons, Inc.).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 26 May 2020
doi: 10.3389/fpsyg.2020.01082

Check for updates

"Are You Telling the Truth?" — Testing Individuals' Ability to Differentiate Between Truth and Deceit in Soccer

Chris Englert[1,2†] and Geoffrey Schweizer[3*†]*

[1] *Institute of Sports Sciences, Department of Sports Psychology, Goethe University Frankfurt, Frankfurt, Germany,* [2] *Institute of Education, Department of Educational Psychology, University of Bern, Bern, Switzerland,* [3] *Institute of Sports Sciences, Department of Sports Psychology, Heidelberg University, Heidelberg, Germany*

Edited by:
Yair Galily,
IDC, Israel

Reviewed by:
Itay Basevitch,
Anglia Ruskin University,
United Kingdom
Guy Hochman,
Interdisciplinary Center Herzliya, Israel

***Correspondence:**
Chris Englert
englert@sport.uni-frankfurt.de;
christoph.englert@edu.unibe.ch
Geoffrey Schweizer
geoffrey.schweizer@issw.uni-heidelberg.de

†*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Movement Science and Sport Psychology,
a section of the journal Frontiers in Psychology

Received: *24 February 2020*
Accepted: *28 April 2020*
Published: *26 May 2020*

Citation:
Englert C and Schweizer G (2020) "Are You Telling the Truth?" — Testing Individuals' Ability to Differentiate Between Truth and Deceit in Soccer. Front. Psychol. 11:1082.
doi: 10.3389/fpsyg.2020.01082

In the present paper, we tested the ability of individuals to judge correctly whether athletes are lying or telling the truth. For this purpose, we first generated 28 videos as stimulus material: in half of the videos, soccer players were telling the truth, while in the other half, the same soccer players were lying. Next, we tested the validity of these video clips by asking $N = 65$ individuals in a laboratory experiment (Study 1a) and $N = 52$ individuals in an online experiment (Study 1b) to rate the level of veracity of each video clip. Results suggest that participants can distinguish between true and false statements, but only for some clips and not for others, indicating that some players were better at deceiving than others. In Study 2, participants again had to make veracity estimations, but we manipulated the level of information given, as participants ($N = 145$) were randomly assigned to one of three conditions (regular video clips, mute video clips, and only the audio stream of each statement). The results revealed that participants from the mute condition were less accurate in their veracity ratings. The theoretical and practical implications of these findings are discussed.

Keywords: deception, lying, truth, referee, soccer

INTRODUCTION

Antisocial behavior in sports and exercise contexts has been documented in several studies and can be understood as intentional behavior designed to disadvantage other individuals (Kavussanu et al., 2006). Research has focused on, among other topics, the prevalence of antisocial behavior (e.g., Kavussanu, 2006), the use of illegal performance-enhancing drugs (e.g., Momaya et al., 2015), or the reasons why athletes make the decision to display antisocial behavior in the first place (e.g., Kavussanu and Roberts, 2001; Ommundsen et al., 2003). In the present paper, we focus on deception in sports, which involves "making someone believe something that is not true in order to get what you want" (Hsu, 1997), p. 167; for a review, see Güldenpenning et al. (2017). To gain an advantage, athletes are oftentimes motivated to deceive the referee, as deception might change the course of a match, for instance, when a soccer player is asking for a penalty during the last minutes of a match even though there had been no foul (e.g., Traclet et al., 2011; Sabag et al., 2018).

Being able to detect deception is not only relevant during a sporting competition but also in the criminal justice system (e.g., Akehurst et al., 1996) or in educational contexts (e.g., Marksteiner et al., 2013). In fact, most studies on lie detection have been conducted in the context of the criminal justice system, which is not surprising, as it is extremely important to classify a statement correctly in court as being true or false. In general, individuals are not highly accurate when it comes to detecting truths and lies, as a meta-analysis revealed they are only slightly better than chance level (accuracy rate of 54%; e.g., Bond and DePaulo, 2006). Overall, individuals are better at identifying a true statement correctly (accuracy rate of 61%) than at identifying a lie correctly (accuracy rate of 47%). Similar accuracy rates have been reported in the field of sport psychology, for instance, in a study by Renden et al. (2014) in which participants were asked to judge whether tackle situations in soccer matches on television were either fouls or dives. While there are plenty of correlational and qualitative publications on judgment and decision making of sports officials (for an overview see Aragão e Pina et al., 2018), there has been little experimental research on referees' ability to differentiate correctly between a true statement and an invented one (e.g., Morris and Lewis, 2010). Experimental designs would allow one to draw causal conclusions concerning which factors have a direct influence on referees' judgment and decision making (e.g., Morris and Lewis, 2010; Sabag et al., 2018).

Therefore, which factors determine whether individuals are capable of estimating accurately the truth of a statement? A meta-analysis by Aamodt and Custer (2006) did not find empirical evidence of a significant effect of gender, age, self-confidence, or certain personality traits (e.g., extraversion) on accuracy rates. Furthermore, expertise does not automatically lead to judgements that are more accurate, meaning that laypersons oftentimes do not differ significantly from experts in their accuracy rates (Aamodt and Custer, 2006; Bond and DePaulo, 2006; del Campo et al., 2018). In sports, results partially suggest that the level of expertise might have an influence on the accuracy rates of judgments (e.g., Renden et al., 2014). However, given the small number of studies on lie detection in sports, future research is needed. Theoretically, the ability to identify correctly both true and false statements hinges on two factors: first, on the presence of cues that differentiate between true and false statements (i.e., valid cues), and second, on individuals' ability to perceive these cues and to use them in a correct manner (so-called cue usage). This means that individuals must use only valid cues and neglect non-valid cues. Furthermore, they must know how specific cues relate to the probability of a statement being true or false. There are several potential factors which can influence the ability to differentiate between true and false statements [e.g., Need for Cognition (NFC); e.g., Reinhard, 2010]. However, in the present studies, we were primarily interested in (a) developing a valid experimental lie detection research paradigm in sports and (b) investigating individuals' ability to properly judge critical game situations.

In line with these considerations, Vrij et al. (2006) propose that one potential explanation for the low detection rates seems to be that individuals oftentimes hold inadequate beliefs about valid cues related to deception. For instance, in laypersons and experts, there is a common stereotypic belief that liars have a tendency to avoid eye contact and to display strong nervous body movements (e.g., Bond and DePaulo, 2006). However, DePaulo et al. (2003) did not find any empirical support for these non-verbal and para-verbal cues. Research has repeatedly demonstrated that relying on these invalid cues when making a judgment affects the accuracy rates negatively (e.g., Frank and Ekman, 2004). In general, non-verbal cues, such as the aforementioned ones, are less strongly related to deception than verbal cues (i.e., the content of the respective statement). Several studies identified the following valid verbal cues to deception: lies are not as logically structured as true statements, lies are less plausible, lies do not contain as many relevant details, and lies are more ambivalent than true statements (e.g., DePaulo et al., 2003). Therefore, it seems beneficial to focus on verbal cues instead of non-verbal or para-verbal cues to make an accurate veracity judgment (Forrest et al., 2004).

As previously mentioned, there has been little experimental research in sports-related contexts on antisocial behavior in general and on lie detection in particular. Instead, most studies on antisocial behavior collected data either by interviewing coaches regarding specific behaviors (e.g., Stuart and Ebbeck, 1995) or by asking individuals how they would behave in hypothetical situations (e.g., Kavussanu and Ntoumanis, 2003). One of the only studies that used an experimental design to investigate lie detection in sports was conducted by Morris and Lewis (2010). In their study, they first created five video clips as stimulus material in which they instructed amateur soccer players to exaggerate the effects of a tackle by an opponent in a convincing manner. They videotaped these sequences and asked participants in another study to watch these video clips and to make a judgment regarding the level of exaggeration of the tackled player depicted in the video clip. The results of this study revealed that participants were fairly accurate in estimating the level of exaggeration.

In the present paper, we were not interested in the prevalence of deception or in the reasons why athletes decide to deceive the referee; instead, we focused on the ability of individuals to detect lies. We first created the stimulus material, which consisted of 28 videos in which soccer players were either telling the truth or lying (see also Morris and Lewis, 2010). In Study 1a (laboratory) and Study 1b (online), we tested these video clips by asking participants to rate the veracity of each video (for a similar approach, see Morris and Lewis, 2010). The participants did not see the actual game situation, but only the interview with the respective player which took place after the critical incident. There were two reasons for replicating Study 1a: First, we wanted to make sure that both participants' ability to discriminate between true and false statements and potential differences between video pairs found in Study 1a reflect systematic differences instead of simply random variation. Second, we wanted to make sure that results do not depend on a laboratory setting, but can also be obtained in an online setting (for a discussion on the replication crisis, see also Klein Richard et al., 2014). In Study 2, we manipulated the type of information presented to the participants, where participants watched the original video clips (i.e., original condition), watched the original video clips without containing any auditory information (i.e., mute condition), or only listened to the audio stream without seeing the actual video clip (i.e., audio condition). In line with previous research, which has shown that non-verbal cues are less reliable than verbal cues (e.g., DePaulo et al., 2003), we assumed that participants from the mute condition would be less accurate in their veracity assessments than participants from the other two conditions. We will explain the stimulus material and the studies in more detail in the following sections. The local ethics committee approved all studies reported in this paper.

GENERATION OF THE STIMULUS MATERIAL

Fourteen male soccer players (M_{age} = 23.36, SD_{age} = 4.77) from a club from the sixth highest league in Germany (out of 11 leagues) volunteered to participate to create the stimulus material (for this procedure, see Marksteiner et al., 2015; see also Levine et al., 2011). On average, the players had played soccer for 18 years (SD = 3.72), and there were defensive, offensive, and midfield players among them. The study was conducted in single sessions on a regular soccer pitch. We obtained written informed consent from each participant before commencing the study.

Each participant played the part of a defender twice, leading to two scenarios. In both scenarios, two confederates acted as attacking players from the opposing team (for the setup, see **Figure 1**). One confederate played a long pass toward the goal line for his teammate (i.e., the second confederate). In one scenario, the defender's job was to prevent the opposing player from reaching the ball and to let the ball cross the goal line so that the defender's team would get the goal kick. In the second scenario, however, the defender was instructed to touch the ball slightly before it crossed the goal line, so that the correct decision would actually be a corner kick instead of a goal kick. In both scenarios, after the ball had crossed the line, the referee blew his whistle, requested the defender to come over to a marked position immediately where a video camera was set up, and asked him a series of seven questions (1. Who was the last player to touch the ball? 2. Are you sure? 3. Why is the other player saying something different? 4. Are you sure? 5. Again, who was the last player to touch the ball? 6. Are you sure about that? 7. Why should I believe you?). These questions were developed in cooperation with an official German B-level referee and an A-level soccer coach (both handed out by the German Football Association) to ensure the questions would be as realistic as possible. The defender was instructed to state in both scenarios that the attacker was the last player to touch the ball and that the referee should decide on a goal kick. That way, we generated two videos from each participant: one video in which he was telling the truth and one video in which he was lying (order counterbalanced). The referee was blind to the condition and did not see the actual critical game scene, as he turned away when the tackling happened. To increase participants' motivation, we offered tickets to a Bundesliga soccer match (first German soccer division) to the one player who was the most convincing. That way, we generated 28 videos in total (14 true statements, 14 lie statements). Each video lasted approximately 28 s (M = 27.5, SD = 6.27) and contained the same amount of questions asked by the same non-visible referee. Each player's upper torso, face, and legs could be seen on all video tapes and the sound quality was the same in all video clips.

FIGURE 1 | Illustration of the experimental setup for the generation of the stimulus material. The player wearing the jacket is a confederate acting as an attacking player, the player wearing the white jersey is a confederate acting as the teammate of the attacking player, and the player wearing the black jersey is the target player acting as the defender. The referee is standing on the right, observing the scene.

STUDY 1A

The aim of Study 1a was twofold: first, we wanted to investigate the question of whether participants are able to distinguish between true and false statements. Based on existing research, we expected participants to be able to distinguish between true and false statements; however, we expected a small to medium effect size at best (e.g., Bond and DePaulo, 2006). Second, we aimed to determine whether participants' ability to distinguish between true and false statements differed between video pairs. In other words, we wanted to find out whether some of the players we filmed were better liars than others were. We were interested in determining which video pairs were more ambiguous (i.e., the veracity ratings of the true and untrue statements of the respective target players do not differ significantly) and which video pairs were less ambiguous (i.e., the veracity ratings of the true and untrue statements of the respective target players differ significantly).

Methods

Participants

A total of N = 65 university students from a German university participated voluntarily in this study (26 female; M_{Age} = 24.66 years, SD_{Age} = 4.24). Eight participants had refereeing experience (M = 3.88 years, SD = 3.87). All participants delivered written informed consent before taking part in the research.

Design, Procedure, and Measures

The study was conducted in the university's laboratory and the videos and all the instructions were administered on a regular computer screen using an online survey software (Unipark). After delivering demographic information (age, sex, mother tongue, and refereeing experience), the 28 videos were displayed in a random order. The participants were informed that in each video clip a player would be asked a series of questions by a

TABLE 1 | Study 1a: Veracity ratings for the true and the false statements of each video pair.

Video pair number	True statement		False statement			
	M	*SD*	*M*	*SD*	*F*	η^2_p
1	6.08	2.53	5.89	2.61	0.19	0.00
2	5.83	2.52	4.09	2.47	26.27**	0.29
3	6.37	2.36	5.12	2.49	9.34**	0.13
4	4.26	2.32	3.89	2.10	0.87	0.01
5	8.35	1.87	6.46	2.50	22.03**	0.26
6	5.22	2.80	3.55	2.29	17.82**	0.22
7	7.09	2.19	5.62	2.73	14.97**	0.19
8	6.14	2.74	6.60	2.67	1.35	0.02
9	7.06	2.46	6.17	2.88	4.51*	0.07
10	6.26	2.60	5.65	2.70	2.18	0.03
11	3.15	2.39	4.25	2.65	10.66**	0.14
12	6.37	2.44	5.78	2.81	1.83	0.03
13	5.12	2.45	5.71	2.66	2.04	0.03
14	3.97	2.51	4.03	2.22	0.03	0.00

Note. N = 65. Each video was rated on a continuous scale ranging from 1 (not at all true) to 10 (totally true). $^*p < 0.05$. $^{**}p < 0.01$.

professional referee and that they had to rate the veracity of each video clip. All participants were wearing regular stereo headphones and the sound was played at a constant volume. Following each video clip, participants rated the truth of each statement on a continuous scale ranging from 1 (*not at all true*) to 10 (*totally true*; for this procedure, see Marksteiner et al., 2013). Finally, the participants were debriefed and thanked for their participation.

Results

Overall, participants were able to distinguish between false and true statements. The veracity ratings of false statements ($M = 5.20$, $SD = 0.78$) were lower on average than the veracity ratings of true statements ($M = 5.81$, $SD = 0.82$). A within-subject analysis of variance (ANOVA) indicates that this difference is a significant one, $F(1,64) = 28.29$, $p < 0.0001$, $\eta^2_p = 0.31$. Furthermore, results suggest that participants were indeed able to distinguish between true and false statements for some but not all video pairs.

To test for which video pairs participants were able to distinguish between true and false statements, we conducted within-subject ANOVAs for each video pair. These analyses resulted in seven ambiguous video pairs (i.e., there were no significant differences between the veracity ratings for the true and the deceptive statements) and seven less-ambiguous video pairs (i.e., the veracity ratings for the true and the deceptive statements differed significantly). The detailed results are illustrated in **Table 1**.

Discussion

The results of Study 1a suggest that, overall, participants are able to distinguish between true and false statements in our stimulus material; however, this ability differs for different video pairs. Furthermore, they revealed that out of the 14 video pairs, seven video pairs were ambiguous, meaning the veracity ratings of the true and deceptive statements of these target players did not differ significantly. On the contrary, for the other seven video pairs, participants were able to differentiate correctly between the true and deceptive statements.

STUDY 1B

The primary aim of Study 1b was to test whether the pattern found in Study 1a is robust in a novel study to ensure that both participants' ability to discriminate between true and false statements and the differences between video pairs found in Study 1a reflects a systematic difference between true and false statements and between video pairs. Furthermore, we wanted to test whether the results depend on a laboratory setting or whether they will also emerge in an online setting (for a discussion on the replication crisis, see also Klein Richard et al., 2014). For this cause, we posted an online link on various social platforms that led participants to an anonymous survey containing all videos and questionnaires (Unipark).

Methods

Participants

In total, $N = 94$ individuals clicked the online link and $n = 52$ individuals finished the study. The individuals who accessed the online survey were informed of the purpose of the study, delivered informed consent, and confirmed that they agreed to participate voluntarily. All the following analyses were conducted only with the participants who finished the study (29 female; $M_{\mathrm{Age}} = 36.54$ years, $SD_{\mathrm{Age}} = 15.76$). Three participants had refereeing experience ($M = 3.33$ years, $SD = 4.04$).

Design, Procedure, and Measures

The design was identical to the design of Study 1a, with the only difference being that Study 1b was conducted online. Participants delivered demographic information and rated the veracity of all 28 videos (Marksteiner et al., 2013). Finally, we thanked the participants for their participation and debriefed them.

Results

Main Analyses

Just as in Study 1a, overall, participants were able to distinguish between false and true statements. The veracity ratings of the false statements ($M = 5.38$, $SD = 1.21$) were lower on average than the veracity ratings of the true statements ($M = 5.82$, $SD = 1.07$). A within-subject ANOVA indicates that this difference is significant, $F(1,51) = 14.69$, $p < 0.0001$, $\eta^2_p = 0.22$.

We again ran within-subject ANOVAs for each video pair to determine which video pairs were ambiguous and which were non-ambiguous (for detailed results, see **Table 2**). In Study 1b, for seven video pairs, there were no significant differences between the veracity ratings for the true and the deceptive statements

TABLE 2 | Study 1b: Veracity ratings for the true and the false statements of each video pair.

Video pair number	True statement		False statement			
	M	*SD*	*M*	*SD*	*F*	η^2_p
1	5.33	2.80	5.35	2.62	0.01	0.00
2	5.25	2.61	4.96	2.71	0.36	0.01
3	7.02	2.44	5.94	2.49	5.40*	0.10
4	4.58	2.67	5.25	2.88	3.15	0.06
5	7.52	2.41	6.21	2.47	9.88**	0.16
6	6.04	2.61	4.87	2.69	8.74**	0.15
7	6.83	2.47	5.54	2.85	7.64**	0.13
8	5.56	2.73	5.04	2.60	1.55	0.03
9	7.33	2.10	6.60	2.80	3.59*	0.07
10	5.60	3.12	5.67	2.71	0.03	0.00
11	4.06	2.86	5.17	2.69	15.76**	0.24
12	6.17	2.61	5.65	2.66	1.39	0.03
13	5.15	2.52	5.02	2.52	0.11	0.00
14	5.08	2.50	4.08	2.47	9.27**	0.15

*Note. N = 52. Each video was rated on a continuous scale ranging from 1 (not at all true) to 10 (totally true). *p < 0.05. **p < 0.01.*

(i.e., ambiguous video pairs), while seven other video pairs were non-ambiguous.

Additional Analyses

As exploratory analyses, given that the basic designs of Studies 1a (laboratory setting) and 1b (online setting) were identical, we merged the data from both studies into a single data sheet and ran additional analyses. A within-subject ANOVA confirmed that overall participants were able to distinguish between false statements ($M = 5.28$, $SD = 0.99$) and true statements ($M = 5.81$, $SD = 0.94$), $F(1,116) = 42.83$, $p < 0.0001$, $\eta^2_p = 0.27$.

Additional within-subject ANOVAs for each video pair were also conducted to determine which video pairs were ambiguous and which were non-ambiguous (for detailed results, see **Table 3**). Taken together, the analyses confirmed that for seven video pairs there were no significant differences between the veracity ratings for the true and the deceptive statements (i.e., ambiguous video pairs), while for the seven other video pairs there were significant differences (i.e., non-ambiguous video pairs).

Discussion

When comparing the results of Study 1a and Study 1b, six video pairs were classified as ambiguous in both studies and six video pairs were classified as non-ambiguous in both studies. In addition, one video pair was classified as ambiguous in Study 1a but not as such in Study 1b, and one video pair was classified as ambiguous in Study 1b but not as such in Study 1a. When merging the data from both studies into a single data file, the results remained stable. The fact that the results of both studies were so similar, even though Study 1a was conducted in a laboratory while Study 1b was conducted online, suggests that our stimulus material is also suited to be applied in future online research.

STUDY 2

The aim of Study 2 was to investigate whether the type of information given to the participants influences the accuracy of the veracity ratings. Previous research suggests that verbal behavior (i.e., the content of the statement) is more reliable than non-verbal and para-verbal behavior (e.g., Bond and DePaulo, 2006). This is why we tested the assumption that participants watching mute versions of the 28 video clips would be less accurate in their veracity ratings than participants watching the regular video clips and participants only listening to the verbal statements without actually seeing the video clips.

Methods

Participants

The sample consisted of $N = 145$ students from a Swiss and a German university who volunteered to take part in this study (87 female; $M_{Age} = 27.03$ years, $SD_{Age} = 7.4$); 22 participants had refereeing experience ($M = 1.85$ years, $SD = 0.36$). Before starting the experimental procedure, each participant delivered written informed consent.

Design, Procedure, and Measures

The experiment took place in the universities' laboratories, and the videos, as well as all the instructions, were displayed on a regular computer screen. The general experimental setup was identical to the two previous studies, with the only difference being that we manipulated the type of information given to the participants. Participants were randomly assigned to a condition that included the original video clips (i.e., original condition; $n = 49$), one that included the original video clips without any auditory information (i.e., mute condition; $n = 47$), or one that only included the audio stream of each video clip without any visual information (i.e., audio condition; $n = 49$). The resulting design was a 2 (true vs. false statements) × 3 (original condition vs. mute condition vs. audio condition) mixed design with repeated measurement on the first factor and a between-participants manipulation on the second. In all three conditions, participants were wearing regular stereo headphones and, as in the previous studies, participants rated the truth of each statement on a continuous scale from 1 (*not at all true*) to 10 (*totally true*; Marksteiner et al., 2013). Finally, we thanked the participants for their participation and debriefed them.

Results

Main Analyses

We conducted a 2 (type of statement: true vs. false) × 3 (experimental group: original condition vs. mute condition vs. audio condition) mixed-design ANOVA to test our hypotheses. There was a significant main effect of the type of statement, $F(1,142) = 10.45$, $p = 0.002$, $\eta^2_p = 0.07$. We also found a significant main effect of the experimental group, $F(2,142) = 9.35$, $p < 0.0001$, $\eta^2_p = 0.12$. As expected, there was also a significant interaction between the type of statement (true or false) and the type of experimental group (original condition vs. mute condition vs. audio condition), $F(2,142) = 6.91$, $p = 0.001$, $\eta^2_p = 0.09$. Mean estimations of statements' veracity suggest

TABLE 3 | Study 1a and b combined: Veracity ratings for the true and the false statements of each video pair.

Video pair number	True statement		False statement			
	M	*SD*	*M*	*SD*	*F*	η^2_p
1	5.74	2.67	5.65	2.61	0.86	0.00
2	5.57	2.56	4.48	2.60	14.13**	0.11
3	6.66	2.41	5.49	2.51	14.74**	0.11
4	4.40	2.48	4.50	2.56	0.11	0.00
5	7.98	2.16	6.35	2.48	31.58**	0.21
6	5.58	2.74	4.14	2.55	26.49**	0.19
7	6.97	2.31	5.58	2.77	22.25**	0.16
8	5.88	2.74	5.91	2.74	0.10	0.00
9	7.18	2.30	6.36	2.84	8.09*	0.07
10	5.97	2.85	5.66	2.70	0.97	0.01
11	3.56	2.63	4.66	2.70	24.45**	0.17
12	6.28	2.51	5.73	2.73	3.25	0.03
13	5.14	2.47	5.40	2.61	0.83	0.01
14	4.46	2.56	4.05	2.33	2.68	0.02

*Note. N = 117. Each video was rated on a continuous scale ranging from 1 (not at all true) to 10 (totally true). *p < 0.05. **p < 0.01.*

TABLE 4 | Study 3: Mean veracity ratings for the true and false statements, separated by condition (original, mute, audio).

Condition	True statements		False statements	
	M	*SD*	*M*	*SD*
Original	5.60	0.79	5.20	0.86
Mute	4.97	0.85	5.10	0.76
Audio	5.89	0.86	5.52	0.99

N = 145. Each video was rated on a continuous scale ranging from 1 (not at all true) to 10 (totally true).

that only participants in the original condition and in the audio condition were able to distinguish between true and false statements, but not participants in the mute condition (**Table 3**). Follow-up t-tests suggest that, indeed, participants in the original condition ($t[48] = 3.14$, $p = 0.003$) and in the audio condition ($t[48] = 3.24$, $p = 0.002$) were able to distinguish between true and false statements, whereas participants in the mute condition were not ($t[46] = -1.42$, $p = 0.16$). Importantly, both main effects are qualified by this interaction.

It is an interesting question whether the above-described interaction between the type of statement and the type of experimental group is driven by both the true and false statements or solely by one group of statements. Furthermore, it is an open question whether the finding is driven more strongly by the more ambiguous or the less ambiguous statements. Therefore, we conducted some additional analyses to address these questions. Contrary to the main analysis, these additional analyses were exploratory in nature.

Additional Analyses

When looking at the true and false statements separately, data suggest that participants in the original condition and in the audio condition were better at identifying true statements as true than participants in the mute condition. Mean estimates for the true statements (on a scale where high values indicate truth) are higher by almost one standard deviation in the original condition and in the audio condition than in the mute condition (**Table 4**). As can be expected, a one-way between-group ANOVA delivered a significant main effect of the type of information given on the ratings of the true statements, $F(2,142) = 10.55$, $p < 0.0001$, $\eta^2_p = 0.17$, which can be considered a large effect (Cohen, 1988). There is no such difference for the false statements, $F(2,142) = 3.03$, $p = 0.051$, $\eta^2_p = 0.04$.

Studies 1a and 1b indicated that individuals were unable to differentiate between the true and the deceptive statements in six video pairs, meaning the deceptive statements of six target players were rather difficult to differentiate from their true statements. Therefore, we were also interested in whether the type of information given affected the veracity ratings of these six target players (the high-ambiguous pairs) differently from the veracity ratings of the other eight target players (the low-ambiguous pairs). To investigate this question, we conducted a 2 (type of statement: true vs. false) × 2 (ambiguity of statement: high vs. low) × 3 (experimental group: original condition vs. mute condition vs. audio condition) mixed-design ANOVA with repeated measures on the first and second factors. The crucial triple interaction does not become significant, suggesting that the pattern described above does not differ significantly for the high-ambiguous and low-ambiguous video pairs, $F(2,142) = 1.28$, $p = 0.282$, $\eta^2_p = 0.018$.

Discussion

Participants in the original condition and in the audio condition were able to distinguish between true and false statements, whereas participants in the mute condition were not. As hypothesized, this pattern suggests that verbal information (i.e., content of statements) is necessary for distinguishing true and false statements, whereas non-verbal behavior alone is not sufficient. Further analyses suggest that the above-described effect was driven primarily by the true statements: participants in the original condition and in the audio condition rated the true statements as being truer, but they did not rate the false statements as being more false. This pattern suggests tentatively that the verbal content of the true statements included cues signaling their truth-value. Alternatively, one might also reason that visual content is not as important in the true statements as in the false statements. We cannot rule out the possibility that a third variable that is confounded with our experimental manipulation might at least partly drive the differences between the experimental conditions. For example, it might be that the experimental manipulations induce different levels of cognitive load (e.g., Paas et al., 2003) or mental fatigue (e.g., Englert, 2019). To the extent that cognitive load as well as mental fatigue influence the ability to distinguish between true and false statements, this would explain the differences between the experimental conditions. Therefore, we suggest that future research utilizes control variables (e.g., a scale assessing cognitive load or mental fatigue) in order to rule out alternative explanations.

GENERAL DISCUSSION

Taken together, results from the three studies suggest that participants are able to distinguish between true and false statements when confronted with our stimulus material. When looking at all 14 video pairs at the same time, the mean difference between the veracity ratings of true and false statements is small, indicating that participants are not very good at discriminating true and false statements, as prior research suggests (e.g., Bond and DePaulo, 2006). Interestingly, however, when looking at all 14 video pairs separately, it becomes apparent that the generally small difference when considering all pairs at the same time is caused by averaging differences of varying sizes: for some video pairs, there are large differences between true and false statements, indicating that participants were able to distinguish well between truth and lies. For other pairs, these differences are smaller, and for some, there are no differences at all, indicating that, on average, participants were unable to distinguish between truths and lies for these pairs. Importantly, this pattern is almost identical in Studies 1A and 1B, suggesting that the differences described above do not merely capture random variation, but constitute systematic differences between videos.

There are two possible (and admittedly speculative) explanations for this pattern of results: first, it might simply be that some video pairs contained valid cues, whereas others did not. Second, it might be that all video pairs contained valid cues, but participants did not use them in some of the videos. For example, participants might not have used some cues because they were not in line with their preconceptions about what constitutes valid cues for lie detection, as suggested by some previous research (e.g., Vrij et al., 2006). The results of Study 2 suggest that the verbal cues were valid, unlike the non-verbal cues. This becomes evident through the observation that the participants in the mute condition were less adept at discriminating between true and false statements. Again, this finding is in line with previous research (e.g., Bond and DePaulo, 2006). We believe that the above described differences between video pairs might be important from theoretical, methodological, and applied perspectives, and thus inspire future research. From a theoretical perspective, future research might shed light onto the reasons underlying differences in discrimination. From a methodological perspective, differences between video pairs emphasize that researchers need to be cautious when averaging single items (e.g., Rose, 2016). From an applied perspective, finally, researchers might want to investigate whether discrimination performance can be improved in general, and for the hard-to-discriminate statements in particular. When there are differences regarding the ratings of the two different videos of one player, it seems obvious to interpret this difference as being caused by the videos' veracity. However, we consider it important to keep in mind that other differences between the videos might have caused the different ratings.

The current results are in line with the interpretation that participants think that the players were in general truthful and thus rated them as such. It is an open question how our results would look like if participants had assumed that players are generally untruthful. Therefore, we suggest that future research investigates into participants' presuppositions regarding players' truthfulness.

The present research has some strengths, but also some limitations. The main strength of the present paper is that we employed an experimental manipulation of lying and telling the truth, which was rather naturalistic at the same time. Obviously, a referee in a real match would never turn away from the action on purpose; however, this was simply our operationalization of a referee not having seen the relevant action. In real matches, it does happen that referees do not see the relevant action, for example, because they were in a position where their line of sight was obstructed. In these situations, referees may well communicate with players. If a scenario like the one from our studies happened in the real world, referees would probably talk to both players involved. In this case, their task would be slightly different from the one employed in our studies: Referees would not necessarily have to judge who is lying and who is saying the truth, but they would have to judge which player they consider to be more trustworthy—even if both of them might appear to be lying (or vice versa). We somewhat simplified this situation for our studies, but the essential task is the same: Judge the veracity of a given statement based on the cues available. Additionally, we consider it to be a strength of the present research that the present findings do not rely on a single study alone but on a set of three studies.

The main limitation of the present research is that, obviously, we cannot make inferences beyond the stimulus material used in our studies. Therefore, we cannot be sure that the findings described above can be generalized to other settings (e.g., other players being filmed, other reasons to lie, or other questions asked). However, we tried to incorporate at least some variation into our stimulus material by filming 14 soccer players and by asking them a set of standardized questions. Additionally, the differences observed between video pairs suggest that we managed to capture at least some variation in answering behavior. Still, future research should replicate the present findings using novel stimulus material. We would also like to mention that our video clips only included male soccer players. Even though previous research from the criminal justice system has revealed that gender does not have a significant effect on veracity judgments (e.g., Bond and DePaulo, 2006), we would recommend a replication of our findings with female soccer players to increase the generalizability.

Future studies might look at a direct comparison between a player who is lying and one who is telling the truth, as suggested above. For this research objective, the scenario from our studies would have to be adapted, so that both players involved can both lie and tell the truth. From a theoretical perspective, this approach might be fruitful not only for research on lie detection in sports, but for research on lie detection generally: In contrast to many other applications of lie detection research, in the situation that we investigated, when one player lies, the other one must be telling the truth. Therefore, this situation allows for investigating into the role of the relative veracity of statements as compared to the role of the absolute one, a comparison that is not possible in many classic lie detection scenarios.

Another important question is which potential factors determine whether individuals have a tendency to rely on valid verbal cues or on stereotypical non-verbal cues. According to Reinhard (2010), NFC is an important factor in this regard. NFC is a personality trait which can be defined as cognitive motivation, meaning the tendency to engage in and enjoy cognitive effort (Cacioppo and Petty, 1982). Previous research has shown that individuals high in NFC primarily base their judgments on valid verbal cues whereas lower levels of NFC are related to a predominant use of non-verbal cues (Reinhard, 2010). Future studies should focus on potential moderators of lie detection performance. One further limitation of the current research is that its relation to real-life situations is not very strong. Thus, we caution against prematurely deriving practical recommendations from this research. Furthermore, we suggest that future research tries to connect to real in-game situations more strongly.

We would also like to acknowledge that our interpretations mostly hinge on comparing the two videos of one pair to each other. That means, we conclude that participants could successfully distinguish lie from truth when the two examples of one video pair differed significantly from each other. However, we did for the most part not take into account the absolute ratings of the respective videos. For example, the ratings of two videos might significantly differ from each other, but both videos are rated in the upper half of the scale, signaling that participants rated both of them being rather true than false, albeit with one of them being rated "truer" than the other. For the goals of our studies, this approach seems to make sense, as we were primarily interested in participants' ability to distinguish truth from lies, and additionally we were interested into first evidence regarding the cues participants base their ratings on. However, research with different goals will probably have to use different comparisons. For example, researchers interested into the question whether different persons are rated differently regarding their trustworthiness will probably want to look at comparisons between persons, and not between video pairs. Likewise, researchers interested into the question whether certain factors may influence participants' general willingness to believe that a statement is true or false (i.e., their prior beliefs) will probably want to investigate whether videos' ratings depend on the variation of the assumed influencing variables. In order to obtain a comprehensive picture of lie detection, researchers would probably have to model veracity ratings as a joint function of (a) participants' prior beliefs about the likelihood of a statement being true or false (while these beliefs might themselves vary based, for example, on context); (b) person characteristics of the potential liars that have been shown to influence veracity ratings (and which might themselves interact with participants' characteristics and prior beliefs; (c) cues inherent in the statements themselves (e.g., verbal and non-verbal cues); and (d) contextual factors influencing the ability to detect lies (e.g., contextual load). Such an approach would surely be able to overcome much of the shortcomings of current research on lie-detection and thus it might be able to provide a more comprehensive picture of lie-detection.

Given the still preliminary and not highly applied nature of our research, we are careful to derive any practical recommendations for referees. Our research appears to suggest that referees should not try to deduce lies or truths from non-verbal behavior, but rather should rely on verbal cues. This suggestion is supported by previous research in domains other than refereeing (Vrij et al., 2010), and it is supported by our present data. Given the overall poor ability of humans to discriminate between true and false statements (e.g., Bond and DePaulo, 2006), at least without formal analysis, as employed in legal proceedings, we suggest that basing a decision on a judgment about the veracity of a statement without further information should probably be employed as a last resort only.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

CE and GS equally contributed to the conceptualization of the studies, review of relevant related work, and writing of the manuscript. Both authors approved the final version of the manuscript and agreed with the order of presentation of the authors.

FUNDING

The authors did not receive any specific funding for this research.

REFERENCES

Aamodt, M. G., and Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Foren. Exam.* 15, 6–11.

Akehurst, L., Koehnken, G., Vrij, A., and Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behavior. *Appl. Cogn. Psychol.* 10, 461–471.

Aragão e Pina, J., Passos, A., Araújo, D., and Maynard, M. T. (2018). Football refereeing: an integrative review. *Psychol. Sport Exerc.* 35, 10–26. doi: 10.1016/j.psychsport.2017.10.006

Bond, C. F., and DePaulo, B. M. (2006). Accuracy of deception judgments. *Pers. Soc. Psychol. Rev.* 10, 214–234. doi: 10.1207/s15327957pspr1003_2

Cacioppo, J. T., and Petty, R. E. (1982). The need for cognition. *J. Pers. Soc. Psychol.* 42, 116–131. doi: 10.1037/0022-3514.42.1.116

Cohen, J. (1988). *Statistical Power Analysis For The Behavioral Sciences*, 2nd Edn, Hillsdale, NJ: Lawrence Earlbaum Associates.

del Campo, V. L., Fariñas, A. C., Márquez, F. J. D., and Martín, J. M. (2018). The influence of refereeing experiences judging offside actions in football. *Psychol. Sport Exerc.* 37, 139–145. doi: 10.1016/j.psychsport.2018.03.004

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., and Cooper, H. (2003). Cues to deception. *Psychol. Bull.* 129, 74–118. doi: 10.1037/0033-2909.129.1.74

Englert, C. (2019). The self-regulation of human performance: a critical discussion and future directions for self-control research. *Perform. Enhan. Health* 6, 156–157. doi: 10.1016/j.peh.2019.04.001

Forrest, J. A., Feldman, R. S., and Tyler, J. M. (2004). When accurate beliefs lead to better lie detection. *J. Appl. Soc. Psychol.* 34, 764–780. doi: 10.1111/j.1559-1816.2004.tb02569.x

Frank, M. G., and Ekman, P. (2004). Appearing truthful generalizes across different deception situations. *J. Pers. Soc. Psychol.* 86, 486–495. doi: 10.1037/0022-3514.86.3.486

Güldenpenning, I., Kunde, W., and Weigelt, M. (2017). How to trick your opponent: a review article on deceptive actions in interactive sports. *Front. Psychol.* 8:917. doi: 10.3389/fpsyg.2017.00917

Hsu, L. (1997). *Cheating and Sports Rules*. Greece: International Olympic Academy.

Kavussanu, M. (2006). Motivational predictors of prosocial and antisocial behaviour in football. *J. Sports Sci.* 24, 575–588. doi: 10.1080/02640410500190825

Kavussanu, M., and Ntoumanis, N. (2003). Participation in sport and moral functioning: does ego orientation mediate their relationship? *J. Sport Exerc. Psychol.* 25, 1–18. doi: 10.1123/jsep.25.4.501

Kavussanu, M., and Roberts, G. C. (2001). Moral functioning in sport: an achievement goal perspective. *J. Sport Exerc. Psychol.* 23, 37–54. doi: 10.1123/jsep.23.1.37

Kavussanu, M., Seal, A., and Phillips, D. (2006). Observed prosocial and antisocial behaviors in male soccer teams: age differences across adolescence and the role of motivational variables. *J. Appl. Sport Psychol.* 18, 326–344. doi: 10.1080/10413200600944108

Klein Richard, A., Ratliff Kate, A., Vianello, M., Adams, R. B., Bahnik, S., Bernstein, M., et al. (2014). Investigating variation in replicability. *Soc. Psychol.* 45, 142–152. doi: 10.1027/1864-9335/a000178

Levine, T. R., Serota, K. B., Shulman, H., Clare, D. D., Park, H. S., Shaw, A. S., et al. (2011). Sender demeanor: individual differences in sender believability have a powerful impact on deception detection judgments. *Hum. Commun. Res.* 37, 377–403. doi: 10.1111/j.1468-2958.2011.01407.x

Marksteiner, T., Ask, K., Reinhard, M. A., and Dickhäuser, O. (2015). Saving cognitive resources when possible: the role of judgment consequences and the judgment tendency of other teachers in teachers' assessment of students. *Soc. Psychol. Educ.* 18, 735–747. doi: 10.1007/s11218-015-9291-0

Marksteiner, T., Reinhard, M. A., Lettau, F., and Dickhäuser, O. (2013). Bullying, cheating, deceiving: teachers' perception of deceitful situations at school. *Intern. J. Educ. Psychol.* 2, 193–220. doi: 10.4471/ijep.2013.24

Momaya, A., Fawal, M., and Estes, R. (2015). Performance-enhancing substances in sports: a review of the literature. *Sports Med.* 45, 517–531. doi: 10.1007/s40279-015-0308-9

Morris, P. H., and Lewis, D. (2010). Tackling diving: the perception of deceptive intentions in association football (soccer). *J. Nonverb. Behav.* 34, 1–13. doi: 10.1007/s10919-009-0075-0

Ommundsen, Y., Roberts, G. C., Lemyre, P. N., and Treasure, D. C. (2003). Perceived motivational climate in male youth soccer: relations to social –moral functioning, sportspersonship and team norm perceptions. *Psychol. Sport Exerc.* 4, 397–413. doi: 10.1016/S1469-0292(02)00038-9

Paas, F., Tuovinen, J. E., Tabbers, H., and Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educ. Psychol.* 38, 63–71. doi: 10.1207/S15326985EP3801_8

Reinhard, M.-A. (2010). Need for cognition and the process of lie detection. *J. Exp. Soc. Psychol.* 46, 961–971. doi: 10.1016/j.jesp.2010.06.002

Renden, P. G., Kerstens, S., Oudejans, R. R., and Cañal-Bruland, R. (2014). Foul or dive? Motor contributions to judging ambiguous foul situations in football. *Eur. J. Sport Sci.* 14, S221–S227. doi: 10.1080/17461391.2012.683813

Rose, T. (2016). *The End of Average*. New York City, NY: Harper Collins.

Sabag, E., Lidor, R., Morgulev, E., Arnon, M., Azar, O., and Bar-Eli, M. (2018). To dive or not to dive in the penalty area? The questionable art of deception in soccer. *Intern. J. Sport Exerc. Psychol.* 4, 1–12. doi: 10.1080/1612197X.2018.1462100

Stuart, M., and Ebbeck, V. (1995). The influence of perceived social approval on moral functioning in youth sport. *Pediat. Exerc. Sci.* 7, 270–280. doi: 10.1123/pes.7.3.270

Traclet, A., Romand, P., Moret, O., and Kavussanu, M. (2011). Antisocial behavior in soccer: a qualitative study of moral disengagement. *Intern. J. Sport Exerc. Psychol.* 9, 143–155. doi: 10.1080/1612197X.2011.567105

Vrij, A., Akehurst, L., and Knight, S. (2006). Police officers', social workers', teachers' and the general public's beliefs about deception in children, adolescents and adults. *Leg. Criminol. Psychol.* 11, 297–312. doi: 10.1348/135532505X60816

Vrij, A., Granhag, P. A., and Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychol. Sci. Public Inter.* 11, 89–121. doi: 10.1177/1529100610390861

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

CONCEPTUAL ANALYSIS
published: 12 January 2021
doi: 10.3389/fpsyg.2020.613469

The Moral Gatekeeper: Soccer and Technology, the Case of Video Assistant Referee (VAR)

Ilan Tamir[1] and Michael Bar-eli[2]*

[1] School of Communication, Ariel University, Ariel, Israel, [2] Business Administration Department, BGU, Ben-Gurion University of the Negev, Beersheba, Israel

Video assistant referee was officially introduced into soccer regulations in 2018, after many years in which referee errors were justified as being "part of the game." The technology's penetration into the soccer field was accompanied by concerns and much criticism that, to a large degree, continues to be voiced with frequency. This paper argues that, despite fierce objections and extensive criticism, VAR represents an important revision in modern professional soccer, and moreover, it completes a moral revolution in the evolution of the sport as a whole. Theoretically speaking, this technology enables an improvement in the sport's professional standards and its public image and prestige, and especially its moral standards – Fair play. Furthermore, the introduction of this technology makes it possible to discover additional weaknesses (Standardization for extra time, a clear definition of a handball offense and more) that professional soccer regulations will probably be forced to address in the future.

Keywords: sport, soccer, technology, VAR, moral

OPEN ACCESS

Edited by:
Edson Filho,
Boston University, United States

Reviewed by:
Danny Rosenberg,
Brock University, Canada
Alan Zaremba,
Northeastern University, United States

****Correspondence:***
Ilan Tamir
ilantamir125@gmail.com

Specialty section:
This article was submitted to Movement Science and Sport Psychology, a section of the journal Frontiers in Psychology

Received: *02 October 2020*
Accepted: *08 December 2020*
Published: *12 January 2021*

Citation:
Tamir I and Bar-eli M (2021) The Moral Gatekeeper: Soccer and Technology, the Case of Video Assistant Referee (VAR). Front. Psychol. 11:613469. doi: 10.3389/fpsyg.2020.613469

REFEREES' MISTAKES IN SPORT

A central ideal of sport competitions is expressed by the traditional notion of "let the best athlete or team win." That is to say sports governors should try minimize as much as possible mistakes and falsification from deciding who the winner is. In order to increase the probability that this will indeed be the case, referees are installed in order to ensure fair competition. In soccer (and other sports as well; see MacMahon et al., 2015) their possible influence on the outcome of games (and the safety of players) is immense. For example, at least two World-Cup finals were decided by controversial decisions of the referees.

On July 30, 1966, England and West-Germany competed at the Wembley stadium in London in the final game of the World cup. With 11 min of extra time gone and a tied score of 2:2, English striker Geoff Hurst received a cross, swiveled and shot from close range. The ball hit the underside of the crossbar and bounced down. Referee Gottfried Dienst was uncertain, but awarded a goal to England after consulting linesman Tofiq Bahramov from Azerbaijan in the USSR, who indicated that it was a goal. The game ended 4:2 for England and this decision has remained controversial ever since then. Furthermore, it led to the creation of the expression "Wembley goal," a phrase used to describe any "Ghost-"or "Phantom-goal" awarded (but actually not scored) in a similar fashion.

On July 7, 1974, West-Germany and Holland competed at the Olympic stadium in Munich in the final game of the World Cup. In the 25th min, when the score was 1:0 in favor of Holland, German striker Bernd Hoelzenbein received the ball at about 40 m diagonally from the Dutch goal, and started to dribble. While entering the penalty area, he was tackled by the Dutch defender Wim

Jansen, and fell down. The English referee, Jack Taylor, awarded a penalty kick to West Germany, which was used by Paul Breitner to tie the game, and the Germans won 2:1. This crucial referee's decision remains controversial until today due to the fact that Hoelzenbein was accused of being dishonest and faking the collision, which he vehemently denied (Sabag et al., 2018). This example involves player deception and is a judgment call by the referee after witnessing an event.

Although these are two different examples, is an eye-witness error and the other a judgment decision which involves a more complicated situation, in both cases the truth remained in doubt.

Bar-Eli (1994) was appointed psychological consultant to the Israel soccer referees association. One major purpose of this intervention was to help referees improve their performances. While discussing the issue with Abraham Klein, an experienced, highly respected international ex-referee (considered by many as the best Israeli referee ever), who served at that time as chairperson of the association, these two cases were discussed. Bar-Eli argued that soccer would only profit from [A] introducing a goal-line camera to decide whether "the whole of the ball passes over the goal line" (as the laws of the game require for approving a goal) or only 97% of the ball (as more recent studies indicate with regard to Hurst's "Wembley Goal"; see, for example, Reid and Zisserman, 1996); [B] using some other available technology, to make better decisions in ambiguous tackling situations (e.g., such as where a penalty kick should eventually be awarded).

Bar-Eli contended that the use of such technologies (i.e., decision aids) will INCREASE the referee's chances to ensure the realization of the very basic notion of "may the best athlete or team win," thereby STREGTHENING the referee's authority on field.

Referee Klein strongly objected to Bar-Eli's suggestions, arguing that the referee's authority on the field will be damaged (at least at the perceived level) as a result of applying such technologies – even with regard to the simple goal-line decision, not to mention the more complex penalty kick situation. However, ongoing controversies such as the abovementioned ones (i.e., from the World Cup finals of 1966 and 1974), and/or other extreme, decisive referees' errors such as in the case of Maradona's notorious 'Hand of God' affair in the 1986 World Cup, accelerated research on referees in soccer (and umpires, officials, linesmen or judges in other sports). Raab and Helsen (2015) argued that referees' performances are determined by physical (e.g., the position of the referee) and perceptual-cognitive (e.g., interpretation of events on the field) factors. They identified an increasing number of publications focusing on physical (n = 67) and perceptual-cognitive (n = 58) aspects of refereeing between 2000 and 2013. Pina et al. (2018) more recent review revealed a similar trend for publications focusing on physical (n = 74) and perceptual-cognitive (n = 90) aspects of refereeing between 2000 and 2016, but with a greater emphasis on perceptual-cognitive aspects such as judgment and decision making (JDM) and perspectival bias (see also Helsen et al., 2019).

The growing interest in researching referees' JDM reflect a "psychology OF sport"- perspective that is, the use of psychological knowledge to cope with a practical problem, in this case, referees' mistakes. This view acknowledges the commercial, media and financial interests in the business of soccer (and sport in general), in which a sound referees' JDM is required to minimize the negative effects of erroneous officiating decisions on match outcomes (Helsen et al., 2019). However, the increasing interests in referees reflect also a "psychology IN sport"-perspective, that is, a view which contends that "studying sport is a great idea, because people make many decisions that matter enormously to them under standard conditions. It is actually one of the places to do this" (Kahneman, 2008). Indeed, referees' JDM has been increasingly viewed as one of the best fields to study human JDM processes in general, as evident from Nobel-Prize winner Daniel Kahneman's quote, and also from earlier reviews of JDM in sport (Bar-Eli et al., 2011) as well as from more recent ones (Raab et al., 2019a,b).

Sport fields are not environments that are conducive to every JDM process. This is true for all the parties involved, from players to coaches, and especially referees. In soccer, the rapid movements of the ball and the players, the varying sight lines, and the heated passions, pressure, and competition expose referees to very complex, sometimes impossible, officiating situations. More specifically, top-referees are conceived as experts who make decisions in dynamic, time-constrained sporting environments (Raab and Helsen, 2015). Many studies from countries around the world that have examined JDM of referees in a wide range of sports, as will be detailed later, have found that referee JDM are affected by significant biases that directly affect games and their results.

PSYCHOLOGICAL BIASES

As mentioned above, much of the "raison d'etre" of the "psychology OF sport"-approach toward the study of referees' JDM, reflects an explicit or implicit intent of trying to minimize the probability of referees' mistakes. Accordingly, this literature has been quite focused on the concept of bias, which was defined as distortion of measurement or evaluation results leading to their misinterpretation (Helsen et al., 2019; Raab et al., 2019a).

However, the very notion of "bias" requires a benchmark of "something which is NOT biased," to be compared with. In other words, we have to clarify what a GOOD decision of a referee actually is.

At first glance, this seems quite easy, because the primary, very essential reason of using referees at all, would be the enforcement and interpretation of the laws of the game, which motivates the referees to be as ACCURATE as possible. This would mean that if the referee is inaccurate, he or she may be biased in some way or another. However, Bar-Eli et al. (2011) – based on previous work conducted by Plessner and his team – maintain that soccer referees can also conduct GAME MANAGENENT, which is intended to ensure the flow of the game and be (or at least appear) unbiased, with accuracy considered less important than "just return home safely" (i.e., officiating a game without any noticeable incidents). These two strategies may point to the same direction, but they can also get into conflict. Thus, while most of the research on referees' biases is concerned with (in)accurate JDM, referees should also be investigated as to the ways in which they adjust their interpretation of incidents to the concrete context of the situation in question (Bar-Eli et al., 2011). The basic premise, therefore, is that increasing the use of objective, seemingly accurate (technological) means will help reduce biases and increase fairness.

More recently, Bar-Eli (2018) reviewed some of his own work on referees (though mainly in basketball) and concluded that in general, referees call fewer fouls than those judged by (basketball) experts as deserving a call. This "conservative" behavior can be explained both by rational reasons and biased JDM. The best objective referees' desire is that nobody question the accuracy and fairness of their calls in the contest. This strategy makes officiating mistakes and interpretations, not only understandable, but also, though paradoxically, rational.

Similar trends were revealed among soccer referees (e.g., Sabag et al., 2018), who strive for accuracy and are also aware of the severe consequences of player dismissals (i.e., red cards; see Bar-Eli et al., 2006) – Referees sometimes try to avoid using measures such as yellow or red cards too soon in the game, thereby producing a possible escalation in which too many dismissals would ruin the game (and the respective TV-ratings and broadcasting income). These considerations leave us again with the open question of what can be considered good referee decisions (Bar-Eli et al., 2011).

Many variables naturally affect decision making in any area. The field of sports, however, accommodates an extraordinary concentration of passions, pressures, and emotions (Hanin, 2007), whose effect on referee JDM process is direct and immediate. One of the manifestations of this effect is compensating bias, which is when referees who make decisions in favor of one team try to even out the situation in subsequent decisions. These referees will impose more stringent criteria against the team that won the previous call (on an offense, sending off a player, etc.) and will lower the criteria when calling against a rival.

A study conducted in Germany (Schwarz, 2011), which examined the penalties in the local soccer league over four decades, is an example of numerous studies that found proof of compensating bias. In the majority of games with two or more penalty kicks, the calls were divided equally in favor of both teams. The referees effectively raised their criteria for awarding a second penalty to the team that already received a penalty call, and relaxed the criteria for the opposing team. The study also examined the timing of the whistle for penalty kicks and found that when two consecutive penalty kicks were awarded to two different teams, the interval between the referees' calls was significantly shorter than when two consecutive penalty kicks were awarded to the same team.

Studies on the NCAA, the US college basketball league (Anderson and Pierce, 2009; Noecker and Roback, 2012) found the exact same pattern: Referees tend to call more fouls against the team that has fewer penalties in order to even out the competition. Moreover, referees also call more penalties against the team that is leading. A similar study (Plessner and Betsch, 2001) found that a referee's previous call affects his subsequent decisions in the game, reflecting a tendency to even out his decisions in the game. The expectation from the referees in the game therefore is to whistle according to the events, without taking into account extraneous considerations, which can impair the accuracy of the decisions.

Another manifestation of the psychological effects that influence referees is related to reputation or prior knowledge bias. The reputation of a team or a player affects referees' calls during a game. The assumption that a referee can ignore all the previous knowledge he has of the players is not borne out by evidence. A study by Jones et al. (2002) illustrates this idea: 38 soccer referees viewed several video clips of segments of taped soccer matches. The same team (wearing a blue uniform) appeared in all the clips, playing against a different rival in each case. The referees were divided into two groups and were instructed to describe the call they would make in each case. The instruction sheet handed to one group contained a comment that the blue team is considered an aggressive team. Interestingly, while no differences emerged in the decisions of the referees in both groups (e.g., the number of fouls they called), the groups differed in the interpretation that they gave to each event. The referees who had prior knowledge of the blue team's "past" awarded more yellow and red cards than the other referees. In other words, referees' previous knowledge of the team directly informed their decision making.

Referees are subject to many more psychological effects, including the pressure imposed on them by the crowd (e.g., Nevill et al., 2002), which has been shown to affect decisions including stoppage time (Garicano et al., 2005; Scoppa, 2008), penalty awards (Dohmen, 2008), and red and yellow carding (Buraimo et al., 2012).

Some effects are less intuitive, for several years, researchers have noted that the color of players' uniforms affect their decisions. For example, red has a positive impact on the outcome of a contest. The fact that contestants are more successful when they wear red has been proven in a broad range of competitions and sports (Hill and Barton, 2005). According to one explanation, cultural and evolutionary variables link the color red to dominance and aggression and thereby psychologically affect contestants (Elliot and Maier, 2014; Meier et al., 2015). Namely, wearing red enhances one's dominance, aggressiveness, and testosterone, which facilitates competitive outcomes. Ilie et al. (2008) even recognized highly significant effect on the performance of red teams in a popular multiplayer first-person-shooter (FPS) computer game. This effect of course does not apply to soccer referees or result of their decisions.

Referees are also affected by the colors that players wear. One study conducted recently (Hagemann et al., 2008) demonstrated how referees (in this case, in taekwondo), give higher scores to competitors who wore the color red. One of the researchers' conclusions was that the prominence of specific colors, and especially red, allow referees to more easily assess the many moves that players in these colors perform. In contrast to other colors that blend into the background and conceal players' moves from the referee's notice, the color red attracts the referee's attention and affects his partiality. Similar findings have also been reported in the England soccer leagues (Attrill et al., 2008; Olde Rikkert et al., 2015), in the Turkish soccer league (Tiryaki, 2005) and more.

PHYSICAL LIMITATIONS

Many sports events are inherently dynamic. Motion and speed are frequently the main elements of a game, and it has become almost impossible for a referee to attend to all the events occurring on the field (the single referee in soccer is certainly a problem toward achieving good officiating decisions). As a result, some sports have multiple referees while other sports use linesmen or a crew of assistant referees, for example. But does the added assistance always meet the challenges of decision making? Not necessarily. Sometimes the angle of a referee's sight of the ball, the players, or other assistant referees, can lead to error. Studies that examined the location of linesmen on the soccer field found that errors regarding offside calls were almost inevitable (Mallo et al., 2012). Studies from various countries indicate that between 17 and 25% of all decisions made by referees and linesmen in a game are inaccurate (Gulec et al., 2018).

The implications are enormous. Linesmen make dozens of calls in every game, and for every ten offside decisions, one or two calls are in error. The error of failing to raise a flag (when the linesman should signal for an offside) occurs more frequently than the error of raising the wrong flag (Mallo et al., 2012). The VAR in this case can be of great help when it oversees the decisions of the linesmen (approve or disqualify a goal as a result of an offside for example; approve or disqualify a penalty kick called by the linesmen, etc.). When we look at the errors made by referees, the number is even larger, although referees are less restricted in their position on the field than linesmen, and a share of these errors can be prevented by improving referees' physical condition and positioning. In any case, referee errors have become such an integral part of the game that the term "ghost goal" has become common in soccer parlance, and is used to express the many cases of debatable or questionable goals (Similar to other examples like "Hand of God" and others).

SPORTS AND TECHNOLOGY

Although the integration of technology in sports, and specifically technology designed to support referee decision making, is warranted in view of the distortions described above, it seems that sports has assumed a double role with respect to technology (Tamir, 2019). On the one hand, the sports industry is leading significant technological revolutions. This is certainly true in the field of media (Galily and Tamir, 2014). Sports is considered a major agent in the introduction of various technologies including plasma screens and HD- and 4K-quality broadcasts in the home. On the other hand, sports' religious-like devotion and commitment to its communities and traditions (Bain-Selbo and Sapp, 2016) highlight the conservative elements of the industry and inhibit new technology adoption.

Loland (2002) argues that the integration of technology into sports should be assessed through one of the following three perspectives, based on the defined aim of sports: If sports is a means to achieve external goals such as political, ideological, or financial prestige, technology's significance lies in its ability to achieve these goals (by creating manipulation and giving an advantage to the athlete). Moral issues related to technological integration do not appear to be relevant in such a case. The history and role of sports in the Cold War (Dimeo, 2007) or, alternatively, in contemporary economic rivalries (Simon et al., 2015) illustrate the relevance of this argument.

If sports is a platform for realizing physical potential, technology's significance lies in its ability to help individuals improve their performance. In this case, too, history indicates that many athletes and sports professionals have used this argument to support the use of prohibited substances. At the same time, within the context of this goal, technology should ensure standardization of performance and ensure balance, credibility, and validity of assessments, evaluations, and judgments.

The third perspective takes a broader approach and views sports as an arena of potential human development. Sports is a sphere with a set of values and encompasses more general human virtues and merits, and the moral virtues of human development. Sports focuses on normative values and behaviors (as respect for the rules of the game and the competitors), and Pierre de Courbrtin's view on reviving the Olympic Games is perhaps emblematic of this view (MacAloon, 2013). In contrast to the previous approaches, which tend to accept any technology that is instrumental in achieving an external goal or enhancing performance, this perspective assumes a clearly moral stance that emphasizes the journey. Technology might lead to better outcomes but if the athlete is unable to control it, it does not have a value. Improving performance without athletic effort has no value.

The involvement of technology in sports can be generally divided into two: technologies that help promotes athletes' achievements and technologies that are used as a governing mechanism of sport. The performance development range is very wide, from body suits in swimming designed to reduced water friction, or shorter alpine skis with radically improved carving capabilities to potentially performance-enhancing genetic technologies. Fundamental resistance to the introduction of new technologies into sports, in this context, is related to the physical component in the definition of sport. That is to say, according to critics, it is the human body and not a machine that should be trained to overcome challenges and natural attributes (Fouché, 2017).

In the second category, many innovations that have penetrated the various sports industries in recent years can be identified, as will be detailed below, mainly with the aim of assisting the referees and regulatory process.

Of course there is a connection between the two, and many times the better the athletic performance the harder it is to make good calls, because of the increased speed and skill (in many sports and certainly in soccer).

TECHNOLOGY AND REFEREEING

Technology's penetration into sports, as a governing mechanism, has been clearly felt in recent decades (Fouché, 2017), creating a significant impact on the entire spectator experience (Dyer, 2015). If the basic assumption was always that the match officials were close to the events and had a better view than anyone else (for example, the high position on the tennis court), the age of television changed the rules (Collins, 2019). The high quality broadcasts and the television replays may put the TV viewers in a better position than the referee when it comes to identifying actions and situations. Naturally, this reality has created a growing sense of unease and distrust among viewers. Therefore, it was only a matter of time before the gaps narrowed. And so, replays (Collins, 2010), photo finishes, goal-line technology, hawk-eye systems, and other technologies, designed to assist the referee, have transformed sports into a more accurate space than before. Some of the technologies provide autonomous assistance, which means giving an indication to the referee during the game (such as in rugby, or in some cases in cricket), and some based on player challenge (such as tennis). As technology will improve, the second type is likely to approach the first, and in each case the intention is, in both cases, to use the technology to reduce the errors of referees and return viewers' confidence in the sport. Although, the introduction of new technology was accompanied by criticism (Dyer, 2015).

For example, the aspiration and expectation of perfect accuracy from the technologically assisted is probably impossible (Collins, 2019). In part, because the lines drawn on sports fields and the edges of balls are not perfectly defined (p. 21) but more importantly, because the measuring devices are based on the world of virtual reality, not the actuality of a physical world. The technology do not show what happened, but a statistical assessment of what probably happened (devices such as ball-trackers, not showing what actually happened but only a statistical estimate of what might have happened) (p. 25).

Another criticism is related to the continuity of the game. The argument is that the use of technology takes a long time to decide and impedes the flow of the game. The frequent stops are tedious and make the game, exhausting and damage the entertainment component of the game (Ryall, 2012). Collins (2019) argues that the guiding principle is to play the game with technology as close as possible to the game without the technology.

Alternatively, some also claim that the use of technology is also not always being correctly applied, because there are some sport situations whereby technology cannot conclusively affirm a correct decision. Some decisions, such as whether the ball crossed the goal line or not, are allegedly simply right or wrong. Immediate and accurate goal technology seems to be a clear step forward. Other decisions involving referee judgment, for instance on player intention and potential sabotage of the game (such as decisions on 'professional fouls' and yellow or red cards), may be more complicated. One might argue that they are best made in the flow and full context of the game, and that video replays, sometimes in slow motion, can lead to misinterpretations.

SOCCER AND VAR

Video assistant referee technology is an example of autonomous assistance given to soccer referees. Although VAR was tested for the first time during the 2012–2013 season, it was officially introduced into the Laws of the Game in 2018 to help referees in reviewing decisions made by the head referee by means of video footage only for three main situations and one administrative incident (Lago-Peñas et al., 2019).

According to the IFAB (International Football Association Board), Principles, a video assistant referee (VAR) is a match official, with independent access to match footage, who may assist the referee only in the event of a 'clear and obvious error' or 'serious missed incident' in relation to:

(a) Goal/no goal; (b) Penalty/no penalty; (c) Direct red card (not second yellow card/caution); and (d) Mistaken identity (when the referee cautions or sends off the wrong player of the offending team). The original decision given by the referee will not be changed unless the video review clearly shows that the decision was a 'clear and obvious error.' The final decision is always taken by the referee, either based on information from the VAR or after the referee has undertaken an 'on field review' (OFR).

In soccer, the introduction of the new technology has been the target of intense criticism. Top soccer executives including former FIFA President Sepp Blatter, voiced their objections to the integration of advanced technologies into soccer games (CBC Sports, 2008), echoing objections made against previous technologies designed to improve the game, such as goal-line technology (Ryall, 2012). The source of opposition in all cases was soccer executives' stance on the nature and authenticity of the game. Their basic assumption was that soccer is the most popular sports in the world due to its simplicity and authenticity. The introduction of technology, they argued, would undermine the deep roots of the most popular game in the world (Walsh, 2011). In other words, human errors are an inevitable part of the game, and even part of its charm. Other objections warned against the time that would be wasted in the game as a result of repeated viewing of video-recorded moves; like other real-time video-replay devices, the criticism of VAR was the possible disruption to the flow and pace of the game due to the stopping and starting (Dyer, 2015). These interruptions to the flow of the game were expected to cause viewers to lose interest. However, according to the FIFA website (FIFA, 2018), technology actually reduces wasted time (time that was apparently taken in the past for arguments with referees). Other critics were concerned about the potential damage to referees' authority (Collins, 2010), and pointed to the large number of penalties in the most recent 2018 World Cup Games in Russia to illustrate how technology changes the game for the worse (29 penalties, more than twice the number of the penalties in the Brazil 2014 games).

Eventually, after years of dispute, soccer executives realized that it was no longer possible to deny that the era of multi-camera HD capture systems and broadcasts had arrived. As viewers at home see each move and referee error in replays showing various angles, soccer and professional leagues should welcome VAR to restore supreme value of the game, namely, fair play. It is important to emphasize that officiating technology will not eliminate all referee errors in sport. As long as human beings are referees there will be officiating mistakes. However, the VAR will help referees make better decisions and rule interpretations, will lessen unfairness, and can be endorsed by the soccer community.

CONCLUSION

The new referee system offers a dramatic contribution to the game of soccer. In Loland's (2002) terms, it gives referees the prestige and honor eroded over the years. The VAR system clearly improves professional decision making during a game, and reinstates in soccer the supreme value of fairness and fair play. After years in which the average viewer saw a large number of referee errors due the fallibility of referees, the VAR has managed to improve fair conditions sought by referees and viewers. From a professional perspective, the VAR system promotes limited impartial and accurate decision making. Although the system does not clarify all potential areas of ambiguity, reviewing a video replay reduces errors and accordingly enhances the professional standards of both referees and players.

Incidents in which players significantly "dive" to obtain a penalty kick or other unprofessional offenses are committed will gradually be reduced (applies to accurate eye-witness reports, and not referee judgments and game management aspects of officiating) and filter out the game's distracting background noises. However, perhaps the most important point is the impact of technology on the moral aspects of the game in relation to human behavior. Fair play is a revered principle of sports. Without ensuring equal chances and fair judgments, sport is sawing off the branch on which it is sitting. That is, it must invest in and make use of technological resources to monitor and enforce the principle of fairness. When circumstances limit the capabilities of referees, especially due to physical limitations, errors could be accepted and contained. However, according to the technological determinism theory (McLuhan, 1964), any new technology that penetrates an industry dictates new standards for evaluating reality, and this is the case for sports (in a limited sense—as far as what cameras focus on and is aired). From the moment viewers at home see moves clearly and sharply, they often voice their disapproval of referees' calls and criticize soccer associations for their errors, demanding accuracy. Officiating technology has the potential to change the game's values and make it more accurate and fairer. Still, the use of VAR is selective, perspectival, two-dimensional and not fully all-encompassing, thus technology does not rule out the efficacy of on-field referees to make accurate and fair decisions.

The argument of the present article is that the use of technologies in sports should be examined in the light of morality and as such, the VAR system, which receives (unsurprisingly) much resistance, should be treated as one of the most important and moral changes in soccer. At the same time, it is important to remember that change in moral conduct requires an internalization by human beings to comprehend wrongfulness, willfully alter their attitudes and then express right actions in sport to the best of their ability.

It is further important to recall that VAR operates under the Laws of the Game. Beyond accuracy, the system also has the potential to reveal additional weaknesses in the Law of the Game that have become integrated into the fabric of acceptable errors (e.g., hand touches, time measurement and other factors).

A good example might be stoppage time in soccer. Stoppage time is an important issue because it can account for more than 10% of the total game time. Studies show that there is not always a connection between the number of minutes a referee decides to add and the time the play was actually stopped (Lago-Peñas and Gómez-López, 2016). Today, as the subjective aspects of decision making have been reduced, and accuracy has become a top priority, stoppage time is yet another issue that can be quantified and redefined through technology.

The introduction of advanced technology in professional soccer has and will continue to significantly improve the game. This is an important step forward for soccer matches and an advancement for all sports. Collins (2019) argues that if referee decisions seem reasonable to both the human eye and television viewers on replay, then a sense of unfairness can be reduced and lessen the annoyance for example of stoppage time inconsistency (based on the principle of continuity). Despite justifiable criticism, the integration of technology in sports to promote the principle of fair play should be encouraged. In the case of soccer, the VAR will increase fairness in officiating and raise the level of impartiality with referee decisions.

AUTHOR CONTRIBUTIONS

Both authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Anderson, K. J., and Pierce, D. A. (2009). Officiating bias: the effect of foul differential on foul calls in NCAA basketball. *J. Sports Sci.* 27, 687–694. doi: 10.1080/02640410902729733

Attrill, M. J., Gresty, K. A., Hill, R. A., and Barton, R. A. (2008). Red shirt colour is associated with long-term team success in English football. *J. Sports Sci.* 26, 577–582. doi: 10.1080/02640410701736244

Bain-Selbo, E., and Sapp, D. G. (2016). *Understanding Sport as a Religious Phenomenon*. London: Bloomsbury.

Bar-Eli, M. (2018). *BOOST! How the Psychology of Sports can Enhance your Performance in Management and Work*. New York, NY: Oxford University Press.

Bar-Eli, M., Plessner, H., and Raab, M. (2011). *Judgement, Decision Making and Success in Sport*. Chichester: Wiley-Blackwell.

Bar-Eli, M., Tenenbaum, G., and Geister, S. (2006). Consequences of players' dismissal in professional soccer: a crisis-related analysis of group-size effects. *J. Sports Sci.* 24, 1083–1094. doi: 10.1080/02640410500432599

Buraimo, B., Simmons, R., and Maciaszczyk, M. (2012). Favoritism and referee bias in European soccer: evidence from the Spanish League and the UEFA Champions League. *Contemp. Econ. Policy* 30, 329–343. doi: 10.1111/j.1465-7287.2011.00295.x

CBC Sports (2008). *FIFA Halts Instant Replay Experiment*. CBC. Available online at: https://www.cbc.ca/sports/fifa-halts-instant-replay-experiment-1.695604.(accessed March 10, 2019).

Collins, H. (2010). The philosophy of umpiring and the introduction of decision-aid technology. *J. Philos. Sports* 37, 135–146. doi: 10.1080/00948705.2010.9714772

Collins, H. (2019). Applying philosophy to refereeing and umpiring technology. *Philosophies* 4, 21–27. doi: 10.3390/philosophies4020021

Dimeo, P. (2007). "Good versus evil? Drugs, sport and the cold war," in *East Plays West: Sport and the Cold War*, eds S. Wagg and D. L. Andrews (London: Routledge), 149–163.

Dohmen, T. J. (2008). The influence of social forces: evidence from the behavior of football referees. *Econ. Inq.* 46, 411–424. doi: 10.1111/j.1465-7295.2007.00112.x

Dyer, B. (2015). The controversy of sports technology: a systematic review. *SpringerPlus* 4:524.

Elliot, A. J., and Maier, M. A. (2014). Color psychology: effects of perceiving color on psychological functioning in humans. *Ann. Rev. Psychol.* 65, 95–120. doi: 10.1146/annurev-psych-010213-115035

FIFA (2018). *VAR Discussed at IFAB Media Briefing. FIFA*. Available online at: https://www.fifa.com/about-fifa/who-we-are/news/var-discussed-at-ifab-media-briefing (accessed March 10, 2019).

Fouché, R. (2017). *Game changer: The Technoscientific Revolution in Sports*. Baltimore, MD: Johns Hopkins University Press.

Galily, Y., and Tamir, I. (2014). A match made in heaven?! Sport, television, and new media in the beginning of the third millennia. *Telev. New Media* 15, 699–702. doi: 10.1177/1527476414541553

Garicano, L., Palacios-Huerta, I., and Prendergast, C. (2005). Favoritism under social pressure. *Rev. Econ. Statist.* 87, 208–216. doi: 10.1162/0034653053970267

Gulec, U., Yilmaz, M., Isler, V., O'Connor, R. V., and Clarke, P. M. (2018). A 3D virtual environment for training soccer referees. *Comput. Stand. Interface* 64, 1–10. doi: 10.1016/j.csi.2018.11.004

Hagemann, N., Strauss, B., and Leißing, J. (2008). When the referee sees red.... *Psychol. Sci.* 19, 769–771. doi: 10.1111/j.1467-9280.2008.02155.x

Hanin, Y. L. (2007). "Emotions in sport: current issues and perspectives," in *Handbook of Sport Psychology*, 3rd Edn, eds G. Tenenbaum and R. C. Eklund (Hoboken, NJ: John Wiley & Sons), 31–58. doi: 10.1002/9781118270011.ch2

Helsen, W. F., MacMahon, C., and Spitz, J. (2019). "Decision making in match officials and judges," in *Anticipation and Decision Making in Sport*, eds A. M. Williams and R. C. Jackson (Abingdon: Routledge), 250–266. doi: 10.4324/9781315146270-14

Hill, R. A., and Barton, R. A. (2005). Red enhances human performance in contests. *Nature* 435:293. doi: 10.1038/435293a

Ilie, A., Ioan, S., Zagrean, L., and Moldovan, M. (2008). Better to be red than blue in virtual competition. *Cyberpsychol. Behav*. 11, 375–377. doi: 10.1089/cpb.2007.0122

Jones, M. V., Paull, G. C., and Erskine, J. (2002). The impact of a team's aggressive reputation on the decisions of association football referees. *J. Sports Sci.* 20, 991–1000. doi: 10.1080/026404102321011751

Kahneman, D. (2008). Personal Communication With Michael Bar-Eli.

Lago-Peñas, C., and Gómez-López, M. (2016). The influence of referee bias on extra time in elite soccer matches. *Percept. Mot. Skills* 122, 666–677. doi: 10.1177/0031512516633342

Lago-Peñas, C., Rey, E., and Kalén, A. (2019). How does Video Assistant Referee (VAR) modify the game in elite soccer? *Int. J. Perform. Anal. Sport* 19, 646–653. doi: 10.1080/24748668.2019.1646521

Loland, S. (2002). Technology in sport: three ideal-typical views and their implications. *Eur. J. Sport Sci.* 2, 1–11. doi: 10.1080/17461390200072105

MacAloon, J. J. (2013). *This Great Symbol: Pierre de Coubertin and the Origins of the Modern Olympic Games*. London: Routledge.

MacMahon, C., Mascarenhas, D. R. D., Plessner, H., Pizzera, A., Oudejans, R., and Raab, M. (2015). *Sport Officials and Officiating: Science and Practice*. Abingdon: Routledge.

Mallo, J., Frutos, P. G., Juárez, D., and Navarro, E. (2012). Effect of positioning on the accuracy of decision making of association football top-class referees and assistant referees during competitive matches. *J. Sports Sci.* 30, 1437–1445. doi: 10.1080/02640414.2012.711485

McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. New-York, NY: McGraw Hill.

Meier, M. A., Hill, R. A., Elliot, A. J., and Barton, R. A. (2015). "Color in achievement contexts in humans," in *Handbook of Color Psychology*, eds A. J. Elliot, M. D. Fairchild, and A. Franklin (Cambridge: Cambridge University Press), 568–584. doi: 10.1017/cbo9781107337930.028

Nevill, A. M., Balmer, N. J., and Williams, A. M. (2002). The influence of crowd noise and experience upon refereeing decisions in football. *Psychol. Sport Exerc.* 3, 261–272. doi: 10.1016/s1469-0292(01)00033-4

Noecker, C. A., and Roback, P. (2012). New insights on the tendency of NCAA basketball officials to even out foul calls. *J. Quant. Anal. Sports* 8, 1–21.

Olde Rikkert, J., Haes, V. D., Barsingerhorn, A. D., Theelen, T., and Olde Rikkert, M. G. (2015). The colour of a football outfit affects visibility and team success. *J. Sports Sci.* 33, 2166–2172. doi: 10.1080/02640414.2015.1064156

Pina, J. A., Passos, A., Araujo, D., and Maynard, M. T. (2018). Football refereeing: an integrative review. *Psychol. Sport Exerc.* 35, 10–26. doi: 10.1016/j.psychsport.2017.10.006

Plessner, H., and Betsch, T. (2001). Sequential effects in important referee decisions: the case of penalties in soccer. *J. Sport Exerc. Psychol.* 23, 254–259. doi: 10.1123/jsep.23.3.254

Raab, M., Bar-Eli, M., Plessner, H., and Araujo, D. (2019a). The past, present and future of research on judgment and decision making in sport. *Psychol. Sport Exerc.* 42, 25–32. doi: 10.1016/j.psychsport.2018.10.004

Raab, M., and Helsen, W. (2015). "How experts make decisions in dynamic, time constrained sporting environments," in *Handbook of Sport Expertise*, eds J. Baker and D. Farrow (Abingdon: Routledge), 64–73. doi: 10.4324/9781315776675-6

Raab, M., MacMahon, C., Avugos, S., and Bar-Eli, M. (2019b). "Heuristics, biases and decision making," in *Anticipation and Decision Making in Sport*, eds A. M. Wiliams and R. C. Jackson (Abingdon: Routledge), 215–231. doi: 10.4324/9781315146270-12

Reid, I., and Zisserman, A. (1996). "Goal-directed video metrology," in *Proceedings of the 4th European Conference on Computer Vision*, eds R. Cipolla and B. Buxton (New York, NY: Springer), 647–658.

Ryall, E. (2012). Are there any good arguments against goal-line technology? *Sport Ethics Philos.* 6, 439–450. doi: 10.1080/17511321.2012.737010

Sabag, E., Lidor, R., Morgulev, E., Arnon, M., Azar, O., and Bar-Eli, M. (2018). To dive or not to dive? The questionable art of deception in soccer. *Int. J. Sport Exerc. Psychol.* 18, 296–307. doi: 10.1080/1612197x.2018.1462100

Schwarz, W. (2011). Compensating tendencies in penalty kick decisions of referees in professional football: evidence from the German Bundesliga 1963–2006. *J. Sports Sci.* 29, 441–447. doi: 10.1080/02640414.2010.538711

Scoppa, V. (2008). Are subjective evaluations biased by social factors or connections? An econometric analysis of soccer referee decisions. *Empir. Econ.* 35, 123–140. doi: 10.1007/s00181-007-0146-1

Simon, R. L., Torres, C. R., and Hager, P. F. (2015). *Fair Play: The Ethics of Sport*. New-York, NY: Routledge.

Tamir, I. (2019). Digital video recorder dodgers–Sport-viewing habits in the face of changing media reality. *Time Soc.* 28, 1319–1332. doi: 10.1177/0961463x17752286

Tiryaki, M. Ş. (2005). Assessing whether black uniforms affect the decisions of Turkish soccer referees: is finding of Frank and Gilovich's study valid for Turkish culture? *Percept. Mot. Skills* 100, 51–57. doi: 10.2466/pms.100.1.51-57

Walsh, D. (2011). 'I don't want GLT, you need it once every 40 years'; michel platini on corruption, the politics of football and why he hates technology. *The Sunday Times* (London), sport section. 8.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINAL RESEARCH
published: 19 October 2020
doi: 10.3389/fpsyg.2020.550271

Development of the Referee Shared Mental Models Measure (RSMMM)

Jorge Sinval[1,2,3,4], João Aragão e Pina[5], João Sinval[6], João Marôco[2], Catarina Marques Santos[7], Sjir Uitdewilligen[8], M. Travis Maynard[9] and Ana Margarida Passos[1]*

[1] Business Research Unit (BRU-IUL), Instituto Universitário de Lisboa (ISCTE-IUL), Lisbon, Portugal, [2] William James Center for Research, ISPA – Instituto Universitário, Lisbon, Portugal, [3] Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal, [4] Faculty of Philosophy, Sciences and Languages of Ribeirão Preto, University of São Paulo, São Paulo, Brazil, [5] ISCTE Business School (IBS), Instituto Universitário de Lisboa (ISCTE-IUL), Lisbon, Portugal, [6] Faculty of Economics, University of Porto, Porto, Portugal, [7] Organisation, Strategy, and Entrepreneurship, Maastricht University School of Business and Economics, Maastricht University, Maastricht, Netherlands, [8] Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands, [9] Department of Management, College of Business, Colorado State University, Fort Collins, CO, United States

OPEN ACCESS

Edited by:
Yair Galily,
Interdisciplinary Center Herzliya, Israel

Reviewed by:
Geoffrey Schweizer,
Heidelberg University, Germany
Asaf Blatt,
Florida State University, United States

***Correspondence:**
Jorge Sinval
jorgesinval@gmail.com

Specialty section:
This article was submitted to Movement Science and Sport Psychology, a section of the journal Frontiers in Psychology

Received: 08 April 2020
Accepted: 28 August 2020
Published: 19 October 2020

Citation:
Sinval J, Aragão e Pina J, Sinval J, Marôco J, Santos CM, Uitdewilligen S, Maynard MT and Passos AM (2020) Development of the Referee Shared Mental Models Measure (RSMMM). Front. Psychol. 11:550271. doi: 10.3389/fpsyg.2020.550271

The concept of shared mental models refers to the shared understanding among team members about how they should behave in different situations. This article aimed to develop a new shared mental model measure, specifically designed for the refereeing context. A cross-sectional study was conducted with three samples: national and regional football referees (n = 133), national football referees and assistant referees and national futsal referees (n = 277), and national futsal referees (n = 60). The proposed version of the Referee Shared Mental Models Measure (RSMMM) has 13 items that are reflected on a single factor structure. The RSMMM presented good validity evidence both based on the internal structure and based on relations to other variables (presenting positive associations with team work engagement, team adaptive performance, and team effectiveness). Such promising psychometric properties point to an optimistic outlook regarding its use to measure shared mental models in futsal and football referee teams.

Keywords: shared mental models, referees, psychometrics, football, futsal

INTRODUCTION

Shared mental models have been examined in numerous contexts (Resick et al., 2010a; Santos et al., 2015a, 2016; Tomás et al., 2017). However, one context where the role of shared mental models has received relatively little attention is sports referee teams (Filho and Tenenbaum, 2012; Aragão e Pina et al., 2018). This gap is interesting as football and futsal refereeing teams are highly interdependent in conducting their team tasks. Namely, they must coordinate several tasks before, during, and after the game (Samuel, 2015; Hancock et al., 2018); share technical and tactical knowledge to gain an adequate understanding of the task and match game needs (Mascarenhas et al., 2006; Mallo et al., 2012; McEwan and Beauchamp, 2014; Hancock et al., 2018); anticipate and adapt to the needs and actions of other members as well as changing task demands (Cannon-Bowers et al., 1993; Hancock et al., 2018); define a communication protocol to facilitate the team decision-making process (Cunningham et al., 2014; Samuel, 2015; Hancock et al., 2018) and engage in continuous learning together with the other team members (Collina, 2004;

Cunningham et al., 2014; McEwan and Beauchamp, 2014). Nevertheless, to date, only one study has addressed the entire football refereeing team (see, Boyer et al., 2015), and currently there is no shared mental model scale adapted specifically for football and futsal refereeing teams.

Shared mental models refer to an organized and common understanding among team members regarding the essential aspects of work (Klimoski and Mohammed, 1994; Mohammed et al., 2010). Team members hold multiple mental models, about different domains, while they work on a task (Cannon-Bowers et al., 1993; Klimoski and Mohammed, 1994). Cannon-Bowers et al. (1993) proposed four types of models, namely, the equipment model, task model, team interaction model, and team model. The equipment model refers to knowledge about the equipment functioning, technology, and tools with which the team members interact. The task model refers to knowledge about the task procedures, task strategies, contingency plans, and environmental constraints. The team interaction model refers to knowledge about the roles and responsibilities of team members, role interdependencies, interaction patterns, and communication channels. The team model regards knowledge about task-relevant attributes of team members, such as knowledge, skills, abilities, preferences, and tendencies (Cannon-Bowers et al., 1993; Mathieu et al., 2000; Mohammed et al., 2010).

Nevertheless, Mathieu et al. (2000) merged those four models into two domains–task mental models (comprising the equipment and task models) and team mental models (comprising the team interaction and team models). Accordingly, over the years, researchers have most commonly analyzed task mental models and team mental models (Mathieu et al., 2000; Lim and Klein, 2006; Santos and Passos, 2013). Task mental models refer to a similar understanding among team members about work objectives, team resources, task procedures and practices, and task duties. Team mental models refer to a similar understanding among team members about interpersonal interaction, team members' roles and responsibilities, and role interdependencies (Mathieu et al., 2000; Mohammed et al., 2010). In this article, this distinction between task and team mental models was made, as it has received extensive support from empirical studies (Mathieu et al., 2000; Lim and Klein, 2006; Santos and Passos, 2013; Santos et al., 2015a). Another dimension recently proposed by Randall et al. (2011) was also considered, namely, strategy mental models, which refer to "an understanding of strategic priorities, the trade-offs, and relationships among strategic alternatives, and the implications of strategic decisions" (p. 527).

Drawing on research on shared mental models in organizational teams, it is arguable that referee teams that develop shared mental models can anticipate each other's needs and adapt their behaviors to fit tightly to task demands (Cannon-Bowers et al., 1993; Mohammed et al., 2010). When the referee team members develop a common understanding on the tools and technologies they interact with, such as audio communication system, electronic flags, or video assistant referees (VARs), on the task procedures and strategies, as well as on the strategic priorities, they similarly, interpret the cues and make effective and quick decisions on the field (Kellermanns et al., 2008; Randall et al., 2011). Furthermore, when referee team members develop a shared understanding regarding each other's roles and responsibilities, as well as on the knowledge, skills, and abilities of each other, this enables them to effectively communicate and work in a coordinated way, which allows them to adapt to unexpected events, and perform their tasks successfully (Mathieu et al., 2000; Muponde and Muchemwa, 2011; Santos and Passos, 2013; Boyer et al., 2015; Diotaiuti et al., 2017; Uitdewilligen et al., 2018). Some psychometric instruments to measure mental models have already been proposed: the Shared Mental Model Scale (SMMS; Santos et al., 2015a) also with a shorter unidimensional version (Santos et al., 2015b), the Team-Related Knowledge Measurement Instrument (TeamKMI; Johnson et al., 2007), and the Perceived Mutual Understanding (PMU) scale (Burtscher and Oostlander, 2019). Yet, none of the existing instruments has been tested among referees. Therefore, by developing a shared mental model measure for referees, scholars could begin to examine shared mental models within the context of referee teams and allow further examination of their antecedents and outcomes (Aragão e Pina et al., 2021).

Research Hypotheses

One of the most desirable psychometric properties of an instrument is its dimensionality stability across different samples (Nunnally and Bernstein, 1994). As so, if an instrument maintains its dimensionality with a good fit to the different sample datasets, one can assume that the items and factors proposed are adequate to measure the desired construct/s over different groups of individuals. It is particularly important to have dimensionality evidence when independent samples of the same population are analyzed with the same instrument (Marôco, 2014). Because the perceptions of mental models can vary from context to context, it is expected that the dimensionality of the proposed measure may have a different number of factors in comparison with the instrument in which this measure was initially based (i.e., three factors). However, it is assumed that the proposed dimensionality will present a good fit to the data (H1; i.e., three different samples). It is expected that the proposed dimensionality on the first sample data will be reproduced with a good fit in the two other independent samples. If such dimensionality (i.e., factor structure) holds in all the samples, there will be promising evidence of dimensionality (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 2014).

Another important psychometric property is the reliability evidence, which can be assessed through internal consistency (Nunnally and Bernstein, 1994). Previous instruments measuring mental models reported acceptable values of internal consistency, as the PMU (Burtscher and Oostlander, 2019) with $\alpha = 0.83$ and $\omega = 0.83$. The TeamKMI reported globally satisfactory internal consistency estimate values (Johnson et al., 2007). Moreover, the SMMS reported satisfactory internal consistency values (Santos et al., 2015b). The second hypothesis (H2) presumes the Referee Shared Mental Models Measure (RSMMM) shows good evidence of the scores' reliability, more specifically in terms of internal consistency (Nunnally and Bernstein, 1994). Such estimates should be desirably high (i.e., ≥ 0.70; Iacobucci and Duhachek, 2003). Adequate internal consistency values will indicate that the items are measuring the same construct, measuring the construct consistently (McDonald, 1999).

The third hypothesis (H3) assumes that the RSMMM will present measurement invariance among referees from different sports. Such property is essential to directly compare groups within the same instrument (Davidov et al., 2014). Measurement invariance has been tested before among referees of different types of sports in a measure of self-efficacy (Myers et al., 2012), also among referees and assistant referees in football (Brandão et al., 2014) and also between elite and non-elite football referees (Johansen et al., 2018).

The extent of the relations of an instrument's scores with external variables constitutes a critical source of validity. This particular source of validity is denominated as validity evidence based on the relation to other variables (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 2014). As such, some related constructs are expected to be associated with shared mental models. Team work engagement is an affective–motivational construct that is expected to be positively related to shared mental models because higher team work engagement means higher team enthusiasm and energy (Costa et al., 2014b). The mental models construct is a cognitive one, which is expected to enhance team members' anticipation of actions and communication, conducting to positive feelings. This is also true regarding team effectiveness, because a higher common understanding of the way the team works will allow predicting behavior patterns that will likely increase the effectiveness of the team (Marks et al., 2002; Mathieu et al., 2009; DeChurch and Mesmer-Magnus, 2010). As such, a positive association between mental models and team effectiveness is expected to be observed. Associated to a higher level of shared mental models is expected to be a higher perception of team adaptive performance. Team members with shared cognitive representations regarding team function will predict the other team members move straightforwardly and consequently improving the ability to react and adjust when necessary (Pulakos et al., 2006). As so, the fourth hypothesis (H4) establishes that the suggested shared mental models measure will present validity evidence based on the relation to other variables, namely, nomological evidence in convergent terms with team work engagement, team effectiveness, and team adaptive performance.

MATERIALS AND METHODS

Sample

This article uses data from three different studies with non-probabilistic convenience samples where data were collected within a cross-sectional survey at the individual level; all participants are Portuguese football referees or assistant referees or futsal referees. Depending on the tournament, football referee teams range from three to seven members, whereas futsal referee teams range from three to four members. Team members within each team are usually the same, with some exceptions (i.e., injuries, not being considered apt in physical or written examinations). However, in the case of the top-class football national referees (i.e., C1 class), rotation between team members is more frequent.

Study I

The sample I data (n = 133) were constituted by national football referees (n = 67), with mean$_{\text{age}}$ = 30.02 (SD_{age} = 3.16) years, mean$_{\text{experience}}$ = 12.29 ($SD_{\text{experience}}$ = 3.15) years; and football regional referees (n = 66) mean$_{\text{age}}$ = 26.37 (SD_{age} = 3.21) years, mean$_{\text{experience}}$ = 8.89 ($SD_{\text{experience}}$ = 3.15). All referees completed the questionnaires.

Study II

The sample II data (n = 277) were composed by football national referees (n = 135) with mean$_{\text{experience}}$ = 12.44 ($SD_{\text{experience}}$ = 5.06) years, futsal national referees (n = 117) with mean$_{\text{experience}}$ = 11.71 ($SD_{\text{experience}}$ = 4.88) years, and football national assistant referees (n = 25) with mean$_{\text{experience}}$ = 18.44 ($SD_{\text{experience}}$ = 4.16) years. All referees completed the questionnaires.

Study III

The sample III data (n = 60) had only futsal national referees with mean$_{\text{age}}$ = 34.54 (SD_{age} = 5.52) years, mean$_{\text{experience in the current team}}$ = 2.72 ($SD_{\text{experience in the current team}}$ = 3.13) years.

Measures

All the self-report measures were collected at the individual level, reflecting the perceptions of the subject about the team.

Shared Mental Models

Shared mental models refer to a multidimensional construct. In this article, three dimensions were considered, namely, task mental models, team mental models, and strategy mental models. Referees must develop a similar understanding of the task procedures, practices, and strategies to make decisions, likely scenarios and contingencies (Cannon-Bowers et al., 1993; Mathieu et al., 2000; Aragão e Pina et al., 2019), and contingency plans (Mohammed et al., 2010). Regarding the task mental models, referee team members must develop a similar understanding about the equipment functioning and equipment limitations (Cannon-Bowers et al., 1993), as well as about the technology and tools with which they interact to make decisions (Mathieu et al., 2000; Mohammed et al., 2010). Example of such equipment are the audio communication system, the Video Assistant Referee (VAR), or the goal-line technology (GLT). Referee team members must also develop a similar understanding of the environmental constraints and the aspects of the task environment that affect team performance (Cannon-Bowers et al., 1993).

Concerning the team mental models, referee team members must develop a similar understanding about the roles and responsibilities of each team member, the role interdependencies, and about interaction patterns and communication channels and patterns (Cannon-Bowers et al., 1993; Mathieu et al., 2000; Mohammed et al., 2010). Besides, they must develop a similar understanding about the knowledge, skills, and abilities of each team member and about the team members' preferences to make decisions during the games (Cannon-Bowers et al., 1993; Mathieu et al., 2000; Aragão e Pina et al., 2019). Regarding the strategy mental models, referee team members must develop a similar understanding of the strategic priorities, as well as the implications of strategic decisions (Randall et al., 2011).

This measure was named as RSMMM (**Table 1**). Based on relevant literature on shared mental models, namely, in other instruments (Santos et al., 2015a,b), an initial pool of 13 items was developed across the three dimensions: task (e.g., "In my team, the team members have a similar understanding about the technology and tools needed to make decisions during a game"); team (e.g., "In my team, the team members have a similar understanding about the knowledge, skills, and abilities of each other"); and strategy (e.g., "In my team, the team members have a similar understanding about the strategic priorities of the game"). It was ensured that the shared mental models' items, in particular, the items of the task dimension, were specific to the context of referee teams (Cannon-Bowers and Salas, 2001) by stating that team members have the knowledge needed to make decisions during a game or by providing examples related to the referees' responsibilities. For instance, "In my team, the team members have a similar understanding about resources needed to make decisions during a game" and "In my team, the team members have a similar understanding about the tasks each team member has to do (e.g., train during the week, prepare the game properly, employ an exemplary behavior, make a difficult decision)." Each item was scored on a seven-point Likert scale (1 = "Totally disagree", 2 = "Strongly disagree", 3 = "Disagree", 4 = "Neither agree, nor disagree", 5 = "Agree", 6 = "Strongly agree", 7 = "Totally agree").

Team Work Engagement

Team work engagement is defined as an emergent state that develops from team members' interactions and that cannot be found in individuals being exclusive to teams (Costa et al., 2016). To measure team work engagement, the Team Work Engagement Scale was used (Costa et al., 2014a). This instrument consists of nine items measured in a seven-point Likert scale (1 = "Totally disagree", 7 = "Totally agree"). Team Work Engagement is seen as a second-order factor (as the individual measure; Sinval et al., 2018b,a) that comprises three first-order dimensions (i.e., vigor, dedication, and absorption). This instrument showed good validity evidence based on the internal structure in previous studies, namely, in terms of reliability, having Cronbach's α of 0.85 a 0.97 for the vigor factor, 0.88 and 0.95 for the dedication factor; and 0.83 and 0.95 for the absorption factor (Costa et al., 2014a). Examples of items are as follows: "At our work, we feel bursting with energy" (vigor), "We are enthusiastic about our job" (dedication), and "We feel happy when we are working intensely" (absorption).

Team Adaptive Performance

Team adaptive performance is defined as an emergent state that occurs as a consequence of the adaptation process, in which individuals and teams cope with the demands of the context (Maynard et al., 2015). The Team Adaptive Performance Scale was used to measure team adaptive performance (Marques-Quinteiro et al., 2015). This instrument has eight items that were answered using a seven-point Likert scale (1 = "Totally disagree", 7 = "Totally agree"). This instrument assumes that team adaptive performance is a second-order latent variable with two first-order latent factors (factor I: problem-solving–oriented factor, six items; and factor II: learning work tasks, technologies, and procedures factor, two items). Examples of items are as follows: "We use creative ideas to manage incoming events" (problem-solving–oriented), and "We remain calm and behave positively under highly stressful events" (learning work tasks, technologies, and procedures).

Team Effectiveness

Team effectiveness is conceived in three criteria: team performance, quality of group experience, and team viability (Aubé and Rousseau, 2005). Team performance has been seen in the function of the assigned team goals (Hackman, 1987). The quality of group experience is defined as the positiveness of the social climate in the team (McGrath, 1991). The team viability consists in the capacity of the team to adapt to external and internal changes and also to the likelihood of team members continuing to work together (Hackman, 1987). The team effectiveness dimension was measured using the Portuguese version of the Scale of Effectiveness of Teams (3Es; Vicente et al., 2014). This instrument has three first-order factors (team performance, quality of group experience, and team viability), which are explained by a hierarchical structure (second-order factor) called effectiveness. The items were scored with a Likert scale from 1 = "Totally disagree" to 7 = "Totally agree". In the original version with the Canadian sample (Aubé and Rousseau, 2005), the authors studied the internal consistency, and good Cronbach's α values were evidenced ($\alpha_{\text{team performance}} = 0.82$, $\alpha_{\text{team viability}} = 0.84$, $\alpha_{\text{quality of group experience}} = 0.96$). Examples of items are as follows: "The members of this team attain their assigned performance goals" (team performance); "The social climate in our work team is good" (quality of group experience); and "Team members adjust to the changes that happen in their work environment" (team viability).

Procedure

For samples I and II studies, the institutional review board, and the National Referees' Committee approved the study. National referees were attending a seminar, and regional referees were attending a promotion seminar compulsory for those wishing to be considered for promotion to the national level. Data were collected at the beginning of each seminar, after providing a brief explanation of the nature of the investigation. The institutional approval of the Portuguese Football Federation was obtained for sample III's study. All referees participated voluntarily, and written or electronic informed consent was obtained from all participants, and confidentiality for their responses was ensured.

Data Analysis

All statistical analyses were performed with *R* (R Core Team, 2020) through RStudio (RStudio Team, 2020). The descriptive statistics were obtained with the *skimr* package (McNamara et al., 2018); the coefficient of variation (*CV*) was calculated through the *sjstats* package (Lüdecke, 2019), and the standard error of the mean (*SEM*) was estimated by the *plotrix* package

TABLE 1 | Referee Shared Mental Models Measure (RSMMM) items.

Item	English version of RSMMM							Portuguese (Portugal) version of RSMMM						
	Totally disagree	**Strongly disagree**	**Disagree**	**Neither disagree nor agree**	**Agree**	**Strongly agree**	**Totally agree**	**Discordo totalmente**	**Discordo muito**	**Discordo em parte**	**Nem concordo, nem discordo**	**Concordo em parte**	**Concordo muito**	**Concordo totalmente**
	1	**2**	**3**	**4**	**5**	**6**	**7**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
	Shared Mental Models							**Modelos Mentais Partilhados**						
1	In my team, members have a similar understanding of the resources that are needed to make decisions during a game.							Na minha equipa, os membros têm um entendimento semelhante sobre os recursos que são necessários para tomar as decisões durante um jogo.						
2	In my team, members have a similar understanding of the technology and tools needed to make decisions during a game.							Na minha equipa, os membros têm um entendimento semelhante sobre a tecnologia e as ferramentas necessárias para tomar as decisões durante um jogo.						
3	In my team, members have a similar understanding of the procedures and practices needed to make decisions during a game.							Na minha equipa, os membros têm um entendimento semelhante sobre os procedimentos e práticas necessários para tomar as decisões durante um jogo.						
4	In my team, even when we are confronted with incidents or problems related to our performance, we have a similar understanding of how to perform our tasks.							Na minha equipa, mesmo quando somos confrontados com incidentes ou problemas relacionados com a nossa atuação, temos um entendimento semelhante sobre como realizar as nossas tarefas.						
5	In my team, members have a similar understanding of what they must do (e.g., train during the week, properly prepare the game, adopt exemplary behavior, make a difficult decision).							Na minha equipa, os membros têm um entendimento semelhante em relação ao que cada um tem que fazer (ex.: treinar durante a semana, preparar adequadamente o jogo, adotar um comportamento exemplar, tomar uma decisão difícil).						
6	In my team, members have a similar understanding of how their roles are related.							Na minha equipa, os membros têm um entendimento semelhante sobre a forma como os papéis de cada um estão relacionados.						
7	In my team, members have a similar understanding of how to interact with each other.							Na minha equipa, os membros têm um entendimento semelhante sobre a forma como interagir uns com os outros.						
8	In my team, members have a similar understanding of what the best methods are for communicating with each other.							Na minha equipa, os membros têm um entendimento semelhante sobre quais os melhores métodos para comunicar uns com os outros.						
9	In my team, members have a similar understanding of each other's knowledge, skills and abilities.							Na minha equipa, os membros têm um entendimento semelhante em relação aos conhecimentos, competências e capacidades de cada um.						
10	In my team, members have a similar understanding of each other's preferences, which are relevant to making decisions during a game.							Na minha equipa, os membros têm um entendimento semelhante em relação às preferências de cada um, que são relevantes para tomar as decisões durante um jogo.						
11	In my team, members have a similar understanding of the game's strategic priorities.							Na minha equipa, os membros têm um entendimento semelhante em relação às prioridades estratégicas do jogo.						
12	In my team, members have a similar understanding of the implications of the strategic decisions that are made.							Na minha equipa, os membros têm um entendimento semelhante em relação às implicações das decisões estratégicas que são tomadas.						
13	In my team, members have a similar understanding of which aspects of the game are most important to team performance.							Na minha equipa, os membros têm um entendimento semelhante sobre quais os aspetos do jogo que são mais importantes para o desempenho da equipa.						

(Lemon, 2006). The mode was calculated with the *DescTools* package (Signorell et al., 2019). Severe univariate normality violations were considered for absolute values of $sk > 3$ and $ku > 7$ (Finney and DiStefano, 2013; Marôco, 2014).

Regarding the exploratory factor analysis (EFA), the Kaiser–Meyer–Olkin (*KMO*) coefficient was used as a measure of sampling adequacy (Kaiser and Rice, 1974). The Bartlett test (Bartlett, 1951) was chosen to test if the correlation matrix was factorable (i.e., the correlations differ from 0) (Revelle, 2019). *KMO* values > 0.8 and Bartlett test significance ≤ 0.05, indicating adequate sampling (Marôco, 2018). The number of factors was determined through the comparison data (CD) approach, as suggested by Ruscio and Roche (2012), which stated that this technique outperforms Parallel Analysis. CD is a variant of Parallel Analysis that reproduces the correlation matrix rather than generating random data (Courtney, 2012). The extraction of the factors was performed using the principal components analysis with a weighted least-squares factoring method on the *polychoric* correlation (ρ_{PC}) matrix with oblimin rotation and weighted least-squares factoring. The cutoff for items' loadings was 0.40. The CD analysis was conducted using the *RGenData* package (Ruscio, 2018). The Bartlett test, the KMO coefficient, factors' extraction and the ρ_{PC} were produced using the *psych* package (Revelle, 2019). As goodness-of-fit index for the EFA, the RMSR (root mean square of the residual) was used.

Confirmatory factor analysis (CFA) was conducted with the *lavaan* package (Rosseel, 2012) using the weighted least-squares means and variances (WLSMV) estimation method for ordinal variables (Muthén, 1983). As goodness-of-fit indices, the *TLI* (Tucker–Lewis index), *NFI* (normed fit index), χ^2/df (ratio chi-square and degrees of freedom), *CFI* (comparative fit index), the *RMSEA*, and the *SRMR* (standardized root mean square residual) were used. For values of $\chi^2/\text{df} < 5$, values of *CFI, NFI*, and *TLI* > 0.95; values of *SRMR* < 0.08; and *RMSEA* < 0.08, the fit of the model was considered good (Hoyle, 1995; Boomsma, 2000; McDonald and Ho, 2002; Byrne, 2010; Marôco, 2014).

To analyze the convergent validity evidence, the average variance extracted (*AVE*) was estimated (Fornell and Larcker, 1981). For values of $AVE \geq 0.5$ (Hair et al., 2019), adequate convergent validity evidence was assumed.

The discriminant validity evidence was tested to verify whether the items that represent a dimension were strongly correlated with other dimensions. To assess such evidence, the Fornell and Larcker's (1981) approach was used: for two factors, x and y, if AVE_x and $AVE_y \geq \rho^2_{xy}$ (squared correlation between the factors x and y), adequate discriminant validity evidence is assumed.

The reliability of the scores was assessed with estimates of internal consistency, α (Cronbach, 1951), and ω (Raykov, 2001), using the *semTools* package (Jorgensen et al., 2019), where higher values were indicative of better internal consistency results. The α coefficient was calculated using the polychoric correlation matrix. The second-order reliability estimates were as follows: the proportion of the second-order factor explaining the total score (ω_{L1}), the proportion of variance explained by second-order factor after partialing the uniqueness of the first-order factor ($\omega_{partialL1}$), and the variance of the first-order factors explained by the second-order factor (ω_{L2}). Such reliability estimates were obtained with the *semTools* package (Jorgensen et al., 2019). The confidence intervals (CIs) for the internal consistency estimates were obtained through the *userfriendlyscience* package (Peters, 2018) and the *boot* package (Davison and Hinkley, 1997; Canty and Ripley, 2020) using 1,000 bootstrap replicates. The bias-corrected and accelerated method was used, which tend to provide better coverage in non-normal sampling distributions (Efron and Tibshirani, 1994; Carpenter and Bithell, 2000).

The measurement invariance was assessed and verified using the *lavaan* package (Rosseel, 2012) and the *semTools* package (Jorgensen et al., 2019). A group of five models was compared: (a) configural invariance; (b) first-order factor loadings; (c) thresholds/intercepts of measured variables (depending on if the items are considered or not as categorical); (d) residual variances of observed variables; and (e) latent means. The latent variable means were compared, and Cohen d was used as the effect size (Cohen, 1988).

RESULTS

The presented results refer to three different studies with three different samples. First, the three samples were merged, and the instrument's expected dimensionality analyzed. Subsequently, the samples were individually analyzed to obtain different validity evidence from each of them.

Merge Samples

Validity Evidence Based on the Internal Structure

The dimensionality, reliability of scores, and measurement invariance of the instrument will be tested to verify the robustness of this source of validity evidence.

Items' distributional properties

As **Table 2** shows, none of the items for samples I and II presented severe problems of univariate normality because all of them presented $|sk| < 3$ and $|ku| < 7$ (Finney and DiStefano, 2013; Marôco, 2014). However, some of sample III items' absolute values of *ku* were greater than 7 (i.e., items 1, 2, 3, 6, and 9; **Table 2**). Item 5 was the one that presented more variability (i.e., *CV*) in the answers in all samples.

Following the recommendations of Finney et al. (2016) with categorical items with six or more points, both maximum likelihood estimation with robust (Huber–White) standard errors (MLR) and diagonal weighted least-squares methods (as the WLSM estimator) can be used. The WLSMV estimator was chosen because it does not require multivariate normality as an assumption. To analyze the validity evidence based on the internal structure of the new measure, several steps were carried (i.e., dimensionality, reliability, and measurement invariance).

Dimensionality

To test the expected three first-order factors of the RSMMM, a CFA was conducted with all the available data from the three collected samples. The CFA is the most appropriate technique to use when there is a definite theory regarding the latent factors and their relationships to the indicators, that is, dimensionality (Brown, 2015; Finch and French, 2015). Items 1 to 4 were used as indicators of the task factor; items 5 to 10 were expected to

TABLE 2 | Items' distributional properties.

Study I items' descriptive statistics (*n* = 133)

Item	*M*	*SD*	*Min*	*Mdn*	*Max*	Histogram	*Mode*	*SEM*	*CV*	*sk*	*ku*
Item 1	5.71	0.93	2	6	7		6.00	0.08	0.16	−1.52	3.07
Item 2	5.85	0.82	3	6	7		6.00	0.07	0.14	−0.70	0.59
Item 3	5.89	0.94	2	6	7		6.00	0.08	0.16	−1.42	2.98
Item 4	5.67	0.92	2	6	7		6.00	0.08	0.16	−0.94	1.76
Item 5	5.43	1.27	1	6	7		6.00	0.11	0.23	−1.03	0.87
Item 6	5.93	0.84	2	6	7		6.00	0.07	0.14	−1.18	3.35
Item 7	6.02	0.84	3	6	7		6.00	0.07	0.14	−0.88	1.23
Item 8	5.73	0.86	3	6	7		6.00	0.07	0.15	−0.65	0.80
Item 9	5.71	0.96	2	6	7		6.00	0.08	0.17	−1.10	2.22
Item 10	5.59	0.93	2	6	7		6.00	0.08	0.17	−0.89	1.35
Item 11	5.71	0.89	2	6	7		6.00	0.08	0.16	−0.91	1.89
Item 12	5.70	0.94	1	6	7		6.00	0.08	0.16	−1.35	4.07
Item 13	5.95	0.80	3	6	7		6.00	0.07	0.13	−0.72	0.85

Study II items' descriptive statistics (*n* = 277)

Item	*M*	*SD*	*Min*	*Mdn*	*Max*	Histogram	*Mode*	*SEM*	*CV*	*sk*	*ku*
Item 1	6.02	0.87	2	6	7		6.00	0.05	0.14	−1.32	3.04
Item 2	6.03	0.88	2	6	7		6.00	0.05	0.15	−1.34	3.22
Item 3	6.10	0.81	2	6	7		6.00	0.05	0.13	−1.51	4.39
Item 4	5.93	0.83	3	6	7		6.00	0.05	0.14	−0.97	1.67
Item 5	5.81	1.14	1	6	7		6.00	0.07	0.20	−1.54	3.09
Item 6	6.09	0.86	2	6	7		6.00	0.05	0.14	−1.31	2.95
Item 7	6.12	0.92	1	6	7		6.00	0.06	0.15	−1.77	5.40
Item 8	6.03	0.93	2	6	7		6.00	0.06	0.15	−1.47	3.66
Item 9	6.08	0.90	2	6	7		6.00	0.05	0.15	−1.50	3.46
Item 10	5.98	0.90	2	6	7		6.00	0.05	0.15	−1.13	2.08
Item 11	6.04	0.85	2	6	7		6.00	0.05	0.14	−1.43	4.20
Item 12	6.01	0.88	2	6	7		6.00	0.05	0.15	−1.46	3.82
Item 13	6.20	0.86	1	6	7		6.00	0.05	0.14	−1.77	6.38

Study III items' descriptive statistics (*n* = 60)

Item	*M*	*SD*	*Min*	*Mdn*	*Max*	Histogram	*Mode*	*SEM*	*CV*	*sk*	*ku*
Item 1	6.10	1.16	1	6	7		6.00	0.15	0.19	−2.74	9.41
Item 2	6.26	1.09	1	6	7		7.00	0.14	0.17	−2.85	10.01
Item 3	6.02	1.07	1	6	7		6.00	0.14	0.18	−2.43	8.42
Item 4	5.67	1.19	1	6	7		6.00	0.15	0.21	−1.85	5.25
Item 5	5.61	1.31	1	6	7		6.00	0.17	0.23	−1.42	2.85
Item 6	6.13	1.02	1	6	7		6.00	0.13	0.17	−2.27	8.31
Item 7	5.92	1.22	1	6	7		6.00	0.16	0.21	−2.20	6.34
Item 8	5.84	1.16	1	6	7		6.00	0.15	0.20	−1.78	4.63
Item 9	6.05	1.22	1	6	7		6.00	0.16	0.20	−2.38	7.19
Item 10	5.62	1.30	1	6	7		6.00	0.17	0.23	−2.04	5.19
Item 11	5.70	1.23	1	6	7		6.00	0.16	0.21	−1.87	4.82
Item 12	5.75	1.30	1	6	7		6.00	0.17	0.23	−1.74	3.82
Item 13	5.83	1.17	1	6	7		6.00	0.15	0.20	−1.76	4.49

be indicators of the dimension team, and items 11 to 13 were developed as potential indicators of the latent variable strategy.

The goodness-of-fit indices were indicative of good fit to the data ($\chi^2(62) = 184.686$, $n = 526$, $\chi^2/\mathrm{df} = 2.979$, $CFI = 0.999$, $NFI = 0.998$, $TLI = 0.998$, $SRMR = 0.031$, $RMSEA = 0.061$, $P(\mathrm{rmsea} \leq 0.05) = 0.032$, 90% CI]0.051; 0.072[). The convergent validity evidence based on the internal structure was good ($AVE_{\mathrm{task}} = 0.78$, $AVE_{\mathrm{team}} = 0.70$, $AVE_{\mathrm{strategy}} = 0.85$). However, the discriminant validity evidence based on the internal structure was not satisfactory, because the latent correlations between the factors were too high ($r_{\mathrm{task} \times \mathrm{team}} = 0.919$, $p < 0.001$; $r_{\mathrm{task} \times \mathrm{strategy}} = 0.870$, $p < 0.001$; $r_{\mathrm{team} \times \mathrm{strategy}} = 0.915$, $p < 0.001$). Comparing the values of the AVE of each pair of factors with their squared correlation value, only one

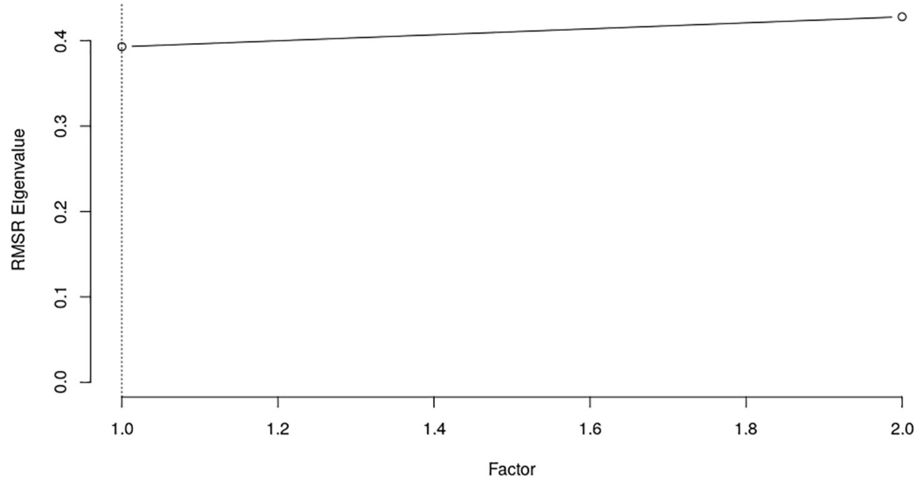

FIGURE 1 | Fit to comparison data (n = 133). The CD analysis suggested that the number of factors to retain is one.

of the three pairs (task and strategy) showed evidence of discriminant validity. The $r^2_{\text{task} \times \text{team}}$ = 0.845 was greater than AVE_{task} = 0.78 and AVE_{team} = 0.70; the $r^2_{\text{task} \times \text{strategy}}$ = 0.757 was smaller than AVE_{task} = 0.78 and AVE_{strategy} = 0.85; and $r^2_{\text{team} \times \text{strategy}}$ = 0.838 was greater than AVE_{team} = 0.70, but smaller than AVE_{strategy} = 0.85. Such finding might be indicative of a unidimensional model, which should be investigated through the appropriate analysis (i.e., EFA).

Reliability of the scores: Internal consistency

The merged data of the three different studies revealed good reliability evidence in terms of internal consistency (α_{task} = 0.93, 95% CI]0.91; 0.94[; ω_{task} = 0.87, 95% CI]0.82; 0.90[; α_{team} = 0.93, 95% CI]0.91; 0.94[; ω_{team} = 0.90, 95% CI]0.88; 0.91[; α_{strategy} = 0.94, 95% CI]0.92; 0.95[; ω_{strategy} = 0.90, 95% CI]0.88; 0.92[).

Because the content explained by the three different factors is similar, the dimensionality was investigated using an exploratory approach (EFA), where the EFA's suggested dimensionality from sample I was then tested (through CFA) in samples II and III's data.

When the empirical evidence lacks regarding the construct expected dimensionality, EFA might be most appropriate than CFA (Finch and French, 2015). The EFA attributes a small burden on the researcher concerning the latent factors and their relationships to the indicators, making possible establishing an interval of the number of factors that can emerge from the indicators (Marôco, 2018).

Study I

Validity Evidence Based on the Internal Structure

Dimensionality

Data obtained from study I met the *KMO* coefficient (0.900) and Bartlett test of sphericity ($\chi^2(78)$ = 963.521; p < 0.001). The CD suggested that the best solution contains only one factor (**Figure 1**).

The one-factor solution was adopted, and the results of the correspondent EFA (**Table 3**) revealed 50.8% of explained variance (*RMSR* = 0.086).

Reliability of the scores: Internal consistency

The study I's data revealed good reliability evidence in terms of internal consistency (α = 0.93, 95% CI]0.91; 0.95[; ω = 0.93, 95% CI]0.91; 0.95[).

To test the proposed structure observed in study I's sample and see if it was adequate for a second and third independent samples from the population, a CFA was also performed on study II and study III samples.

Sample II

As previously mentioned, the use of CFA demands strong theoretical and/or empirical evidence regarding the dimensionality of a psychometric instrument. As such, because study I's sample provided empirical evidence supporting the one-factor solution, the CFA will be used to investigate the RSMMM single-factor model (Finch and French, 2015).

Validity Evidence Based on the Internal Structure

Dimensionality

The goodness-of-fit indices were indicative of an acceptable fit of study II's data to the model (**Figure 2**; $\chi^2(65)$ = 271.199, n = 277, χ^2/df = 4.172, *CFI* = 0.993, *NFI* = 0.991, *TLI* = 0.992, *SRMR* = 0.054, *RMSEA* = 0.107, $P(\text{rmsea} \leq 0.05) < 0.001$, 90% CI]0.094; 0.121[). In terms of convergent validity based on the internal structure, the estimate of *AVE* was good (*AVE* = 0.67).

Reliability of the scores: Internal consistency

Regarding study II's internal consistency, the obtained values revealed good validity evidence in terms of reliability (α = 0.96, 95% CI]0.95; 0.97[; ω = 0.92, 95% CI]0.89; 0.93[). Both the α and ω coefficients were indicative of good evidence in terms of the reliability of the scores.

TABLE 3 | Exploratory factor analysis loadings and total of explained variance.

Items	Factor 1
Item 1	0.826
Item 2	0.779
Item 3	0.762
Item 4	0.631
Item 5	0.440
Item 6	0.697
Item 7	0.731
Item 8	0.805
Item 9	0.677
Item 10	0.685
Item 11	0.721
Item 12	0.724
Item 13	0.713
Total of variance	0.508

Measurement invariance

Measurement invariance between sports refereed (i.e., football and futsal) was tested using study II's sample. Because there were only 25 football assistant referees, the measurement invariance analysis was performed only with the futsal and football referees. To conduct the measurement invariance considering the ordinal nature of the items, it is required that the items in both groups have the same number of thresholds. Because both groups had a different number of thresholds for some items, it was not possible to use WLSMV. As so, the measurement invariance analysis was performed using the MLR estimator because this method has been shown to work well with categorical data with no severe deviations from the normal distribution (Rhemtulla et al., 2012). As **Table 4** shows, full uniqueness measurement invariance was achieved both by the ΔCFI and $\Delta\chi^2$ criteria (Satorra and Bentler, 2001; Cheung and Rensvold, 2002), which allows establishing comparisons between the shared mental models latent scores among the football and futsal referees.

The shared mental models' latent means presented significant differences among the futsal and football referees ($\Delta\chi^2(1) = 5.495$, $p = 0.019$, $d = 0.168$), with the football referees perceiving higher levels of shared mental models than their futsal counterparts.

Sample III

Validity Evidence Based on the Internal Structure

Dimensionality

Study III's CFA revealed an excellent fit to the data (**Figure 3**; $\chi^2(65) = 74.428$, $n = 60$, $\chi^2/\text{df} = 1.145$, $CFI = 0.999$, $NFI = 0.994$, $TLI = 0.999$, $SRMR = 0.059$, $RMSEA = 0.050$, $P(\text{rmsea} \leq 0.05) = 0.482$, 90% CI]0.000; 0.095[). The convergent validity evidence in terms of the internal structure was particularly good ($AVE = 0.74$).

Reliability of the scores: Internal consistency

The internal consistency estimates of study III's sample were like study I's and study II's ones ($\alpha = 0.98$, 95% CI]0.93; 0.99[; $\omega = 0.98$, 95% CI]0.93; 0.99[) and as so were indicative of good evidence in terms of the reliability of the scores.

Validity Evidence Based on the Relation With Other Variables

The validity evidence based on the relation to other variables was investigated using study III's sample. The nomological evidence was verified (i.e., convergent validity with team adaptive performance, team work engagement, and team effectiveness). For such analysis, the structural equation model framework was used.

Measurement model

Because the sample size (i.e., $N_{\text{studyIII}} = 60$) was too small to be used with the WLSMV estimator on this model, the MLR estimator was used in all subsequent analyses. The measurement model of the team work engagement measure revealed an acceptable fit to the data ($\chi^2(25) = 69.366$, $n = 54$, $\chi^2/\text{df} = 2.775$, $CFI = 0.932$, $NFI = 0.900$, $TLI = 0.902$, $SRMR = 0.045$, $RMSEA = 0.181$, $P(\text{rmsea} \leq 0.05) < 0.001$, 90% CI]0.131; 0.233[) after constraining the error variance of the first-order factor (dedication) to 0.01 in order to avoid negative variance. The second-order reliability estimates were good ($\omega_{L1} = 0.95$, 95% CI]0.87; 0.99[; $\omega_{\text{partial}L1} = 0.97$, 95% CI]0.92; 0.99[; $\omega_{L2} = 0.98$, 95% CI]0.93; 1.00[).

The hierarchical model of the team adaptive performance scale revealed an acceptable fit to the data ($\chi^2(19) = 44.166$, $p = 0.001$, $n = 54$; $\chi^2/\text{df} = 2.325$; $CFI = 0.957$; $TLI = 0.936$; $NFI = 0.928$; $SRMR = 0.034$; $RMSEA = 0.157$; $P(\text{rmsea} \leq 0.05) = 0.004$, 90% CI]0.096; 0.217[) after adding one correlation path between item 5's and item 6's residuals ($r = 0.636$; $p = 0.024$). The variance of the problem-solving–oriented factor was constrained to 0.01 to avoid negative variances. The structural weights (γ) of the two factors were constrained to be equal, to solve the model identification problem of two first-order factors in a hierarchical model. The second-order reliability estimates were good ($\omega_{L1} = 0.96$, 95% CI]0.87; 0.98[; $\omega_{\text{partial}L1} = 0.97$, 95% CI]0.91; 0.99[; $\omega_{L2} = 0.96$, 95% CI]0.85; 0.99[).

The team effectiveness second-order model had a good fit to the data ($\chi^2(32) = 58.072$, $p = 0.003$, $n = 57$; $\chi^2/\text{df} = 1.815$; $CFI = 0.964$; $TLI = 0.950$; $NFI = 0.925$; $SRMR = 0.044$; $RMSEA = 0.120$; $P(\text{rmsea} \leq 0.05) = 0.018$, 90% CI]0.068; 0.168[). The second-order reliability estimates were good ($\omega_{L1} = 0.92$, 95% CI]0.69; 0.97]; $\omega_{\text{partial}L1} = 0.95$, 95% CI]0.84; 0.99[; $\omega_{L2} = 0.97$, 95% CI]0.78; 0.98[).

Structural model

Because the used psychometric instruments (i.e., measurement model) presented good validity evidence based on the internal structure, a full structural model was tested for each of the related measures (i.e., team work engagement, team adaptive performance, and team effectiveness). The structural model that related team work engagement and shared mental models revealed an acceptable fit to the data ($\chi^2(206) = 383.091$, $p < 0.001$, $n = 54$; $\chi^2/\text{df} = 1.86$; $CFI = 0.900$; $TLI = 0.888$; $NFI = 0.809$; $SRMR = 0.060$; $RMSEA = 0.126$; $P(\text{rmsea} \leq 0.05) < 0.001$, 90% CI]0.106; 0.146[) with a strong and positive latent

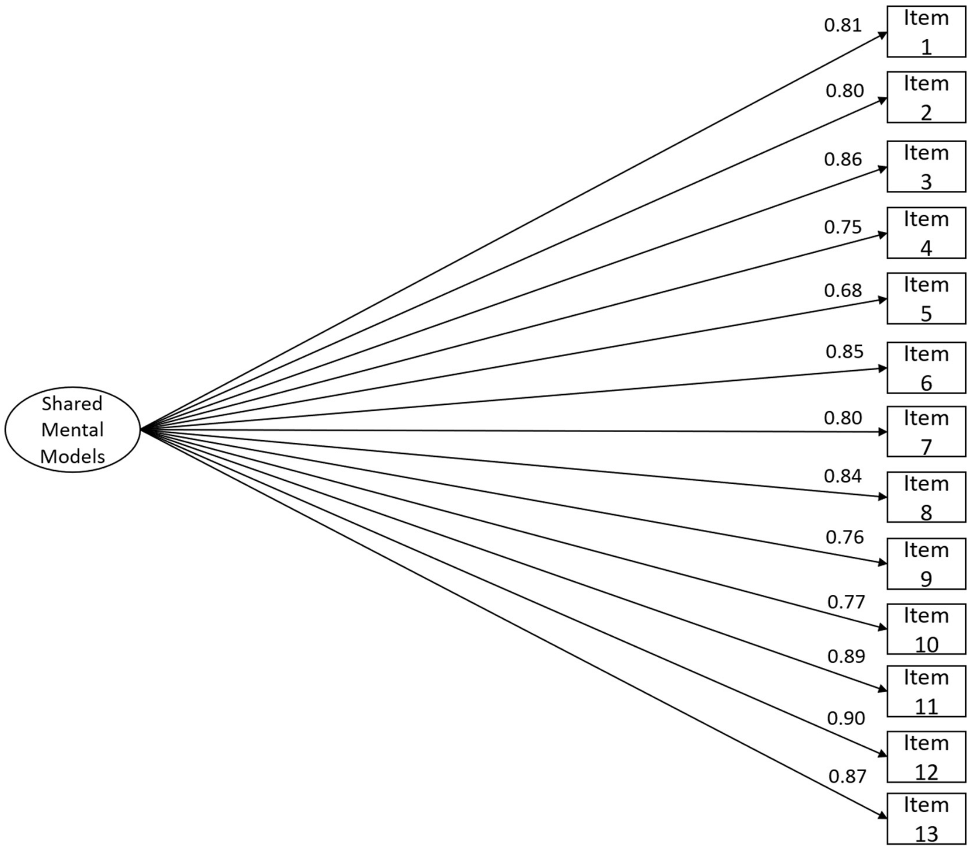

FIGURE 2 | RSMMM one-factor version (13-item) structure fit using study II's sample (n = 277). Factor loadings for each item are shown: $\chi^2(65) = 271.199$, n = 277, χ^2/df = 4.172, *CFI* = 0.993, *NFI* = 0.991, *TLI* = 0.992, *SRMR* = 0.054, *RMSEA* = 0.107, $P(\text{rmsea} \leq 0.05) < 0.001$, 90% CI]0.094; 0.121[.

correlation (H4; $r_{\text{team work engagement} \times \text{shared mental models}} = 0.764$; $p = 0.099$). The raw correlation between the arithmetic mean of the RSMMM's items and the team work engagement's items was strong and positive ($r = 0.776$; $p < 0.001$). The structural model that correlated shared mental models with team adaptive performance presented an acceptable fit to the data ($\chi^2(185) = 329.269$, $p < 0.001$, $n = 54$; χ^2/df = 1.780; *CFI* = 0.916; *TLI* = 0.905; *NFI* = 0.829; *SRMR* = 0.038; *RMSEA* = 0.120; $P(\text{rmsea} \leq 0.05) < 0.001$, 90% CI]0.099; 0.141[) showing a positive strong latent correlation (H4; $r_{\text{adaptive performance} \times \text{shared mental models}} = 0.910$; $p = 0.053$). The raw correlation was also strong and positive between the arithmetic mean of the RSMMM's items and the team adaptive performance's items ($r = 0.888$; $p < 0.001$). Finally, the model that correlated team effectiveness with shared mental models showed an acceptable fit to the data ($\chi^2(226) = 423.832$, $p < 0.001$, $n = 54$; χ^2/df = 1.875; *CFI* = 0.895; *TLI* = 0.882; *NFI* = 0.801; *SRMR* = 0.050; *RMSEA* = 0.124; $P(\text{rmsea} \leq 0.05) < 0.001$, 90% CI]0.106; 0.142[) revealing positive and strong latent correlation (H4; $r_{\text{effectiveness} \times \text{shared mental models}} = 0.909$; $p = 0.068$). The raw correlation between the arithmetic mean of the team effectiveness' items and the RSMMM's items was strong and positive ($r = 0.857$; $p < 0.001$).

TABLE 4 | Measurement invariance analysis among futsal and football referees (study II's sample).

Model	χ^2	*df*	χ^2/df	CFI_{robust}	$\Delta\chi^2$	ΔCFI_{robust}
Sport refereed						
Configural	436.77	130	3.36	0.888	–	–
Metric	442.68	142	3.12	0.893	4.071^{ns}	0.005
Scalar	455.31	154	2.96	0.892	13.485^{ns}	−0.001
Full uniqueness	495.30	167	2.97	0.888	16.641^{ns}	−0.004
Latent means	499.45	168	2.97	0.887	5.495*	−0.002

$^{ns}p < 0.05$; $^{*}p \leq 0.05$.

Such correlation values suggest acceptable nomological evidence–particularly in terms of convergent validity evidence–in relation to the team work engagement scores. However, the correlation values between the shared mental models' scores and the team adaptative performance and the team effectiveness seem too high (constructs overlap), providing poor convergent validity evidence.

Some of the presented models had mediocre *RMSEA* values. However, *RMSEA* point estimates depend on sample size, model degrees of freedom, and model misspecification (MacCallum et al., 1996; Chen et al., 2008). To assess the model's fit to the

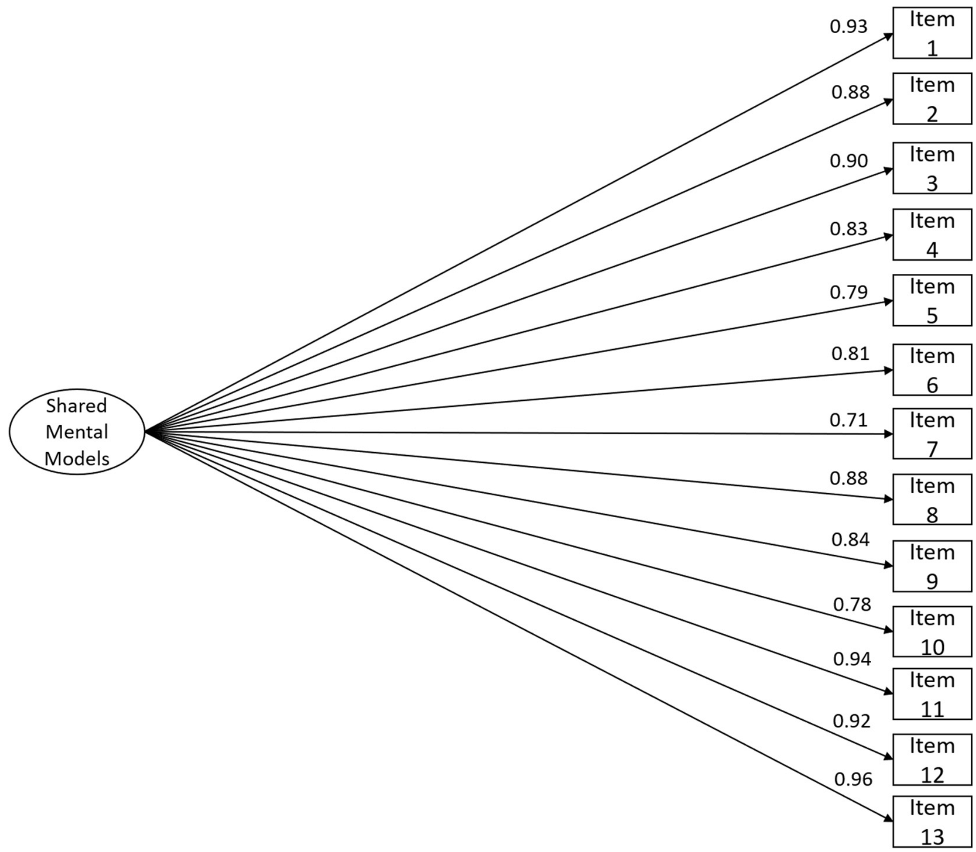

FIGURE 3 | RSMMM one-factor version (13-item) structure fit using study III's sample (n = 60). Factor loadings for each item are shown: $\chi^2(65) = 74.428$, n = 60, $\chi^2/df = 1.145$, *CFI* = 0.999, *NFI* = 0.994, *TLI* = 0.999, *SRMR* = 0.059, *RMSEA* = 0.050, *P*(rmsea $\leq$ 0.05) = 0.482, 90% CI]0.000; 0.095[.

data, other goodness-of-fit indices were presented in conjunction, namely, *SRMR*, which showed acceptable to good estimates. The *SRMR* goodness-of-fit index seems to be more robust than *RMSEA* across all conditions (Maydeu-Olivares et al., 2018).

DISCUSSION

There is a need to more fully examine the team dynamics present within professional sport referee teams (Aragão e Pina et al., 2018). In particular, given that shared mental models have been shown to play an important role in shaping team dynamics and performance in other context (Marks et al., 2000; Mathieu et al., 2000; Mascarenhas et al., 2005), there is a need to investigate the impact of referee teams shared mental models on team functioning and adaptability. However, the shared mental model literature suggests that one needs to adapt the measurement of such cognitive structures to the context within which such teams operate. As such, the primary aim of the current study was to develop a measure of shared mental models within the context of professional football and futsal referee teams.

The proposed new measure revealed good psychometric properties. Namely, the shared mental model measure developed in this article presented good validity evidence across the three different samples of football and/or futsal referees presented here. The RSMMM showed promising validity evidence both based on the internal structure and based on the relation with team work engagement (i.e., nomological evidence). Nomological evidence approaches the network of relations between the constructs. As so, the observed correlations between the latent variables (shared mental models and team work engagement) were aligned with the claims of the literature. Such findings suggest a useful unidimensional measure both for futsal and football referees.

The initial model (three first-order factors) revealed a lack of discriminant validity in terms of internal structure (Fornell and Larcker, 1981), indicating that the content explained by the three factors is similar. The dimensionality that emerged from the CD revealed that the referees on the sample perceive shared mental models as a unidimensional structure. Previous studies in which the RSMMM was based conceptualized it as a three-factor model (Santos et al., 2015a) or as a unidimensional one (Santos et al., 2015b). However, such solutions were not necessarily expected to be found in the referee context. Both the three-factor dimensionality of the SMMS (Santos et al., 2015a) and the unidimensional four-items version (Santos et al., 2015b) were proposed using a sample of teams from diverse contexts that participated on a virtual management challenge. Researchers have identified different dimensions of shared mental models including task, team, and strategy, as these are key aspects of the team work environment (e.g., Mohammed et al., 2010;

Resick et al., 2010b). However, as Mohammed et al. (2000) state, "although the domain of a team model can vary (e.g., individual task work, team task work, team work), it should be viewed as reflecting how team members conceptualize a team-relevant phenomenon" (p. 125). Our study supports Mohammed et al. (2000) argument as our findings consistently suggest (over three different samples) that football and futsal referees have a general understanding of the relevant elements of team work and thereby do not distinguish between the different dimensions. Our findings are following previous studies that analyze the perception of shared mental models (Aubé et al., 2015, 2018; Santos et al., 2015b; Burtscher and Oostlander, 2019). Although conceptually, shared mental models may regard to different aspects of work, practitioners in a domain do not always seem to make this distinction, and results have supported a one-factorial solution (Aubé et al., 2015, 2018; Santos et al., 2015b; Burtscher and Oostlander, 2019). Mental models considerably derive from the occupational context in which they raise (Cannon-Bowers et al., 1993).

Additionally, the referees' tasks are majorly concentrated during the referring of the futsal or football matches. Where everything is interconnected and must be deeply articulated between the team members, such contextual peculiarities might contribute to a perception of mental models as a singular whole. The suggested solution revealed a good fit for the single-factor model in three different samples of referees from two different sports (futsal and football), and as so, the H1 was supported. The RSMMM showed robustness in maintaining its dimensionality even when tested in a different sport other than football. Thus, such stability in the instrument's structure allows for useful perspectives in terms of its implementation within other sports.

The second hypothesis was supported; thus, reliability evidence was good. The internal consistency estimates (i.e., α and ω) values were satisfactory for all the samples, based on the recommended values (Nunnally and Bernstein, 1994). Previous studies that used a similar measure also had good values of internal consistency estimates, namely, the unidimensional shared mental models proposed by Santos et al. (2015b), which had $\alpha = 0.92$, and the PMU (which is another unidimensional measure) had $\alpha = 0.83$ and $\omega = 0.83$ (Burtscher and Oostlander, 2019). The TeamKMI internal consistency values of its five factors ranged from $\alpha_{\text{factor 3}} = 0.75$ to $\alpha_{\text{factor2}} = 0.89$ (Johnson et al., 2007), whereas the SMMS had not its internal consistency values reported in its original study (Santos et al., 2015a). As such, the obtained results are aligned with previous studies using similar measures.

Measurement invariance among futsal and football referees was obtained. Such kind of psychometric property is essential to establish comparisons between mental model scores. Previous studies with referee samples using other instruments (e.g., Referee Self-Efficacy Scale) have not obtained full uniqueness measurement invariance among the sport referred; in fact, only partial factorial invariance was achieved (Myers et al., 2012). While studies using other instruments among football referees samples obtained different levels of measurement invariance, namely, metric invariance between referees and assistant referees (Brandão et al., 2014), and strong invariance (i.e., scalar invariance) among elite and non-elite referees (Johansen et al., 2018). An instrument's mean scores should be compared only if scalar measurement invariance is granted (Marôco, 2014). Besides this fact, latent means comparisons should be implemented instead of raw means comparison, because the raw means do not account for measurement error. The established comparisons between shared mental models' latent means were made only after the achievement of full uniqueness measurement invariance. Football referees perceived significantly higher shared mental models' levels than futsal referees, which might be explained by the fact that in football the referee has a higher concentration of responsibilities in comparison with their assistants, whereas in futsal there is a higher sharing of those responsibilities between referees. As so, the perception of the shared mental models can be affected by the number of tasks with shared responsibility among the team members. The different levels of experience might also explain the differences between the shared mental models' levels because the football referees had more years of experience than their futsal counterparts.

Regarding the validity evidence based on the relation to other variables, the RSMMM revealed acceptable nomological validity evidence in terms of convergent evidence with team work engagement. However, the convergent evidence was poor regarding team effectiveness and team adaptive performance. Thus, H4 was partially verified. The correlations among shared mental models with team adaptive performance; and shared mental models with team effectiveness were too high, suggesting some overlap of the constructs. The correlation among shared mental models with team work engagement was more adequate to the extent of convergence expected. Such empirical evidence allows confirming the expected direction of the associations with work engagement, reflecting partial support for the proposed theoretical nomological network (Lissitz and Samuelsen, 2007). This source of evidence was analyzed using study III's sample (only composed by futsal referees), which might be a particularity of this small sample. Studies with small samples often report anomalously large effect sizes (Funder and Ozer, 2019), and as such, future replication studies might show that those effect sizes were overestimated with the used small sample ($n_{\text{study III}} = 60$) of futsal referees (Cumming, 2012).

This is the first instrument that explicitly measures shared mental models taking into consideration the specificities of football and futsal referee teams. All psychometric properties were indicative of good validity evidence, revealing a promising instrument for other contexts of referring (e.g., handball, basketball, rugby). The accumulated validity evidence seems to support the intended interpretation of the test scores for the RSMMM (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 2014).

Limitations and Directions for Future Research

While the current study provided some promising results about the RSMMM within the domain of football and futsal referees, there are some limitations in this study that must be acknowledged. For starters, this study was solely focused on referee teams within the sports of football and futsal. However, even though these are popular sports, it begs the question of how RSMMM would need to be altered to apply to other professional sport referee teams. Accordingly, it will be pertinent to see future research to examine the RSMMM in other sports and explore how this measure would need to be adjusted to be valuable and useful within other sports contexts.

In the present article, the data were analyzed at the individual level, not aggregated to the team level. Using the data aggregated to the team level could allow for a better understanding of the team's global perceptions of themselves instead of the isolated individuals' perceptions of the team. It is then possible to assess to which extent team members share mental models. For that propose, the level of agreement between team members would have to be considered for the subsequent analyses. It is worth mentioning that some of the referee teams are more stable in their constitution through the season than others (e.g., referees of lower categories tend to have more heterogeneity). The current cross-sectional study only provides a snapshot of the perceptions, which could vary if measured in a time frame (Levin, 2006).

The validity evidence based on the relations to other variables should be deeply investigated in terms of test criterion (e.g., higher team performance). As McNeese et al. (2015) urged, studying team cognition in sport must include a combination of both the shared knowledge and dynamical approaches. Future studies should investigate if shared mental models' levels are associated with performance (e.g., match analysis report ratings or associations' match/season ratings). Regarding the associations' ratings, it might be also interesting to check which of the components of the assessment (there are usually three components: physical performance, performance of the written test regarding rules and the laws of the game, and performance of the match observations attributed by the referees observers) has higher association with shared mental models. The validity evidence based on the relations to other variables should also be investigated in terms of convergent (similar constructs) and discriminant evidence (measures purportedly of different constructs), preferably using different measures of other nature rather than perceptions (i.e., self-report measures). It is challenging to prove that representations exist beyond the boundaries of an individual organism and that such representations can be somehow shared with others. The use of technologies, as multiple eye tracker (Wildman et al., 2014) or hyperbrain networks (Filho et al., 2017), is encouraged. This kind of measures can surpass some of self-report measures limitations (Schwarz, 1999; Baumeister et al., 2007), particularly when it concerns measuring the perception of behaviors instead of behaviors (Lonati et al., 2018). With the robustness of such evidence, the RSMMM might give a step forward in its establishment as a measure of shared mental models among referees.

Study III's sample size is small for structural equation modeling analysis, however, when looking to the number of futsal referees at the national level, it represents a considerable amount (30%) of the population of the Portuguese Football Federation ($N_{\text{season 2018}-2019} = 177$). Nevertheless, future studies should try to increase the number of referees both at the national and regional levels.

Additionally, given that in the collected samples were only a few football assistant referees (n = 25), the measurement invariance was not tested among them. Accordingly, given that assistant referees play an essential role within the football referee team, future studies should account for this and explore what impact having assistant referees more represented in future research samples can alter. Finally, within the current study and the underlying data that were used here, measurement invariance across time could not be examined. As a result, no statements regarding the trends that may exist across time can be made. In response, future research should examine this fact and collect the type of data necessary to be able to assess longitudinal measurement invariance. The assessment of validity evidence is an ongoing and never-ending process (Slaney, 2017); thus, the next steps should seem like a natural on the evolution of the RSMMM as an established measure to approach shared mental models within referees.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JorS, JP, and AP contributed to conception and design of the study. JorS, JP and JoãS organized the database. JorS and JM performed the statistical analysis. JorS wrote the first draft of the manuscript. JP, JoãS, and CS wrote sections of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

FUNDING

This work was supported by Fundação para a Ciência e Tecnologia, grant UIDB/00315/2020. This work was also produced with the support of INCD funded by FCT and FEDER under the project 22153-01/SAICT/2016.

ACKNOWLEDGMENTS

We would like to thank the Portuguese Football Federation for providing access to seminars and for sharing the digital survey with their referees (Study III).

REFERENCES

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Psychological Association.

Aragão e Pina, J., Passos, A. M., Araújo, D., and Maynard, M. T. (2018). Football refereeing: an integrative review. *Psychol. Sport Exerc.* 35, 10–26. doi: 10.1016/j.psychsport.2017.10.006

Aragão e Pina, J., Passos, A. M., Carvalho, H., and Travis Maynard, M. (2019). To be or not to be an excellent football referee: different experts' viewpoints. *J. Sports Sci.* 37, 692–700. doi: 10.1080/02640414.2018.1522940

Aragão e Pina, J., Passos, A. M., Travis Maynard, M., and Sinval, J. (2021). Self-efficacy, mental models and team adaptation: a first approach on football and futsal refereeing. *Psychol. Sport Exerc.* 52:101787. doi: 10.1016/j.psychsport.2020.101787

Aubé, C., and Rousseau, V. (2005). Team goal commitment and team effectiveness: the role of task interdependence and supportive behaviors. *Gr. Dyn. Theory Res. Pract.* 9, 189–204. doi: 10.1037/1089-2699.9.3.189

Aubé, C., Rousseau, V., Brunelle, E., and Marques, D. (2018). The relevance of being "on the same page" to succeed as a project team: a moderated mediation model. *Motiv. Emot.* 42, 804–815. doi: 10.1007/s11031-018-9706-2

Aubé, C., Rousseau, V., and Tremblay, S. (2015). Perceived shared understanding in teams: the motivational effect of being 'on the same page.'. *Br. J. Psychol.* 106, 468–486. doi: 10.1111/bjop.12099

Bartlett, M. S. (1951). The effect of standardization on a χ2 approximation in factor analysis. *Biometrika* 38, 337–344. doi: 10.2307/2332580

Baumeister, R. F., Vohs, K. D., and Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: whatever happened to actual behavior? *Perspect. Psychol. Sci.* 2, 396–403. doi: 10.1111/j.1745-6916.2007.00051.x

Boomsma, A. (2000). Reporting analyses of covariance structures. *Struct. Equ. Model. A Multidiscip. J.* 7, 461–483. doi: 10.1207/S15328007SEM0703_6

Boyer, S., Rix-Lièvre, G., and Récopé, M. (2015). L'arbitrage de haut niveau, une affaire d'équipe. *Mov. Sport Sci. Sci. Mot.* 101, 91–101. doi: 10.1051/sm/2014014

Brandão, M. R. F., Serpa, S., Rosado, A., and Weinberg, R. (2014). Psychometric properties of the burnout inventory for referees. *Mot. Rev. Educ. Física* 20, 374–383. doi: 10.1590/S1980-65742014000400003

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*, 2nd Edn. New York, NY: The Guilford Press.

Burtscher, M. J., and Oostlander, J. (2019). Perceived mutual understanding (PMU): development and initial testing of a German short scale for perceptual team cognition. *Eur. J. Psychol. Assess.* 35, 98–108. doi: 10.1027/1015-5759/a000360

Byrne, B. M. (2010). *Structural Equation Modeling with AMOS*, 3rd Edn. New York, NY: Routledge.

Cannon-Bowers, J. A., and Salas, E. (2001). Reflections on shared cognition. *J. Organ. Behav.* 22, 195–202. doi: 10.1002/job.82

Cannon-Bowers, J. A., Salas, E., and Converse, S. (1993). "Shared mental models in expert team decision making," in *Individual and Group Decision Making: Current Issues*, ed. N. J. Castellan (Hillsdale, NJ: Lawrence Erlbaum), 221–246. doi: 10.4324/9780203772744-20

Canty, A., and Ripley, B. (2020). *boot: Bootstrap R (S-Plus) Functions (R package version 1.3-25) [Computer software]*. Available online at: https://cran.r-project.org/web/packages/boot/boot.pdf

Carpenter, J., and Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Stat. Med.* 19, 1141–1164. doi: 10.1002/(SICI)1097-0258(20000515)19:9<1141::AID-SIM479<3.0.CO;2-F

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J. B., and Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociol. Methods Res.* 36, 462–494. doi: 10.1177/0049124108314720

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equ. Model. A Multidiscip. J.* 9, 233–255. doi: 10.1207/S15328007SEM0902_5

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Hillsdale, NJ: Lawrence Erlbaum Associates.

Collina, P. (2004). *As Minhas Regras do Jogo - O que o Futebol me Ensinou Sobre a Vida*. Lisboa: Editorial Presença.

Costa, P. L., Passos, A. M., and Bakker, A. B. (2014a). Empirical validation of the team work engagement construct. *J. Pers. Psychol.* 13, 34–45. doi: 10.1027/1866-5888/a000102

Costa, P. L., Passos, A. M., and Bakker, A. B. (2014b). Team work engagement: a model of emergence. *J. Occup. Organ. Psychol.* 87, 414–436. doi: 10.1111/joop.12057

Costa, P. L., Passos, A. M., and Bakker, A. B. (2016). The work engagement grid: predicting engagement from two core dimensions. *J. Manag. Psychol.* 31, 774–789. doi: 10.1108/JMP-11-2014-0336

Courtney, M. G. R. (2012). Determining the number of factors to retain in EFA: using the SPSS R-Menu v2.0 to make more judicious estimations. *Pract. Assess. Res. Eval.* 18, 1–14.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334. doi: 10.1007/BF02310555

Cumming, G. (2012). *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. New York, NY: Routledge.

Cunningham, I., Simmons, P., Mascarenhas, D. R. D., and Redhead, S. (2014). Skilled interaction: concepts of communication and player management in the development of sport officials. *Int. J. Sport Commun.* 7, 166–187. doi: 10.1123/IJSC.2013-0098

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., and Billiet, J. (2014). Measurement equivalence in cross-national research. *Annu. Rev. Sociol.* 40, 55–75. doi: 10.1146/annurev-soc-071913-043137

Davison, A. C., and Hinkley, D. V. (1997). *Bootstrap Methods and Their Application*. Cambridge: Cambridge University Press.

DeChurch, L. A., and Mesmer-Magnus, J. R. (2010). Measuring shared team mental models: a meta-analysis. *Gr. Dyn. Theory Res. Pract.* 14, 1–14. doi: 10.1037/a0017455

Diotaiuti, P., Falese, L., Mancone, S., and Purromuto, F. (2017). A structural model of self-efficacy in handball referees. *Front. Psychol.* 8:811. doi: 10.3389/fpsyg.2017.00811

Efron, B., and Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. Boca Raton, FL: Chapman & Hall.

Filho, E., Pierini, D., Robazza, C., Tenenbaum, G., and Bertollo, M. (2017). Shared mental models and intra-team psychophysiological patterns: a test of the juggling paradigm. *J. Sports Sci.* 35, 112–123. doi: 10.1080/02640414.2016.1158413

Filho, E., and Tenenbaum, G. (2012). "Team mental models in sports: an overview," in *Athletic Insight's Writings in Sport Psychology (Sports and Athletics Preparation Performance and Psychology)*, ed. R. Schinke (Hauppauge, NY: Nova Science Pub Inc), 329–342.

Finch, W. H., and French, B. F. (2015). *Latent Variable Modeling with R*. New York, NY: Routledge.

Finney, S. J., and DiStefano, C. (2013). "Non-normal and categorical data in structural equation modeling," in *Structural Equation Modeling: A Second Course*, eds G. R. Hancock and R. O. Mueller (Charlotte, NC: Information Age Publishing), 439–492.

Finney, S. J., DiStefano, C., and Kopp, J. P. (2016). "Overview of estimation methods and preconditions for their application with structural equation modeling," in *Principles and Methods of Test Construction: Standards and Recent Advances*, eds K. Schweizer and C. DiStefano (Boston, MA: Hogrefe Publishing), 135–165.

Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Mark Res.* 18, 39–50. doi: 10.2307/3151312

Funder, D. C., and Ozer, D. J. (2019). Evaluating effect size in psychological research: sense and nonsense. *Adv. Methods Pract. Psychol. Sci.* 2, 156–168. doi: 10.1177/2515245919847202

Hackman, J. R. (1987). "The design of work teams,", in *Handbook of Organizational Behaviour*, ed. J. W. Lorsch (Englewood Cliffs, NJ: Prentice-Hall), 315–342.

Hair, J. F., Babin, B. J., Anderson, R. E., and Black, W. C. (2019). *Multivariate Data Analysis*, 8th Edn. Andover: Cengage Learning.

Hancock, D. J., Martin, L. J., Evans, M. B., and Paradis, K. F. (2018). Exploring perceptions of group processes in ice hockey officiating. *J. Appl. Sport Psychol.* 30, 222–240. doi: 10.1080/10413200.2017.1349208

Hoyle, R. H. (ed.) (1995). *Structural Equation Modeling: Concepts, Issues and Applications*. Thousand Oaks, CA: SAGE Publications.

Iacobucci, D., and Duhachek, A. (2003). Advancing alpha: measuring reliability with confidence. *J. Consum. Psychol.* 13, 478–487. doi: 10.1207/S15327663JCP1304_14

Johansen, B. T., Ommundsen, Y., and Haugen, T. (2018). Referee efficacy in the context of Norwegian soccer referees – a meaningful construct? *Psychol. Sport Exerc.* 38, 184–191. doi: 10.1016/j.psychsport.2018.06.013

Johnson, T. E., Lee, Y., Lee, M., O'Connor, D. L., Khalil, M. K., and Huang, X. (2007). Measuring sharedness of team-related knowledge: design and validation of a shared mental model instrument. *Hum. Resour. Dev. Int.* 10, 437–454. doi: 10.1080/13678860701723802

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., and Rosseel, Y. (2019). *semTools: Useful Tools for Structural Equation Modeling (R package version 0.5-1.933) [Computer software]*. Available online at: https://cran.r-project.org/package=semTools

Kaiser, H. F., and Rice, J. (1974). Little Jiffy, Mark IV. *Educ. Psychol. Meas.* 34, 111–117. doi: 10.1177/001316447403400115

Kellermanns, F. W., Floyd, S. W., Pearson, A. W., and Spencer, B. (2008). The contingent effect of constructive confrontation on the relationship between shared mental models and decision quality. *J. Organ. Behav.* 29, 119–137. doi: 10.1002/job.497

Klimoski, R., and Mohammed, S. (1994). Team mental model: construct or metaphor? *J. Manage.* 20, 403–437. doi: 10.1177/014920639402000206

Lemon, J. (2006). Plotrix: a package in the red light district of R. *R-News* 6, 8–12.

Levin, K. A. (2006). Study design III: cross-sectional studies. *Evid. Based. Dent.* 7, 24–25. doi: 10.1038/sj.ebd.6400375

Lim, B., and Klein, K. J. (2006). Team mental models and team performance: a field study of the effects of team mental model similarity and accuracy. *J. Organ. Behav.* 27, 403–418. doi: 10.1002/job.387

Lissitz, R. W., and Samuelsen, K. (2007). A suggested change in terminology and emphasis regarding validity and education. *Educ. Res.* 36, 437–448. doi: 10.3102/0013189X07311286

Lonati, S., Quiroga, B. F., Zehnder, C., and Antonakis, J. (2018). On doing relevant and rigorous experiments: review and recommendations. *J. Oper. Manag.* 64, 19–40. doi: 10.1016/j.jom.2018.10.003

Lüdecke, D. (2019). *sjstats: Statistical Functions for Regression Models (R package version 0.17.3)*.

MacCallum, R. C., Browne, M. W., and Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychol. Methods* 1, 130–149. doi: 10.1037/1082-989X.1.2.130

Mallo, J., Frutos, P. G., Juárez, D., and Navarro, E. (2012). Effect of positioning on the accuracy of decision making of association football top-class referees and assistant referees during competitive matches. *J. Sports Sci.* 30, 1437–1445. doi: 10.1080/02640414.2012.711485

Marks, M. A., Sabella, M. J., Burke, C. S., and Zaccaro, S. J. (2002). The impact of cross-training on team effectiveness. *J. Appl. Psychol.* 87, 3–13. doi: 10.1037/0021-9010.87.1.3

Marks, M. A., Zaccaro, S. J., Mathieu, J. E., Gessner, D., Klimoski, R., Sanchez, J., et al. (2000). Performance implications of leader briefings and team-interaction training for team adaptation to novel environments. *J. Appl. Psychol.* 85, 971–986. doi: 10.1037//0021-90I0.85.6.971

Marôco, J. (2014). *Análise de Equações Estruturais: Fundamentos Teóricos, Software & Aplicações*, 2nd Edn. Pêro Pinheiro: ReportNumber, Lda.

Marôco, J. (2018). *Análise Estatística Com o SPSS Statistics*, 7th Edn. Pêro Pinheiro: ReportNumber, Lda.

Marques-Quinteiro, P., Ramos-Villagrasa, P. J., Passos, A. M., and Curral, L. (2015). Measuring adaptive performance in individuals and teams. *Team Perform. Manag. An Int. J.* 21, 339–360. doi: 10.1108/TPM-03-2015-0014

Mascarenhas, D. R. D., Collins, D., Mortimer, P. W., and Morris, B. (2005). Training accurate and coherent decision making in rugby union referees. *Sport Psychol.* 19, 131–147. doi: 10.1123/tsp.19.2.131

Mascarenhas, D. R. D., O'Hare, D., and Plessner, H. (2006). The psychological and performance demands of association football refereeing. *Int. J. Sport Psychol.* 37, 99–120.

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., and Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *J. Appl. Psychol.* 85, 273–283. doi: 10.1037/0021-9010.85.2.273

Mathieu, J. E., Rapp, T. L., Maynard, M. T., and Mangos, P. M. (2009). Interactive effects of team and task shared mental models as related to air traffic controllers' collective efficacy and effectiveness. *Hum. Perform.* 23, 22–40. doi: 10.1080/08959280903400150

Maydeu-Olivares, A., Shi, D., and Rosseel, Y. (2018). Assessing fit in structural equation models: a monte-carlo evaluation of RMSEA versus SRMR confidence intervals and tests of close fit. *Struct. Equ. Model. A Multidiscip. J.* 25, 389–402. doi: 10.1080/10705511.2017.1389611

Maynard, M. T., Kennedy, D. M., and Sommer, S. A. (2015). Team adaptation: a fifteen-year synthesis (1998–2013) and framework for how this literature needs to "adapt" going forward. *Eur. J. Work Organ. Psychol.* 24, 652–677. doi: 10.1080/1359432X.2014.1001376

McDonald, R. P. (1999). *Test Theory: A Unified Treatment*. Mahwah, NJ: Routledge.

McDonald, R. P., and Ho, M.-H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychol. Methods* 7, 64–82. doi: 10.1037/1082-989X.7.1.64

McEwan, D., and Beauchamp, M. R. (2014). Teamwork in sport: a theoretical and integrative review. *Int. Rev. Sport Exerc. Psychol.* 7, 229–250. doi: 10.1080/1750984X.2014.932423

McGrath, J. E. (1991). Time, interaction, and performance (TIP): a theory of groups. *Small Gr. Res.* 22, 147–174. doi: 10.1177/1046496491222001

McNamara, A., Arino de la Rubia, E., Zhu, H., Ellis, S., and Quinn, M. (2018). *skimr: Compact and Flexible Summaries of Data (R package version 1.0.3) [Computer software]*. Available online at: https://cran.r-project.org/package=skimr

McNeese, N. J., Cooke, N. J., Fedele, M. A., and Gray, R. (2015). Theoretical and methodical approaches to studying team cognition in sports. *Proc. Manuf.* 3, 1211–1218. doi: 10.1016/j.promfg.2015.07.201

Mohammed, S., Ferzandi, L., and Hamilton, K. (2010). Metaphor no more: a 15-year review of the team mental model construct. *J. Manage.* 36, 876–910. doi: 10.1177/0149206309356804

Mohammed, S., Klimoski, R., and Rentsch, J. R. (2000). The measurement of team mental models: we have no shared schema. *Organ. Res. Methods* 3, 123–165. doi: 10.1177/109442810032001

Muponde, R., and Muchemwa, K. (2011). Dictatorships, disasters, and African soccer: reflections on a moment in Zimbabwean soccer. *African Ident.* 9, 279–290. doi: 10.1080/14725843.2011.591227

Muthén, B. O. (1983). Latent variable structural equation modeling with categorical data. *J. Econom.* 22, 43–65. doi: 10.1016/0304-4076(83)90093-3

Myers, N. D., Feltz, D. L., Guillén, F., and Dithurbide, L. (2012). Development of, and initial validity evidence for, the referee self-efficacy scale: a multistudy report. *J. Sport Exerc. Psychol.* 34, 737–765. doi: 10.1123/jsep.34.6.737

Nunnally, J. C., and Bernstein, I. H. (1994). *Psychometric Theory*, 3rd Edn. New York, NY: McGraw-Hill.

Peters, G.-J. Y. (2018). *userfriendlyscience: Quantitative Analysis Made Accessible (R package version 0.7.2)*.

Pulakos, E. D., Dorsey, D. W., and White, S. S. (2006). "Adaptability in the workplace: selecting an adaptive workforce," in *Understanding adaptability: A prerequisite for effective performance within complex environments*, eds C. S. Burke, L. G. Pierce, and E. Salas (Amsterdam: Elsevier), 41–71.

R Core Team (2020). *R: A Language and Environment for Statistical Computing (version 4.0.0) [Computer software]*. Vienna: R Foundation for Statistical Computing.

Randall, K. R., Resick, C. J., and DeChurch, L. A. (2011). Building team adaptive capacity: the roles of sensegiving and team composition. *J. Appl. Psychol.* 96, 525–540. doi: 10.1037/a0022622

Raykov, T. (2001). Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *Br. J. Math. Stat. Psychol.* 54, 315–323. doi: 10.1348/000711001159582

Resick, C. J., Dickson, M. W., Mitchelson, J. K., Allison, L. K., and Clark, M. A. (2010a). Team composition, cognition, and effectiveness: examining mental model similarity and accuracy. *Gr. Dyn. Theory Res. Pract.* 14, 174–191. doi: 10.1037/a0018444

Resick, C. J., Murase, T., Bedwell, W. L., Sanz, E., Jimenez, M., and DeChurch, L. A. (2010b). Mental model metrics and team adaptability: a multi-facet multi-method examination. *Gr. Dyn. Theory Res. Pract.* 14, 332–349. doi: 10.1037/a0018822

Revelle, W. (2019). *psych: Procedures for Psychological, Psychometric, and Personality Research (R package version 1.9.4) [Computer software]*. Available online at: https://cran.r-project.org/package=psych.

Rhemtulla, M., Brosseau-Liard, P. É, and Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychol. Methods* 17, 354–373. doi: 10.1037/a0029315

Rosseel, Y. (2012). lavaan: an R package for structural equation modeling. *J. Stat. Softw.* 48, 1–21. doi: 10.18637/jss.v048.i02

RStudio Team (2020). *RStudio: Integrated development for R (version 1.4.672) [Computer software]*. Available online at: http://www.rstudio.com/

Ruscio, J. (2018). *RGenData: Generates Multivariate Nonnormal Data and Determines How Many Factors to Retain (R package version 1.0) [Computer software]*. Available online at: https://cran.r-project.org/package=RGenData

Ruscio, J., and Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychol. Assess.* 24, 282–292. doi: 10.1037/a0025697

Samuel, R. D. (2015). A psychological preparation framework for elite soccer referees: a practitioner's perspective. *J. Sport Psychol. Action* 6, 170–187. doi: 10.1080/21520704.2015.1065938

Santos, C. M., and Passos, A. M. (2013). Team mental models, relationship conflict and effectiveness over time. *Team Perform. Manag. An Int. J.* 19, 363–385. doi: 10.1108/TPM-01-2013-0003

Santos, C. M., Passos, A. M., and Uitdewilligen, S. (2016). When shared cognition leads to closed minds: temporal mental models, team learning, adaptation and performance. *Eur. Manag. J.* 34, 258–268. doi: 10.1016/j.emj.2015.11.006

Santos, C. M., Uitdewilligen, S., and Passos, A. M. (2015a). A temporal common ground for learning: the moderating effect of shared mental models on the relation between team learning behaviours and performance improvement. *Eur. J. Work Organ. Psychol.* 24, 710–725. doi: 10.1080/1359432X.2015.1049158

Santos, C. M., Uitdewilligen, S., and Passos, A. M. (2015b). Why is your team more creative than mine? The influence of shared mental models on intra-group conflict, team creativity and effectiveness. *Creat. Innov. Manag.* 24, 645–658. doi: 10.1111/caim.12129

Satorra, A., and Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika* 66, 507–514. doi: 10.1007/BF02296192

Schwarz, N. (1999). Self-reports: how the questions shape the answers. *Am. Psychol.* 54, 93–105. doi: 10.1037/0003-066X.54.2.93

Signorell, A., Aho, K., Alfons, A., Anderegg, N., Aragon, T., Arppe, A., et al. (2019). *DescTools: Tools for Descriptive Statistics (R package version 0.99.28) [Computer software]*. Available online at: https://cran.r-project.org/package=DescTools

Sinval, J., Marques-Pinto, A., Queirós, C., and Marôco, J. (2018a). Work engagement among rescue workers: psychometric properties of the Portuguese UWES. *Front. Psychol.* 8:2229. doi: 10.3389/fpsyg.2017.02229

Sinval, J., Pasian, S. R., Queirós, C., and Marôco, J. (2018b). Brazil-Portugal transcultural adaptation of the UWES-9: internal consistency, dimensionality, and measurement invariance. *Front. Psychol.* 9:353. doi: 10.3389/fpsyg.2018.00353

Slaney, K. (2017). *Validating Psychological Constructs: Historical, Philosophical, and Practical Dimensions*. London: Palgrave Macmillan.

Tomás, J. M., de los Santos, S., Alonso-Andres, A., and Fernández, I. (2017). Validation of the Maslach burnout inventory-general survey on a representative sample of dominican teachers: normative data. *Span. J. Psychol.* 19:E83. doi: 10.1017/sjp.2016.91

Uitdewilligen, S., Rico, R., and Waller, M. J. (2018). Fluid and stable: dynamics of team action patterns and adaptive outcomes. *J. Organ. Behav.* 39, 1113–1128. doi: 10.1002/job.2267

Vicente, Â, Rodrigues, B. S., and D'Oliveira, T. C. (2014). Escala de Eficácia de Equipas (3Es): adaptação para o contexto Português [Scale of Effectiveness of Teams (3Es): adaptation to the Portuguese context]. *Lab. Psicol.* 12, 3–13. doi: 10.14417/lp.884

Wildman, J. L., Salas, E., and Scott, C. P. R. (2014). Measuring cognition in teams: a cross-domain review. *Hum. Fact. J. Hum. Fact. Ergon. Soc.* 56, 911–941. doi: 10.1177/0018720813515907

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 16 February 2021
doi: 10.3389/fpsyg.2021.633228

Superior Performance in Skilled Golfers Characterized by Dynamic Neuromotor Processes Related to Attentional Focus

Kuo-Pin Wang[1,2], Cornelia Frank[3], Yen-yu Tsai[4], Kao-Hung Lin[4], Tai-Ting Chen[4], Ming-Yang Cheng[5], Chung-Ju Huang[6], Tsung-Min Hung[4,7] and Thomas Schack[1,2]**

[1] *Center for Cognitive Interaction Technology (CITEC), Bielefeld University, Bielefeld, Germany,* [2] *Neurocognition and Action – Biomechanics Research Group, Faculty of Psychology and Sports Science, Bielefeld University, Bielefeld, Germany,* [3] *Sports and Movement Group, Department of Sports and Movement Sciences, School of Educational and Cultural Studies, University of Osnabrück, Osnabrück, Germany,* [4] *Department of Physical Education, National Taiwan Normal University, Taipei, Taiwan,* [5] *School of Psychology, Shanghai University of Sport, Shanghai, China,* [6] *Graduate Institute of Sport Pedagogy, University of Taipei, Taipei, Taiwan,* [7] *Institute for Research Excellence in Learning Science, National Taiwan Normal University, Taipei, Taiwan*

OPEN ACCESS

Edited by:
Yair Galily,
Interdisciplinary Center Herzliya, Israel

Reviewed by:
Takehiro Iwatsuki,
Penn State Altoona, United States
Gal Ziv,
Academic College at Wingate, Israel

****Correspondence:***
Tsung-Min Hung
ernesthungkimo@yahoo.com.tw
Thomas Schack
thomas.schack@uni-bielefeld.de

Specialty section:
This article was submitted to Movement Science and Sport Psychology, a section of the journal Frontiers in Psychology

Received: *24 November 2020*
Accepted: *25 January 2021*
Published: *16 February 2021*

Citation:
Wang K-P, Frank C, Tsai Y-y, Lin K-H, Chen T-T, Cheng M-Y, Huang C-J, Hung T-M and Schack T (2021) Superior Performance in Skilled Golfers Characterized by Dynamic Neuromotor Processes Related to Attentional Focus. Front. Psychol. 12:633228. doi: 10.3389/fpsyg.2021.633228

The meshed control theory assumes that cognitive control and automatic processes work together in the natural attention of experts for superior performance. However, the methods adopted by previous studies limit their capacity to provide in-depth information on the neuromotor processes. This experiment tested the theory with an alternative approach. Twelve skilled golfers were recruited to perform a putting task under three conditions: (1) normal condition, with no focus instruction (NC), (2) external focus of attention condition (EC), and (3) internal focus of attention condition (IC). Four blocks of 10 putts each were performed under each condition. The putting success rate and accuracy were measured and electroencephalographies (EEGs) were recorded. The behavioral results showed that the NC produced a higher putting success rate and accuracy than the EC and IC. The EEG data showed that the skilled golfers' attentional processes in the NC initially resembled those in the EC and then moved toward those in the IC just before putting. This indicates a switch from more automatic processes to cognitive control processes while preparing to putt. The findings offer support for the meshed control theory and indicate the dynamic nature of neuromotor processes for the superior performance of athletes in challenging situations.

Keywords: electroencephalography, precision sports, attention, constrained action hypothesis, meshed control theory

INTRODUCTION

Attentional focus is a crucial factor in superior skilled performance (Wulf and Su, 2007; Wulf and Lewthwaite, 2016). In recent years, a number of research studies have specifically examined the internal vs. the external categories of attentional focus (Wulf and Prinz, 2001). An internal focus refers to attention being directed toward specific body actions, whereas an external focus relates to the effect that those body actions have on the environment (Wulf et al., 1998). A growing number of studies using attentional instructions have consistently demonstrated the advantage for motor performance of adopting an external focus relative to an internal focus (Bell and Hardy, 2009; Kearney, 2015). For instance, an external focus has been found to increase movement effectiveness, such as accuracy (Wulf and Su, 2007; Bell and Hardy, 2009), and movement efficiency, such as maximum force production (Marchant et al., 2011), speed, and endurance (Stoate and Wulf, 2011), in highly skilled performers. One possible explanation for the advantages of an external focus is provided by the constrained action hypothesis (Wulf et al., 2001; Mcnevin et al., 2003). This hypothesis assumes that individuals who direct their attention internally during skill execution experience conscious control of their movements (i.e., constraints on the motor system), which interferes with automatic control processes. In contrast, individuals who direct their attention to the effects of the movement experience less movement control. It promotes the use of automatic processes during motor preparation (Wulf et al., 2001). However, Schücker et al. (2014) and Vitali et al. (2019) who defined focusing on physical sensations as being internal sensations observed that internal sensation does not impair performance. They suggested that the internal sensations do not disrupt automated motor processes. Given that the definition of internal focus is different from the original definition of Wulf et al. (1998), it is possible to focus internally in more than one way which may lead to different results.

Furthermore, an external focus can be further distinguished as proximal or distal (Mcnevin et al., 2003). A proximal focus is close to the body (i.e., the club motion and the clubface when putting in golf) whereas a distal focus is further from the body (i.e., the desired trajectory of the ball). Previous studies that have adopted attentional instructions to examine proximal and distal focus in skilled performance have demonstrated that a distal focus benefits performance effectiveness relative to a proximal focus in golf chipping (Bell and Hardy, 2009) and putting (Kearney, 2015). These findings suggested that performers tend to move their attention toward higher levels of representation (i.e., the ball trajectory) and engage in less conscious monitoring of the lower-level features of the action (i.e., the clubface and its motion), resulting in adequate cognitive resources for motor preparation and a better coordinated action (Wulf and Su, 2007). However, these studies only adopted dichotomous focus manipulations and may not represent the actual type of attentional processes. According to the meshed control theory, the cognitive control and automatic processes of experts under natural conditions work in a synergistic manner for superior performance (Christensen et al., 2016). Furthermore, cognitive control processes focus on the higher-order strategic control (e.g., the ball trajectory) and situation control (e.g., adjusting the body action appropriately). Although both foci of attention are different from each other, both involve similar cognitive control processing since elite performers intend to pay their attention to self-monitoring processes (Bortoli et al., 2012; Wang et al., 2019). That is, both foci of attention are similar to the moment-to-moment awareness in mindfulness. For example, skilled performers need to be aware to both their goal (i.e., the higher-order strategic control) and to their action (situation control), without disrupting automated motor processes on implementation control (e.g., the clubface and its motion). In qualitative studies, Oliver et al. (2020) observed that skilled golfers switched their attentional focus between an external focus of attention and an internal focus of attention during motor preparation. Similarly, Bahmani et al. (2019) uncovered those athletes who switched both foci of attention in a difficult situation showed superior skill execution. Given that the nature of attentional focus is a complex and dynamic aspect of the superior skilled performance of athletes, it is important to investigate attention over time during skilled performance. However, the methods adopted in the aforementioned behavioral research are limited as they cannot provide in-depth information regarding the neuromotor processes over time. Electroencephalography (EEG) is an ideal method to detect dynamic processes in the superior skilled performance of athletes.

Electroencephalography provides a high temporal resolution of neural activity and therefore provides a window into understanding the dynamics of neuromotor processes in preparation for action. Specifically, the alpha 2 band (10–12 Hz) has been associated with an internal and an external focus during motor preparation (Radlo et al., 2002; Ellmers et al., 2016). For example, Radlo et al. (2002) examined the effects of an external and internal attentional focus on electrocortical activity and observed that novices who adopted an external focus relative to an internal focus had lower left hemisphere (T7 and O1) and right temporal (T8 and O2) alpha power in a dart throwing task. This suggested that novices who adopted an external focus had similar EEG patterns to skilled archers in their superior performance (Landers et al., 1991), which indicated more efficient neural processing. Furthermore, measures of cortico-cortical communication are critical to the determination of psychomotor efficiency during motor preparation (Hatfield, 2018). EEG coherence, the connection between different cortical areas, is an excellent measure of cortico-cortical communication (Deeny et al., 2003, 2009). Specifically, higher coherence indicates stronger cotico-cortical communication, whereas lower coherence implies the functional autonomy of cortex. Ellmers et al. (2016) utilized T7 (verbal-analytical) with Fz (motor-planning) and T8 (visuospatial) with Fz in assessing attentional focus in a golf putting postural task. Ellmers et al. (2016) observed that adopting an external focus relative to an internal focus of attention in novices resulted in decreased Fz-T7 alpha 2 coherence. Ellmers et al. (2016) suggested that this decreased Fz–T7 alpha 2 coherence reflected less use of verbal-analytical processes, and thus more automatic processes, whereas increased Fz–T7 alpha 2 coherence would be associated with more cognitive control processes. Although Fz–T7 alpha 2 coherence is valid for assessing the neuromotor processes underlying attentional focus, the aforementioned evidence did not reveal the dynamics of cortico-cortical communication underlying the nature of attentional focus in skilled performers. The simultaneous measurements of both processes in two time windows (one from −2 to −1 s and one from −1 s to task execution) and in a normal focus condition (i.e., no instruction) compared to an internal and an external focus condition (i.e., instruction) is needed to shed light on the dynamic nature of attentional processes in skilled performance.

To summarize, the present study examines the meshed control theory, which assumes that cognitive control and automatic processes work together in attentional processes underlying superior performance (Christensen et al., 2016). We adopted an approach to compare the well-known psychological states induced by different attentional focus manipulations (i.e., an external focus vs. an internal focus) with unmanipulated attentional processes, which have not been examined in previous studies. In addition, Bernier et al. (2011) and Oliver et al. (2020) observed that skilled golfers commonly and naturally – without any manipulation – adopted a distal external focus (i.e., visualizing the trajectory of ball) during motor preparation in a challenging condition. Along these lines, we assume that skilled golfers naturally adopt a distal external focus (i.e., visualizing the trajectory of the ball) during motor preparation. Therefore, not manipulating a distal external focus could enable a better understanding of the nature of attentional focus processes in skilled golfers. To examine the dynamics of neuromotor processes in skilled performers during motor preparation, we used two time windows for detecting EEG coherence, one from −2 to −1 s before putting and one from −1 s to execution of the putt. In addition, based on previous research showing that lower Fz–T7 alpha 2 coherence is associated with an external focus (i.e., more automatic processes) relative to an internal focus of attention (Ellmers et al., 2016), we expected skilled golfers adopting an external focus relative to an internal focus to show decreased Fz–T7 alpha 2 coherence (e.g., less verbal-analytical processes), indicating more automatic processes. Furthermore, most importantly, given the meshed control theory and Williams et al. (2015) assumes attention shifts from an external focus to an internal focus associated with superior motor performance in skilled performers in challenging conditions, we expected skilled performers in a no instruction condition (NC; exhibiting the actual attentional processes without manipulation) to have initial Fz–T7 alpha 2 coherence similar to that in the external focus condition (EC; more automatic processes), and then to switch to a coherence similar to that in an internal focus condition (IC; more cognitive control processes) for superior performance.

MATERIALS AND METHODS

Participants

The number of participants was determined by means of power analysis software (G*Power 3.1). Consistent with previous EEG study in attentional instruction (Ellmers et al., 2016), we set the following input parameters for using a repeated measures ANOVA with alpha = 0.05, power = 0.80, effect size = 0.33~0.50 (corresponds to η_p^2 = 0.10~0.20), and actual power = 0.80. The resulting sample size specification as in SPSS recommendation was N = 8~16. Being aware of the potential for power analysis biases in the neuroscience field (Albers and Lakens, 2018) and large samples of elite athletes are typically hard to recruit for scientific studies, 12 skilled golfers (four females, eight males; mean age = 21.08 ± 4.64) with a mean golf experience of 9.66 years (SD = 3.57) were recruited, with a mean handicap of 3.25 (SD = 0.97). According to United States Golf Association (USGA) statistics, a handicap range of 2.0–5.9 reflects golf skill that is above 87.7% of female elite golfers and 98.27% of male elite golfers in the country (United States Golf Association, 2018). Thus, the skilled golfers could be defined as elite golfers at a high competitive level (Swann et al., 2015; Scharfen and Memmert, 2019). In addition, all of the recruited participants met the following selection criteria: (1) no history of neurological disease, (2) right-handed (Oldfield, 1971). All participants gave an informed written consent, and the study was approved by the Research Ethics Committee of National Taiwan Normal University. All methods were carried out in accordance with the relevant guidelines and regulations of Research Ethics Committee.

Measures

Golf Putting Task

The golfers were set a putting task similar to that used by Wang et al. (2019). The putting task was executed in the laboratory on an artificial putting green (600 cm × 90 cm). Participants used their own golf putters to putt regular-size white golf balls (4.27 cm diameter) towards a standard-size hole (diameter = 10.8 cm). We instructed participants to putt from a distance that was chosen to set a difficult task, such that the average of five putting success rate was 40–60% in the warm-up phase. For example, the individual putting distance was designated 40–60% putting success rate. All participants putted 300 cm in the beginning distance. They performed five putts and the distance was adjusted relying on whether the average of five putting success rate was 40–60% or not. If the success rate was between 40 and 60%, the putting distance was set at 300 cm. If the success rate was above 60%, the putting distance would increase 30 cm and then they performed extra five putts to ensure the success rate reached 40–60%. On the contrary, if the success rate was below 40%, the putting distance would decrease 30 cm and then they performed extra five putts to ensure the success rate reached 40–60%. After the appropriate putting distance was decided, the participant performed 40 putts in each condition. Furthermore, the average distance that was related to 40–60% success is mean = 302 ± 24 cm. To avoid the learning effect during the task, the ball placed on different points of the circumference of the individual putting distance. For measuring EEG activity during motor preparation, the motor preparation period was defined as the time between placing the putter behind the ball and initiating the backswing (Lam et al., 2010), with the event-marker initiated *via* an infrared sensor that detected the movement of the backswing during each trial. Putting performance was judged by using a measuring tape to measure the distance between the ball and the hole; when a ball was holed, we registered the putt as having a distance of 0 cm.

Experimental Conditions

Based on previous research on attentional focus induced by giving explicit instructions to skilled golfers (Bell and Hardy, 2009; Toner and Moran, 2011), our experimental design was similar to those of Bell and Hardy (2009) and Toner and Moran (2011) who adopted three conditions – no instruction, an external focus, and an internal focus – for understanding the effects of attentional focus on performance effectiveness. The three conditions represented a putting task in which participants received instructions to:

1. Putt as they normally would in the NC;
2. Focus on the position of the clubface in the EC;
3. Focus on adjusting direction with hand movements and feeling sensation of hand movement in the IC.

Manipulation Check

To ensure that all participants had adopted the focus as instructed, we asked the participants to rate their experience on a five-point Likert scale, ranging from 0 (*not at all*) to 4 (*very much*; Bell and Hardy, 2009). After each putt, participants were asked to report the extent to which they focused on three types of foci:

1. Distal focus – to what extent were you focusing on the ball path?
2. Proximal focus – to what extent were you focusing on adjusting direction with the clubface?
3. Internal focus – to what extent were you focusing on adjusting direction with hand movements and feeling sensation of hand movement?

EEG Recording

Consistent with Kao et al. (2013) and Wang et al. (2020), we used an electro cap (Quik-Caps, Neuroscan, Charlotte, NC, United States) to record the EEG activity and followed the International 10–20 EEG system (Jasper, 1958). The ground electrode was at the FPz site. An average-ear reference offline was taken from the left mastoids (A1) and right mastoids (A2) in 32 scalp locations. We also recorded vertical and horizontal electrooculograms (VEOG and HEOG) in bipolar configurations located superior and inferior to the right eye and on the left and right orbital canthi. All EEG data were recorded with a band-pass filter that was set at 1–100 Hz with the notch filter at 60 Hz. The impedance at each electrode site was below 5 kΩ. The data were obtained at a sampling rate of 1,000 Hz using Neuroscan 4.5 software and stored using Neuroscan NuAmps acquisition amplifiers (Neuroscan, Charlotte, NC, United States).

Procedure

Before the testing day, we asked the participants to not consume any food or beverages containing alcohol or caffeine for 24 h. Before beginning the experiment, the participants were informed of their right to withdraw from the study at any time, and they provided their informed consent to participate. They were then fitted with a Lycra electrode cap, and asked to practice putting from a beginning distance of 300 cm between the hole and ball on the green to calculate their individual putting distances (Chen et al., 2019). We followed that of Chen et al. (2019) and Wang et al. (2020), who adopted an individual task difficulty. It can ensure the same level of difficulty for each performer to control confounding factors. The individual putting distance was determined 40–60% putting success rate (i.e., a challenging condition). Although participants performed an unequal number of putts during warm-up (Mean = 19 ± 10 putts), fatigue and learning effects are not probable because golfers were skilled athletes (Wang et al., 2020). The participants were then shown a list of the action components associated with a distal, a proximal, and an internal focus of attention when putting, and asked for confirmation that those action components were familiar to them. Furthermore, we adopted the same order of conditions for each participant (starting with the NC, then the EC, and finally the IC) because Perkins-Ceccato et al. (2003) who adopted a counterbalance design in attentional focus instructions observed attention order effect (external-internal vs. internal-external). Specifically, the external focus of attention instructions first resulted in lower variability in overall condition than internal instructions first in skilled golfers. They suggested that an external focus of attention instruction first may decrease a potential confounding factor for the results. Furthermore, the participants then performed 40 putts (in four blocks of 10 putts) with each of instructions (NC, EC, and IC) separately. Before each set of 10 trials, the experimenter reminded the participants to adhere to the instruction. After each putt, they were asked to respond to the manipulation check questions on the visual analogue scale. To increase the reliability of the manipulation checks, the participants were also asked to describe precisely what they were focusing on in each condition.

Data Analysis

Behavioral Data

To measure the performance outcome, we calculated the putting success rate per condition as the number of balls holed out of 40 putts. That is, a certain number of balls into the hole divided by the total putts in each condition. For example, if participant putt 10 balls into the hole in first condition, 10 balls divide by 40 balls, so we get 25% putting success rate in first condition. In addition, we followed that of Moore et al. (2012) who adopted mean radial error (MRE) as accuracy data, defined as a subject's average distance between ball after putt and hole in centimeters. Zero was recorded and calculated in MRE on trials where the putt was holed.

EEG Data

Following Semlitsch et al. (1986), an EOG correction was applied to the EEG data to eliminate the effects of blinking. Furthermore, we set a band-pass FIR filter from 1 to 30 Hz with 12 dB/oct for the EEGs and EOG channels. EEG data collected in the 2 s before the putt and containing amplitudes exceeding ± 100 μV were eliminated from subsequent analysis (Kao et al., 2014). Fast Fourier transforms with a Hanning window were used to transform all of the trials for coherence analysis and to maintain minimum spectral leakage. Coherence was defined as $|Cxy(f)|^2$, where:

$$C_{xy}(f) = \frac{\sum_i \{Xi(f) - X(f)\}\{Yi(f) - Y(f)\}^*}{\sqrt{\sum_i |Xi(f) - X(f)|^2 \sum_i |Yi(f) - Y(f)|^2}}$$

and where $Xi(f)$ and $Yi(f)$ represent the Fourier transforms of the time series for electrode sites X and Y, respectively. Coherence was calculated in 1 Hz frequency bins and averaged across the appropriate frequencies to obtain the coherence values for the bandwidths. The electrode pairings of interest were Fz–T7 and Fz–T8 for 10–12 Hz (Ellmers et al., 2016). We applied a Fisher z-transformation to ensure an approximately normal distribution across subjects before conducting the statistical analysis.

Statistical Analysis

SPSS 24.0 software was used for statistical analysis. First, we ran Friedman's ANOVAs to analyze the self-reported manipulation check responses for each of the three conditions. Furthermore, we used Friedman's ANOVAs by ranks to analyze differences in participants who reported attentional strategies for each condition. Second, we ran Friedman's ANOVAs by ranks to analyze the putting success rate and ran a repeated-measures ANOVA over the three conditions for and MRE to evaluate the behavioral performance. Third, we ran a 3 (condition: NC, EC, and IC) × 2 (time: T1 = −2,000 ~ −1,000 ms, T2 = −1,000 ~ 0 ms) × 2 (coherence site: Fz–T7, Fz–T8) repeated-measures ANOVA on the 10–12 Hz EEG data to assess the dynamic changes in neuromotor processes. Fourth, to ensure that only the 10–12 Hz band was altered by attentional focus while the other frequency bands remained the same (i.e., to check for frequency specificity), we analyzed the flanking EEG frequency bands (Cheng et al., 2015). For this purpose, we ran a 3 (condition: NC, EC, and IC) × 2 (time: T1 = −2,000 ~ −1,000 ms, T2 = −1,000 ~ 0 ms) × 2 (coherence site: Fz-T7, Fz-T8) repeated-measures MANOVA on the 4–7 Hz and 16–20 Hz EEG data.

When the ANOVA detected significant effects, we performed *post hoc* calculations of the least significant difference (LSD) and false discovery rate (FDR), with the latter used to control for inflation of the Type I error value due to the multiple comparisons. The alpha level was set at 0.05 for all analyses before FDR (Genovese et al., 2002).

RESULTS

Manipulation Check

There was a significant difference in the extent to which participants reported their attention on the three types of foci during motor preparation in no focus instruction condition, χ^2 (2, N = 12) = 12.809, p = 0.002, in internal focus of attention condition, χ^2 (2, N = 12) = 16.174, p < 0.001, and in external focus of attention condition (EC), χ^2 (2, N = 12) = 11.872, p = 0.003 separately. A follow-up Wilcoxon signed-rank test with FDR corrections revealed that the distal focus had a significantly higher rate than the internal focus (p = 0.005) and the proximal focus (p = 0.002) in the NC. No significant difference was observed between the internal focus and the proximal focus (p = 0.477). In the IC, the internal focus had a significantly higher rate than the distal focus (p = 0.003) and proximal focus (p = 0.002). Again, no significant difference was observed between the distal focus and proximal focus (p = 0.139). In the EC, the proximal focus had a significantly higher rate than the distal focus (p = 0.023) and internal focus (p = 0.002). There was no significant difference between the distal focus and internal focus (p = 0.756; **Figure 1**).

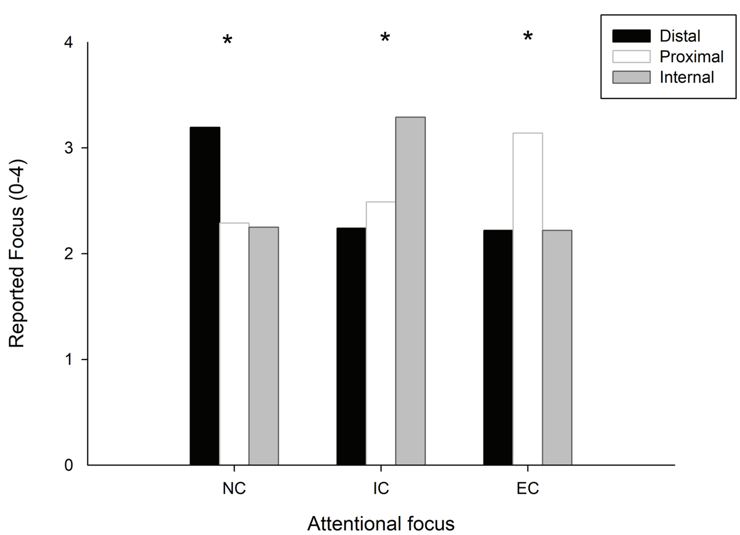

FIGURE 1 | The self-reported manipulation check for each of the three attentional focuses. The participants rate their experience on a five-point Likert scale, ranging from 0 (not at all) to 4 (very much). NC, normal condition; IC, internal focus condition; and EC, external focus condition. *Significant difference, p < 0.05.

Behavioral Results

Putting Success Rate and Accuracy

Friedman's ANOVA by ranks revealed significant differences between the putting success rate (higher is better) in NC (Mdn = 2.67), EC (Mdn = 1.63), and IC (Mdn = 1.71), χ^2 (2, N = 12) = 9.415, p = 0.009. Follow-up pairwise comparisons conducted using Wilcoxon signed ranks tests revealed that the putting success rate in NC was significantly higher than that of EC (p = 0.024 FDR corrected) and IC (p = 0.043 FDR corrected). No statistically significant difference was observed between the EC and IC (p = 0.527 FDR corrected).

In addition, the accuracy data with a repeated-measures ANOVA indicated a significant effect of condition, $F(2,22) = 5.562$, $p = 0.011$, $\eta_p^2 = 0.336$. *Post hoc* analysis indicated that the NC (M = 11.22 ± 4.22 cm) had better accuracy than the EC (M = 16.32 ± 6.63 cm, p = 0.049 FDR corrected) and IC (M = 14.61 ± 5.14 cm, p = 0.036 FDR corrected). No statistically significant difference was observed between the EC and IC (p = 0.20).

EEG Parameters

A 3 (condition: NC, EC, and IC) × 2 (time: T1 = −2,000~−1,000 ms, T2 = −1,000~0 ms) × 2 (coherence site: Fz–T7, Fz–T8) repeated-measures ANOVA for 10–12 Hz showed a significant Condition × Time × Coherence Site interaction, $F(2,22) = 14.349$, $p < 0.001$, $\eta_p^2 = 0.566$, Power = 0.996. A closer look at the simple effect analysis demonstrated a significant interactive effect in Condition × Coherence Site at T1, $F(2,22) = 26.933$, $p < 0.001$, $\eta_p^2 = 0.710$, and at T2, $F(2,22) = 3.917$, $p = 0.034$, $\eta_p^2 = 0.265$. A simple main effect analysis revealed a significant condition effect at Fz–T7 at T1, $F(2,22) = 6.117$, $p = 0.008$, $\eta_p^2 = 0.357$; **Figure 2A**, and at T2, $F(2,22) = 5.417$, $p = 0.012$, $\eta_p^2 = 0.332$; **Figure 2B**. Similarly, a simple main effect analysis revealed a significant condition effect at Fz–T8 at T1, $F(2,22) = 5.316$, $p = 0.013$, $\eta_p^2 = 0.326$. *Post hoc* analyses showed (a) higher Fz–T7 10–12 Hz coherence at T1 in the IC than in the NC (p = 0.024 FDR corrected) and EC (p = 0.043 FDR corrected); (b) lower Fz–T8 10–12 Hz coherence at T1 in the IC than in the EC (p = 0.036 FDR corrected) and NC (p = 0.048 FDR corrected); and (c) lower Fz-T7 10–12 Hz coherence at T2 in the EC than in the NC (p = 0.048 FDR corrected) and IC (p = 0.024 FDR corrected). As expected, skilled golfers in the NC showed dynamic changes in neuromotor processes.

Control Analysis

Frequency Specificity

We compared adjacent frequency bands at Fz–T7 and Fz–T8 during motor preparation. The repeated-measures MANOVA on the 4–7 and 16–20 Hz data found no significant interaction effects: Condition × Time × Coherence site, Wilks' lambda = 0.895, $F(4,8) = 0.235$, $p = 0.911$, $\eta_p^2 = 0.105$; Condition × Time, Wilks' lambda = 0.567, $F(4,8) = 1.528$, $p = 0.282$, $\eta_p^2 = 0.433$; Condition × Coherence Site, Wilks' lambda = 0.850, $F(4,8) = 0.353$, $p = 0.835$, $\eta_p^2 = 0.150$.

Learning and Fatigue Effect on EEG Coherence and Putting Performance

To eliminating the effects of learning and fatigue on brain activity and performance, we followed Ksao et al. (2013). Specifically, artifact-free trials and the putting accuracy data from each participant in each condition were divided into two parts (i.e., first 15 and last 15 putts) and averaged it. We hypothesized that the average Fz–T7 and Fz–T8 alpha 2, and putting accuracy were not significantly different between the first 15 putts and last 15 putts in each condition. First, we run a 2 (Putting Session: Session 1, Session 2) × (Coherence site: Fz–T7, Fz–T8) repeated-measures

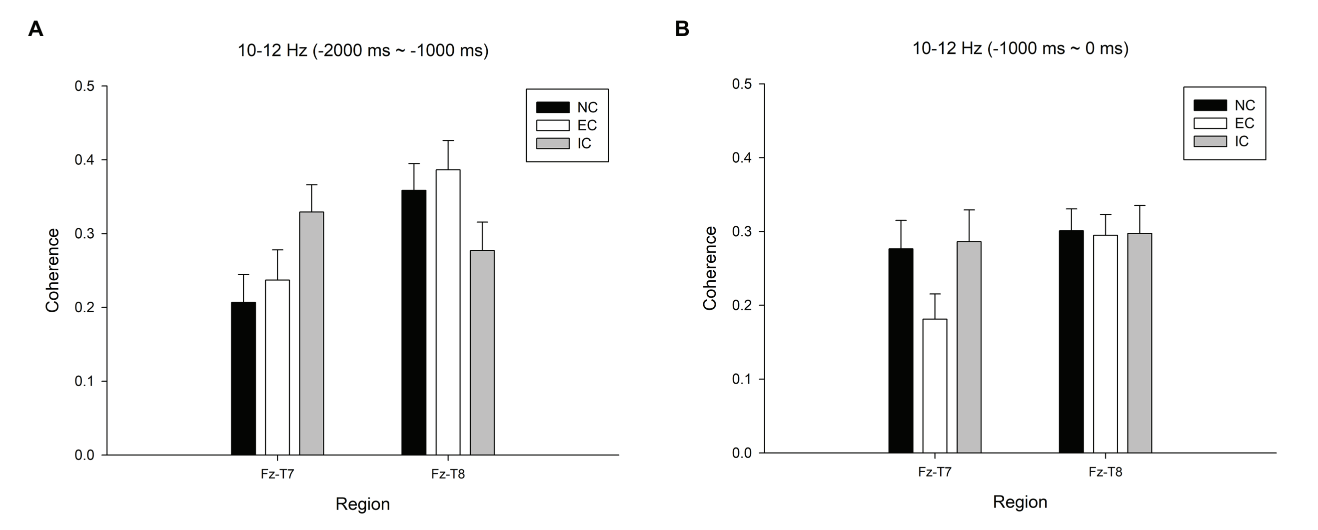

FIGURE 2 | (A) Mean values (±SE) for 10–12 Hz coherence in the three conditions for the Fz-T7 and Fz-T8 electrode pairs in T1 (−2000 ~ −1,000 ms). **(B)** Mean values (±SE) for 10–12 Hz coherence in the three conditions for the Fz-T7 and Fz-T8 electrode pairs in T2 (−1,000 ~ 0 ms).

ANOVA on the 10–12 Hz EEG data in each condition separately. The results indicated no significant Putting Session × Coherence site interaction during the putting task in NC, $F(1,11) = 0.103$, $p = 0.775$, $\eta_p^2 = 0.009$, in EC, $F(1,11) = 0.007$, $p = 0.936$, $\eta_p^2 = 0.001$, and in IC, $F(1,11) = 3.222$, $p = 0.100$, $\eta_p^2 = 0.227$. Second, we run a paired t test in each condition to ensure that the putting performance did not change over time during the putting task in each condition. The results showed that the MRE was not significantly different between the first 15 putts and last 15 putts in the NC ($p = 0.696$), in EC ($p = 0.286$), and in IC ($p = 0.657$). Taken together, these results suggest that skilled golfers in present study did not change 10–12 Hz coherence and did not improve or decrease performance by accumulating trials. Thus, the control analysis eliminated the effects of learning and fatigue.

DISCUSSION

This study examined the dynamic neuromotor processes underlying the nature of attention in skilled golfers. We compared EEG coherence in three different focus conditions (i.e., no instruction condition; NC, external focus condition; EC, and internal focus condition; IC). The findings are consistent with those of previous work (Bernier et al., 2011; Fairbrother et al., 2016; Bahmani et al., 2019) and support the meshed control theory (Christensen et al., 2016). Our results showed that skilled golfers in the NC had similar Fz–T7 and Fz–T8 10–12 Hz coherence as they did in the EC, before switching to a state similar to the IC. Skilled golfers seem to operate with a self-regulated state of attention that optimally combines automatic and controlled processes. That is, superior performance cannot be directly improved by an external or internal focus instruction (Wulf, 2008). In addition, the results showed that adopting an internal focus of attention (i.e., being asked to focus on adjusting direction with hand movements and feeling sensation of hand movement) relative to an external focus of attention (i.e., being asked to focus on the position of the clubface) did not degrade performance, which suggests that skilled golfers use different types of information to stabilize their performance. The results also showed that adopting an external focus generated lower Fz–T7 10–12 Hz coherence, reflecting more automatic processes.

With regard to EEG coherence, lower Fz–T7 10–12 Hz coherence during motor preparation in EC compared to IC supports our hypothesis and corresponds with previous research. Ellmers et al. (2016) observed that novices who adopted an external focus relative to an internal focus had lower Fz–T7 10–12 Hz coherence in a postural task. That study suggested that performers who utilized an external focus could promote their automatic control processes because lower 10–12 Hz coherence at FZ–T7 has been associated with less verbal-analytic processes or language processing in motor planning (Deeny et al., 2003, 2009), which reflects more automatic processes (Cheng et al., 2017; Lo et al., 2019; Wang et al., 2020). However, our findings extend those of Ellmers et al. (2016) by showing that this holds not only for novices but also for skilled athletes. As such, our finding further specifies that external focus could promote their automatic control processes through 10–12 Hz coherence at Fz–T7 (i.e., verbal-analytic processes) in skilled athletes. Interestingly, we additionally found that skilled golfers in the EC had increased Fz–T8 10–12 Hz coherence at T1 (i.e., −2,000 ~ −1,000 ms) before putting. Increased Fz–T8 10–12 Hz coherence is associated with engaging in visuospatial processes in motor planning (Deeny et al., 2003, 2009). Although Fz–T8 10–12 Hz coherence may not be sensitive to changes in attentional focus in a simple voluntary sway task (Ellmers et al., 2016), it could detect changes in complex visuo-motor tasks (i.e., golf putting). Given that golf putting requires complex visuo-motor coordination, it is reasonable to suggest that Fz–T8 10–12 Hz coherence may be associated with visuospatial processing (i.e., a shift in focus on ball trajectory or end point) in motor planning.

Turning to the behavioral results between IC and EC, our findings contrasted with previous studies that have shown negative effects of internal focus in a golf chipping task (Wulf and Su, 2007; Bell and Hardy, 2009). Researchers have suggested that skilled performers' movement control is relatively automatic and thus adopting an internal focus may engage unnecessary information processes, resulting in sub-optimal performance (Wulf et al., 2001). The constrained action hypothesis, an explanation for this phenomenon, assumes that individuals who direct their attention internally interfere with automatic control processes. This interfere constraints on movement, thus resulting in inferior performance (Wulf et al., 2001; Mcnevin et al., 2003). However, our finding is inconsistent with the conventional interpretation. This may be explained by two possible reasons. First, the definition of internal focus in present study is different from the original definition of Wulf et al. (1998). Wulf's definition of internal focus of attention is that individuals direct their attention to control their actions in a relatively conscious movement. In contrast to Wulf's definition of internal focus of attention, the definition of internal focus in present study is that individuals direct their attention to internal awareness on movement. In previous studies, Schücker et al. (2014) and Vitali et al. (2019) observed the attention focus on physical sensations does not disrupt performance. These findings suggested that performers who consciously monitor on their physical sensations did not constraint on the motor system, thus maintaining high-performance effectiveness under challenging conditions (Hanin and Hanina, 2009; Bortoli et al., 2012; Toner and Moran, 2015). Therefore, it is possibly leading to different results in our study. Second, the effects of attentional focus may be modulated by familiarity with attentional focus conditions in skilled performers (Maurer and Munzert, 2013). That is, highly practiced athletes may develop a specific skill-internal focus which does not have a disruptive influence on performance effectiveness (Toner and Moran, 2011). For instance, Maurer and Munzert (2013) observed that placing highly skilled performers in a familiar internal focus condition did not degrade performance effectiveness relative to a familiar external focus condition. Moreover, Wang et al. (2019) and Bertollo et al. (2016) revealed that skilled performers could allocate appropriate degrees of attention to the core components of action (i.e., adjustment of movements) for achieving optimal performance under a challenging task. Given that the present study ensured that skilled golfers were familiar with the internal focus instruction and the manipulation check also ensured that they adhered to internal focus instruction, we suggest that the use of adjusting direction with hand movements or feeling sensation of hand movement may not degrade performance effectiveness compared with adopting the proximal focus in skilled golfers.

In the NC, the data revealed that the skilled golfers most commonly adopted the strategy of focusing on the intended ball path as a distal external focus. Furthermore, the skilled performers operating under the NC showed similar results to EC processes at −2,000 ~ −1,000 ms, before switching to a similar state to that of IC processes just before executing the putt for superior performance. This finding extends previous qualitative studies, which have reported that athletes switched their attentional focus under challenging conditions for superior performance (Bernier et al., 2011; Fairbrother et al., 2016; Bahmani et al., 2019), and supports the meshed control theory. According to the meshed control theory, cognitive control and automatic processes work together to contribute to superior performance in challenging situations (Christensen et al., 2016). Cognitive control processes typically focus on the higher strategic control of the primary skill with its main goals (e.g., focusing on the ball trajectory and the hole when putting) and on the situational control with the control of action in immediate the situation (e.g., adjusting movement in the performance context). Meanwhile, automatic processes typically focus on implementation control that involves performing relatively stable actions (e.g., keeping clubface in the right direction). As such, taking a ball path focus involves visualizing a line from the ball with clubface (similar to an external focus) to the target, and then checking the final position (i.e., making technical adjustments or feeling sensation of the core action component) before putting (Wulf and Su, 2007; Kearney, 2015). There is no surprise in our finding that skilled golfers in the NC had dynamic neuromotor processes (Williams et al., 2015; Oliver et al., 2020). The finding not only supports the meshed control theory but also further specifies the dynamic neuromotor processes underlying the nature of attention in skilled golfers for superior performance.

In terms of implications for coaches and athletes, our findings suggest that no focus instruction in skilled performers may result in a superior cognitive-motor processing and performance when they face challenging situations (Beilock and Gray, 2007; Toner and Moran, 2015). In addition, our EEG results showed that the skilled golfers' attentional processes initially resembled in an external focus of attention and then moved toward an internal focus of attention. We recommend that practitioners should encourage athletes to develop the attentional strategies including the familiarity with an external focus of attention and an internal sensation focus of attention during motor preparation in challenging situations (Bortoli et al., 2012; Bertollo et al., 2013).

Our control analysis showed frequency specificity at Fz–T7 and Fz–T8, and the manipulation check indicated that participants adhered to the instructions. Nonetheless, some limitations should be noted. First, our sample was relatively small compared with previous studies (e.g., Ellmers et al., 2016, $N = 24$; Radlo et al., 2002, $N = 20$) and thus, although our study was sufficiently powered to detect the interaction effects detailed above, the results should be interpreted with caution until they are replicated in a larger sample. Second, to improve the spatial resolution of the EEG, a high-density EEG recording and a source localization algorithm could be used in future studies to confirm the origin of Fz, T7, and T8. Third, to test the meshed control theory more thoroughly, future research should compare tasks at a range of difficulty levels (from easy to highly challenging) and include participants across a range of skill levels (from novices to skilled performers). This approach would provide a clearer overall picture of whether cognitive control processes come to play an increasing role in more highly skilled performers. Fourth, we acknowledge that not manipulating a distal external focus in our instruction may be a limitation in our research because a distal focus of attention is not only about visualizing the trajectory of the ball, but also towards a target (i.e., golf hole). It would be worth studying the differential effect of a distal external focus instruction and the actual type of attentional processes with EEG. It would provide more a comprehensive picture of the underlying mechanisms. Fifth, we adopted the same order of conditions in our study because Perkins-Ceccato et al. (2003) who adopted a counterbalance design observed that the external focus of attention instructions first resulted in lower variability in overall condition than internal instructions first in skilled golfers. This finding raises the concern that the external focus of attention instructions first may decrease a potential confounding factor for the results. However, in another study, Wulf and Su (2007) adopted the counterbalanced design to reduce order effects. These methodological differences may impact on the results. It is recommended that future studies should replicate Perkins-Ceccato et al. (2003) study to examine whether the different order of attentional instruction affects the performance in other precision sport (e.g., golf putting, dart throwing, and archery).

In conclusion, the present study extends previous findings by specifying that skilled performers receiving an external focus of attention instruction had reduced verbal-analytic processes (i.e., more neuromotor supported automatic processes) relative to when they received an internal focus of attention instruction. In addition, adopting an internal focus did not always degrade the performance of skilled performers executing a challenging task relative to adopting an external focus. This indicates that the action-related content of the focus plays a major role. Finally, the present study found that skilled performers receiving no focus instructions first adopted a state similar to external focus processes, which include reduced verbal-analytic and increased visuospatial processes, and then shifted to a state similar to internal focus processes, which include increased verbal-analytic processes, just before putting for superior performance. These findings not only support the meshed control theory but also highlight the neuro-temporal dynamics of these processes.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Research Ethics Committee of National Taiwan Normal University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was financially supported by the Ministry of Science and Technology, Taiwan (MOST 103-2410-H-003-113-MY3).

REFERENCES

Albers, C., and Lakens, D. (2018). When power analyses based on pilot data are biased: inaccurate effect size estimators and follow-up bias. *J. Exp. Soc. Psychol.* 74, 187–195. doi: 10.1016/j.jesp.2017.09.004

Bahmani, M., Bahram, A., Diekfuss, J. A., and Arsham, S. (2019). An expert's mind in action: assessing attentional focus, workload and performance in a dynamic, naturalistic environment. *J. Sports Sci.* 37, 1–13. doi: 10.1080/02640414.2019.1631102

Beilock, S. L., and Gray, R. (2007). "Why do athletes choke under pressure?" in *Handbook of sport psychology*. *3rd Edn.* eds. G. Tenenbaum and R. C. Eklund (Hoboken, NJ: John Wiley and Sons Inc.), 425–444.

Bell, J. J., and Hardy, J. (2009). Effects of attentional focus on skilled performance in golf. *J. Appl. Sport Psychol.* 21, 163–177. doi: 10.1080/10413200902795323

Bernier, M., Codron, R., Thienot, E., and Fournier, J. F. (2011). The attentional focus of expert golfers in training and competition: a naturalistic investigation. *J. Appl. Sport Psychol.* 23, 326–341. doi: 10.1080/10413200.2011.561518

Bertollo, M., Bortoli, L., Gramaccioni, G., Hanin, Y., Comani, S., and Robazza, C. (2013). Behavioural and psychophysiological correlates of athletic performance: a test of the multi-action plan model. *Appl. Psychophysiol. Biofeedback* 38, 91–99. doi: 10.1007/s10484-013-9211-z

Bertollo, M., Di Fronso, S., Filho, E., Conforto, S., Schmid, M., Bortoli, L., et al. (2016). Proficient brain for optimal performance: the MAP model perspective. *PeerJ* 4:e2082. doi: 10.7717/peerj.2082

Bortoli, L., Bertollo, M., Hanin, Y., and Robazza, C. (2012). Striving for excellence: a multi-action plan intervention model for shooters. *Psychol. Sport Exerc.* 13, 693–701. doi: 10.1016/j.psychsport.2012.04.006

Chen, T. T., Wang, K. P., Cheng, M. Y., Chang, Y. T., Huang, C. J., and Hung, T. M. (2019). Impact of emotional and motivational regulation on putting performance: a frontal alpha asymmetry study. *PeerJ* 7:e6777. doi: 10.7717/peerj.6777

Cheng, M. Y., Huang, C. J., Chang, Y. K., Koester, D., Schack, T., and Hung, T. M. (2015). Sensorimotor rhythm neurofeedback enhances golf putting performance. *J. Sport Exerc. Psychol.* 37, 626–636. doi: 10.1123/jsep.2015-0166

Cheng, M. Y., Wang, K. P., Hung, C. L., Tu, Y. L., Huang, C. J., Koester, D., et al. (2017). Higher power of sensorimotor rhythm is associated with better performance in skilled air-pistol shooters. *Psychol. Sport Exerc.* 32, 47–53. doi: 10.1016/j.psychsport.2017.05.007

Christensen, W., Sutton, J., and Mcilwain, D. J. F. (2016). Cognition in skilled action: meshed control and the varieties of skill experience. *Mind Lang.* 31, 37–66. doi: 10.1111/mila.12094

Deeny, S. P., Haufler, A. J., Saffer, M., and Hatfield, B. D. (2009). Electroencephalographic coherence during visuomotor performance: a comparison of cortico-cortical communication in experts and novices. *J. Mot. Behav.* 41, 106–116. doi: 10.3200/JMBR.41.2.106-116

Deeny, S. P., Hillman, C. H., Janelle, C. M., and Hatfield, B. D. (2003). Cortico-cortical communication and superior performance in skilled marksmen: an EEG coherence analysis. *J. Sport Exerc. Psychol.* 25, 188–204. doi: 10.1123/jsep.25.2.188

Ellmers, T. J., Machado, G., Wong, T. W. L., Zhu, F., Williams, A. M., and Young, W. R. (2016). A validation of neural co-activation as a measure of attentional focus in a postural task. *Gait Posture* 50, 229–231. doi: 10.1016/j.gaitpost.2016.09.001

Fairbrother, J. T., Post, P. G., and Whalen, S. J. (2016). Self-reported responses to player profile questions show consistency with the use of complex attentional strategies by expert horseshoe pitchers. *Front. Psychol.* 7:1028. doi: 10.3389/fpsyg.2016.01028

Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* 15, 870–878. doi: 10.1006/nimg.2001.1037

Hanin, Y., and Hanina, M. (2009). Optimization of performance in top-level athletes: an action-focused coping approach. *Int. J. Sports Sci. Coach.* 4, 47–91. doi: 10.1260/1747-9541.4.1.47

Hatfield, B. D. (2018). Brain dynamics and motor behavior: a case for efficiency and refinement for superior performance. *Kinesiol Rev* 7, 42–50. doi: 10.1123/kr.2017-0056

Jasper, H. H. (1958). The ten twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.* 10, 371–375.

Kao, S. C., Huang, C. J., and Hung, T. M. (2013). Frontal midline theta is a specific indicator of optimal attentional engagement during skilled putting performance. *J. Sport Exerc. Psychol.* 35, 470–478. doi: 10.1123/jsep.35.5.470

Kao, S. C., Huang, C. J., and Hung, T. M. (2014). Neurofeedback training reduces frontal midline theta and improves putting performance in expert golfers. *J. Appl. Sport Psychol.* 26, 271–286. doi: 10.1080/10413200.2013.855682

Kearney, P. E. (2015). A distal focus of attention leads to superior performance on a golf putting task. *Int J Sport Exerc Psychol* 13, 371–381. doi: 10.1080/1612197X.2014.993682

Lam, W. K., Masters, R. S. W., and Maxwell, J. P. (2010). Cognitive demands of error processing associated with preparation and execution of a motor skill. *Conscious. Cogn.* 19, 1058–1061. doi: 10.1016/j.concog.2008.11.005

Landers, D. M., Petruzzello, S. J., Salazar, W., Crews, D. J., Kubitz, K. A., Gannon, T. L., et al. (1991). The influence of electrocortical biofeedback on performance in pre-elite archers. *Med. Sci. Sports Exerc.* 23, 123–129. doi: 10.1249/00005768-199101000-00018

Lo, L. C., Hatfield, B. D., Wu, C. T., Chang, C. C., and Hung, T. M. (2019). Elevated state anxiety alters cerebral cortical dynamics and degrades precision cognitive-motor performance. *Sport Exerc. Perform. Psychol.* 8, 21–37. doi: 10.1037/spy0000155

Marchant, D. C., Greig, M., Bullough, J., and Hitchen, D. (2011). Instructions to adopt an external focus enhance muscular endurance. *Res. Q. Exerc. Sport* 82, 466–473. doi: 10.1080/02701367.2011.10599779

Maurer, H., and Munzert, J. (2013). Influence of attentional focus on skilled motor performance: performance decrement under unfamiliar focus conditions. *Hum. Mov. Sci.* 32, 730–740. doi: 10.1016/j.humov.2013.02.001

Mcnevin, N. H., Shea, C. H., and Wulf, G. (2003). Increasing the distance of an external focus of attention enhances learning. *Psychol. Res.* 67, 22–29. doi: 10.1007/s00426-002-0093-6

Moore, L. J., Vine, S. J., Cooke, A., Ring, C., and Wilson, M. R. (2012). Quiet eye training expedites motor learning and aids performance under heightened anxiety: the roles of response programming and external attention. *Psychophysiology* 49, 1005–1015. doi: 10.1111/j.1469-8986.2012.01379.x

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Oliver, A., Mccarthy, P. J., and Burns, L. (2020). Using a "think aloud" protocol to understand meta-attention in club-level golfers. *Int J Sport Exerc Psychol*. 19, 1–14. doi: 10.1080/1612197X.2020.1766536

Perkins-Ceccato, N., Passmore, S. R., and Lee, T. D. (2003). Effects of focus of attention depend on golfers' skill. *J. Sports Sci.* 21, 593–600. doi: 10.1080/0264041031000101980

Radlo, S. J., Steinberg, G. M., Singer, R. N., Barba, D. A., and Melnikov, A. (2002). The influence of an attentional focus strategy on alpha brain wave activity, heart rate and dart-throwing performance. *Int. J. Sport Psychol.* 33, 205–217.

Scharfen, H. E., and Memmert, D. (2019). Measurement of cognitive functions in experts and elite athletes: a meta-analytic review. *Appl. Cogn. Psychol.* 33, 843–860. doi: 10.1002/acp.3526

Schücker, L., Knopf, C., Strauss, B., and Hagemann, N. (2014). An internal focus of attention is not always as bad as its reputation: how specific aspects of internally focused attention do not hinder running efficiency. *J. Sport Exerc. Psychol.* 36, 233–243. doi: 10.1123/jsep.2013-0200

Semlitsch, H. V., Anderer, P., Schuster, P., and Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology* 23, 695–703. doi: 10.1111/j.1469-8986.1986.tb00696.x

Stoate, I., and Wulf, G. (2011). Does the attentional focus adopted by swimmers affect their performance? *Int. J. Sports Sci. Coach.* 6, 99–108. doi: 10.1260/1747-9541.6.1.99

Swann, C., Moran, A., and Piggott, D. (2015). Defining elite athletes: issues in the study of expert performance in sport psychology. *Psychol. Sport Exerc.* 16, 3–14. doi: 10.1016/j.psychsport.2014.07.004

Toner, J., and Moran, A. (2011). The effects of conscious processing on golf putting proficiency and kinematics. *J. Sports Sci.* 29, 673–683. doi: 10.1080/02640414.2011.553964

Toner, J., and Moran, A. (2015). Enhancing performance proficiency at the expert level: considering the role of 'somaesthetic awareness'. *Psychol. Sport Exerc.* 16, 110–117. doi: 10.1016/j.psychsport.2014.07.006

United States Golf Association (2018). Handicap index statistics. Available at: https://www.usga.org/content/usga/home-page/handicapping/handicapping-stats.html (Accessed March 01, 2019).

Vitali, F., Tarperi, C., Cristini, J., Rinaldi, A., Zelli, A., Lucidi, F., et al. (2019). Action monitoring through external or internal focus of attention does not impair endurance performance. *Front. Psychol.* 10:535. doi: 10.3389/fpsyg.2019.00535

Wang, K. P., Cheng, M. Y., Chen, T. T., Chang, Y. K., Huang, C. J., Feng, J., et al. (2019). Experts' successful psychomotor performance was characterized by effective switch of motor and attentional control. *Psychol. Sport Exerc.* 43, 374–379. doi: 10.1016/j.psychsport.2019.04.006

Wang, K. P., Cheng, M. Y., Chen, T. T., Huang, C. J., Schack, T., and Hung, T. M. (2020). Elite golfers are characterized by psychomotor refinement in cognitive-motor processes. *Psychol. Sport Exerc.* 50:101739. doi: 10.1016/j.psychsport.2020.101739

Williams, J. M., Nideffer, R. M., Wilson, V. E., and Sagal, M. S. (2015). "Concentration and strategies for controlling it" in *Applied sport psychology: Personal growth to peak performance*. *7th Edn.* eds. J. M. Williams and V. Krane (McGraw Hill: New York), 304–325.

Wulf, G. (2008). Attentional focus effects in balance acrobats. *Res. Q. Exerc. Sport* 79, 319–325. doi: 10.1080/02701367.2008.10599495

Wulf, G., Höß, M., and Prinz, W. (1998). Instructions for motor learning: differential effects of internal versus external focus of attention. *J. Mot. Behav.* 30, 169–179. doi: 10.1080/00222899809601334

Wulf, G., and Lewthwaite, R. (2016). Optimizing performance through intrinsic motivation and attention for learning: the OPTIMAL theory of motor learning. *Psychon. Bull. Rev.* 23, 1382–1414. doi: 10.3758/s13423-015-0999-9

Wulf, G., Mcnevin, N., and Shea, C. H. (2001). The automaticity of complex motor skill learning as a function of attentional focus. *Q. J. Exp. Psychol.* 54, 1143–1154. doi: 10.1080/713756012

Wulf, G., and Prinz, W. (2001). Directing attention to movement effects enhances learning: a review. *Psychon. Bull. Rev.* 8, 648–660. doi: 10.3758/BF03196201

Wulf, G., and Su, J. (2007). An external focus of attention enhances golf shot accuracy in beginners and experts. *Res. Q. Exerc. Sport* 78, 384–389. doi: 10.1080/02701367.2007.10599436

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 21 December 2020
doi: 10.3389/fpsyg.2020.597913

Cognitive Interaction Technology in Sport—Improving Performance by Individualized Diagnostics and Error Prediction

Benjamin Strenge[1], Dirk Koester[2] and Thomas Schack[1]*

[1] *Neurocognition and Action Group, Faculty of Psychology and Sports Science, Center for Cognitive Interaction Technology (CITEC), Bielefeld University, Bielefeld, Germany,* [2] *Sport Psychology, Faculty Business and Management, BSP Business School Berlin, Berlin, Germany*

OPEN ACCESS

Edited by:
Edson Filho,
Boston University, United States

Reviewed by:
Itay Basevitch,
Anglia Ruskin University,
United Kingdom
Nataniel Boiangin,
Barry University, United States

***Correspondence:**
Benjamin Strenge
benjamin.strenge@uni-bielefeld.de

Specialty section:
This article was submitted to Movement Science and Sport Psychology, a section of the journal Frontiers in Psychology

Received: *22 August 2020*
Accepted: *27 November 2020*
Published: *21 December 2020*

Citation:
Strenge B, Koester D and Schack T (2020) Cognitive Interaction Technology in Sport—Improving Performance by Individualized Diagnostics and Error Prediction. Front. Psychol. 11:597913. doi: 10.3389/fpsyg.2020.597913

The interdisciplinary research area Cognitive Interaction Technology (CIT) aims to understand and support interactions between human users and other elements of socio-technical systems. Important reasons for the new interest in understanding CIT in sport psychology are the impressive development of cognitive robotics and advanced technologies such as virtual or augmented reality systems, cognitive glasses or neurotechnology settings. The present article outlines this area of research, addresses ethical issues, and presents an empirical study in the context of a new measurement and assessment system for training in karate. Recent advances in the field of cognitive assistance systems enabled largely automatized assessments of individual mental representation structures for action sequences, such as choreographed movement patterns in dance or martial arts. Empirical investigations with karate practitioners of different skill levels demonstrate that advanced software-based survey and algorithmic analysis procedures based on cognitive models generate individualized performance predictions for a movement sequence from the *Kanku-dai kata* (a pre-defined karate movement sequence), which correlated significantly not only with formal expertise (*kyu/dan* rank) but also with the actual likelihood of mistakes in action execution. This information could prospectively be used to define individual training goals for deliberate practice and incorporated into cognitive interaction technology to provide appropriate feedback. We argue that the development of cognitive interaction systems for sport should explicitly take ethical issues into consideration and present a particular developed engineering approach. The potential benefits of such an assistance system for intermediate and advanced practitioners include more effective and flexible practice, as well as supportive effects, and more flexible training schedules. Furthermore, we argue that researchers from the field of sport psychology can benefit from advances in technological systems that enhance the understanding of mental and motor control in skilled voluntary action.

Keywords: karate athletes/performance, karate kata, SDA-M, ethical issue recognition, mental representation structures, cognitive assistance systems

1. INTRODUCTION

For over a decade numerous researchers from psychology, computer science, engineering, biology, linguistics, and sports science shaped the interdisciplinary field of Cognitive Interaction Technology (CIT) in order to establish the scientific and technological basics for creating systems that are capable of interacting at different levels of cognitive complexity (Ritter and Sagerer, 2009). Pursuing the vision of intuitive, human-friendly technology that adapts to users' needs (Wachsmuth et al., 2012) by offering intuitive and personalized support in daily routines (Wrede et al., 2017), CIT comprises research topics such as motion intelligence, attentive systems, situated communication, memory and learning (Wachsmuth, 2008; Schack and Ritter, 2013). A major goal is *"to develop memory systems that can approximate some of the key features of human memory, such as flexible association, scalability and learning at different levels"* (Ritter, 2010, p. 230). While classic artificial intelligence concentrates on modeling the mind, CIT research focuses more on interactions that take place in the physical world (Wachsmuth et al., 2012) and combines algorithmic approaches with insights from analyses of human and animal motion to establish "a coherent picture about the internal representation of our movement abilities" (Ritter, 2010, p. 230). On the technical side, CIT combines visualization, sonification, haptic, and augmented reality devices, motion capture, simulated agents in virtual worlds, and attentive user interfaces in novel ways (Ritter, 2010). This led to a broad range of technological advancements such as embodied anthropomorphic robots that can aid humans (Ritter, 2010; Wachsmuth et al., 2012), intelligent glasses for cognitive assistance (Essig et al., 2016), and smart environments systems with mobile service robots for ambient assisted living (Wrede et al., 2017).

Sport psychology researchers and practitioners have been traditionally concerned with topics like analyzing and improving human performance but started to develop new technologies (e.g., Schack and Ritter, 2013) to support sport performance several years ago, e.g., using motion tracking, eye tracking, heart-rate variability or EEG measurements to provide biofeedback with sonification, virtual and augmented reality systems (see e.g., Schack et al., 2014a, 2020; Hagan et al., 2018, for overviews). A main question is how to inform assistive technologies about the cognitive background (memory) and motion intelligence (motor skills) of the user. From a traditional cognitive psychology perspective (see Anderson, 2010), the development of human expertise is commonly characterized by *proceduralization*: The learner integrates declarative knowledge into procedural rule sets so that less declarative memory needs to be used, which reduces brain activation in areas like the hippocampus, prefrontal cortex, and anterior cingulate, and decreases latency. Fitts and Posner (1967) famously described this process as a three-stage model, which transitions from an initial "cognitive stage" to an intermediate "associative stage" and terminates in the "autonomous stage." Research has also found that, while potential performance improvements are limited by factors like musculature and age, the time required for cognitive processing may converge against zero as a power function of practice (Anderson, 2010). This characterization of human expertise development has been challenged by the sport psychological theory of *deliberate practice*, which means engaging in training that focuses on improving specific tasks and involves providing immediate feedback, time for problem-solving and evaluation, and opportunities for repeated performance in order to refine behavior (Ericsson, 2008). This obviously requires that practitioners are given specific tasks with well-defined goals (Ericsson, 2007). Purportedly, deliberate practice continually improves performance, because *"expert performers counteract automaticity by developing increasingly complex mental representations to attain higher levels of control of their performance and will therefore remain within the cognitive and associative phases"* (Ericsson, 2008, p. 991).

Based on a Cognitive Action Architecture Approach (CAA-A), sport psychology researchers described the building blocks and levels of the action system that enable us to control movements such as striking the tennis ball at the right time, or coordinating steps and arm movements in dancing or golf, and demonstrated how the measurement of mental representation can be used for applied work in sport, new pathways in mental training (imagery), and to inform technical systems (Tenenbaum et al., 2009; Frank et al., 2014; Schack, 2020). A highly promising application of interactive technology in sport psychology is to provide helpful assistance to athletes in the context of learning. In coaching, trainees' capabilities to respond to an expert's assistance and the coaching system's ability to activate users' learning potential can be observed (Schack, 2020). Coaching a trainee at different interaction levels while practicing and learning a motor task constitutes an interesting scenario not only for supporting motor learning processes but also to understand the effectiveness of current coaching principles (see also Schack, 2020). Based on mental representation analyses in sport (Schack and Mechsner, 2006; Schack and Hackfort, 2007; Schack, 2020), we investigate how coaching could become more individualized and adaptive in the real world and in Virtual or Augmented Reality settings (Schack et al., 2020). To this extent, it is clearly advantageous for a real or virtual coach to know how mental structures form, stabilize, and change in sports (Schack, 2020). Coaches who possesses such knowledge are better able to address the individual athlete on his or her current level of learning and shape instructions to improve training and performance (Schack, 2020).

In this line of research, numerous studies found that the differing mental representation structures of experts and novices can be measured with the *"structural-dimensional analysis of mental representations"* (SDA-M) method (Schack, 2012) and influenced by appropriate training (e.g., Heinen et al., 2002; Schack, 2004; Schack and Mechsner, 2006; Schack and Hackfort, 2007; Frank et al., 2013, 2014; Schack et al., 2014b). A methodological review and evaluation of research in expert performance in sport by Hodges et al. (2007, p. 164) noted that the SDA-M method *"is expected to aid in our understanding of the usually non-declarative motor representations underlying expert performance in fast, complex coordinative actions and in identifying the problems novices encounter in understanding motor problems."* Recently, Strenge et al. (2019) described advanced algorithms for automatized analyses of task-related mental representation structures based on SDA-M related to action sequences. These algorithmic approaches might be useful as a component of future CIT systems, like cognitive glasses, to measure and improve human performance in sport. In this context, SDA-M and its recent algorithmic extensions could serve as a measurement and assessment tool, and smart glasses or other portable devices could provide corresponding feedback for deliberate practice.

The present study reports on an empiric study in karate as a proof of concept for this assessment approach. To this end, the article first recapitulates the SDA-M method and its algorithmic extensions and then describes the specific study-related methods. Subsequently, potential ethical benefits and risks, as well as links to ethical aspects of technical system development, are discussed.

2. RETRIEVAL AND ANALYSIS OF MENTAL REPRESENTATION STRUCTURES WITH SDA-M

The SDA-M method can be used to analyze human memory structures with respect to a specified set of items (e.g., basic actions in sports). SDA-M consists of several survey and analysis steps, which are briefly outlined in the following. The theoretical, methodical and algorithmic foundations of SDA-M have been described in detail by Schack (2012), and Strenge et al. (2019) presented and exemplified recent algorithmic extensions for automatic assessment of SDA-M data concerning individual likelihoods of errors in action sequences.

2.1. Task Analysis

In a preparatory step, it is generally important to understand the motor task (here: a karate movement sequence) and characterize its task-adequate functional organization, e.g., in cooperation with athletes of different levels of expertise or coaches. The activity is hereby split into "basic action concepts" (BACs, see Schack, 2012), which are represented by textual descriptions and/or images. This can be done by researchers with the help of a functional movement analysis (Hossner et al., 2015) and together with domain experts to establish a plausible and workable set of BACs.

2.2. Step A: Split Procedure and Distance Scaling

During the split procedure these action items (BACs) are shown to study participants on a screen using specialized software such as the *QSplit* SDA-M tool (see **Figure 1**). The n actions of the analyzed task or activity are presented in random order as reference objects or "targets," and all $n-1$ other actions are then compared to the current target (also in random order). For each pair of actions the participant must decide whether or not these are directly associated during execution of the analyzed activity (e.g., a movement sequence). The SDA-M software then calculates correlation and distance values between all pairs of actions.

2.3. Step B: Hierarchical Clustering and Visualization

The results from step A can be used to create a hierarchical agglomerative average-linkage clustering of the actions (BACs). The SDA-M software visualizes this clustering with a dendrogram to enable manual assessment of participants' mental representation structures (see **Figures 2**, **3** for examples from the present study). For many SDA-M applications this is the last necessary analysis step (see e.g., Heinen and Schwaiger, 2002; Heinen et al., 2002; Heinen and Schack, 2004; Schack, 2004; Schack and Hackfort, 2007). Further steps like investigating the feature dimensions of the representation or invariance measures are possible (Schack, 2012).

2.4. Step C: Automatized Algorithmic Assessment of SDA-M Data

Two different algorithmic approaches for predicting human error based on SDA-M data have been developed and presented by Strenge et al. (2019): *Analysis of Most Probable Actions* (AMPA) and *Correct Action Selection Probability Analysis* (CASPA). These new algorithmic approaches automatized the process of assessing memory structures based on SDA-M data to predict probable errors in action sequences, which eliminated the previous need for manual assessments using dendrograms.

The basic approach of AMPA is to determine if the set of actions that have lowest distance to the previously executed action, which corresponds to strongest association, contains a correct follow-up action. This results in a simple binary assessment $\in \{0, 1\}$ for each action a_i with $i \in \{1, ..., n-1\}$ that indicates whether participants would be able to select a correct action for the next step or not. The more sophisticated CASPA algorithm is based on parts of the computational cognitive architecture "Adaptive control of thought–rational" (ACT-R) (Anderson and Lebiere, 1998; Anderson et al., 2004). CASPA uses the same individual SDA-M data as AMPA but outputs a continuous measure $p_i \in [0, 1]$ to estimate the probability of correct action selection after action a_i. Arbitrary thresholds for p_i can be used to decide if assistance will be needed. This binarized output of CASPA is referred to as CASPA_d when using a default threshold of 0.5, whereas usage of an empirically informed task-specific threshold is denoted as CASPA_i.

Strenge et al. (2019) speculated that the results of AMPA or CASPA could be used by technical systems like intelligent glasses to provide anticipatory action support.

3. METHODS

Karate practitioners of different skill levels were analyzed regarding a choreographed sequence of distinct movements (karate techniques) from the beginning of the so-called *Kanku-dai kata*. These karate techniques include blocks and strikes using one's hands or feet, and occasionally both at the same time, delivered from specific stances. While executing a *kata* movement sequence, technique executions may be accompanied by hip rotations and/or a transition to another stance, thus effectively making them full-body movements

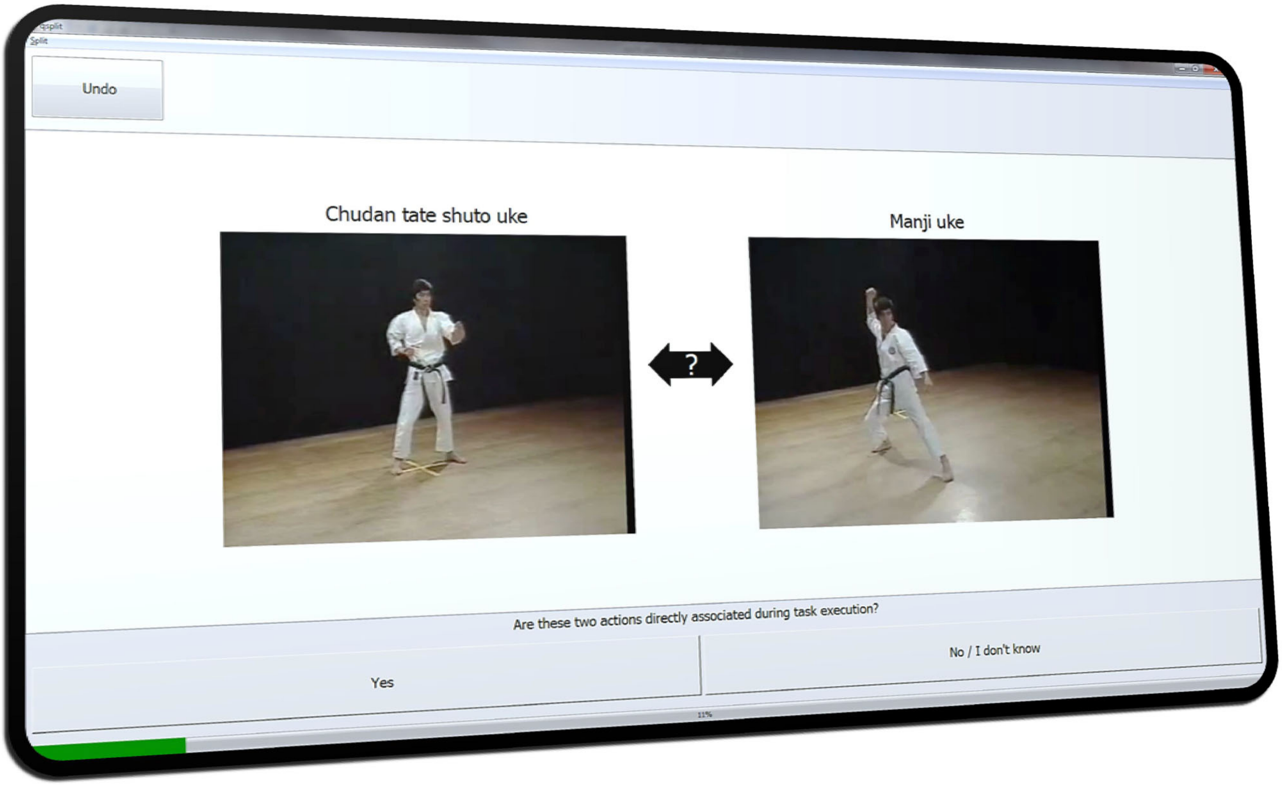

FIGURE 1 | User interface of the QSplit software for the SDA-M split procedure showing two basic actions (karate blocking techniques) from the *Kanku-dai* movement sequence.

despite the fact that they are commonly denoted just by the type of block or strike. For example, each successive *shuto uke* ("knife hand block") in the *Kanku-dai* sequence implicitly includes moving one step forward in a *kokutsu dachi* (a defensive back stance where most body weight rests on the rear leg) switching the front and rear leg on each step (see **Figure 4**). Instructors of the popular *Shotokan* style of karate commonly introduce the *Kanku-dai* at some point during students' preparation for the first *dan* black belt or "master" level. The *Kanku-dai kata* can be understood as a long compilation and rearrangement of subsequences from preliminary *katas*, especially the so-called *Bassai-dai* and *Heian katas*, which should be well-known by then. Therefore, most intermediate practitioners supposedly possess extensive experience with some or all of the preliminary *katas* but have limited, if any, knowledge of the *Kanku-dai*. Even advanced practitioners might commonly fall prey to memory interference effects due to wrong matching and association of the corresponding movement patterns. This constitutes an interesting and challenging scope of application for analyzing mental representation structures, error prediction and performance assessment. The study focused on the first 17 moves from the beginning of *Kanku-dai* up to the first *manji uke* blocking technique:

1. *Morote age shuto uke* (rising knife hand block with both hands)
2. *Chudan shuto uchi* (inside knife hand strike)
3. *Jodan haishu uke* with left arm (back hand block at head height)
4. *Jodan haishu uke* with right arm (back hand block at head height)
5. *Chudan tate shuto uke* (inside vertical knife hand block)
6. *Chudan choku zuki* with right arm (straight punch at middle level)
7. *Uchi uke* with right arm (forearm block)
8. *Chudan choku zuki* with left arm (straight punch at middle level)
9. *Uchi uke* with left arm (forearm block)
10. *Jodan uraken uchi* + *yoko geri* (back fist strike at head height + side kick)
11. 1st *shuto uke* (left arm) (knife hand block)
12. 2nd *shuto uke* (right arm) (knife hand block)
13. 3rd *shuto uke* (left arm) (knife hand block)
14. *Chudan gohon nukite* (five-finger spear hand strike at middle level)
15. *Age shuto uke* (rising knife hand block)
16. *Jodan mae geri* (front kick at head height)
17. *Manji uke* (swastika-shaped double block).

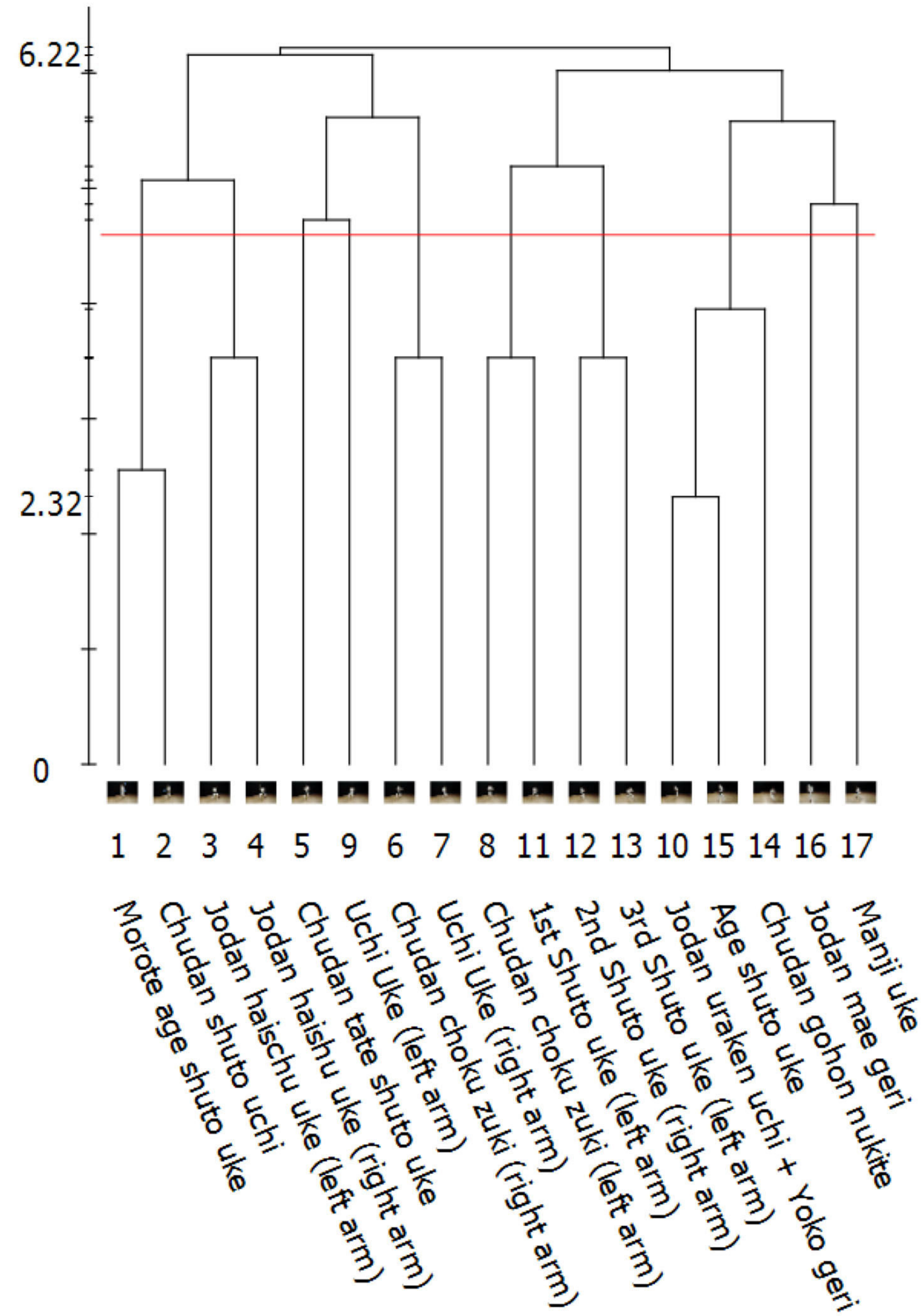

FIGURE 2 | Visualization of a 5th *kyu* (blue belt) karate practitioner's individual mental representation structure related to the *Kanku-dai* movement sequence by an SDA-M dendrogram. Numbers below the dendrogram indicate designated positions of each action within the sequence.

3.1. Statement of Ethical Approval

The study has been approved by the ethics committee of Bielefeld University in written form according to the guidelines of the German Psychological Society (DGPs) and the Association of German Professional Psychologists (BDP). All participants gave informed and written consent to participate in the study.

3.2. Participants

Twelve individuals between 18 and 63 years with a mean age of 30.7 years ($SD = 13.3$) participated in the study. The majority (75%) of participants were male. Some basic experience in karate, as indicated by holding at least the sixth *kyu* rank ("green belt"), was required to enable proper determination of individual techniques. This was necessary since

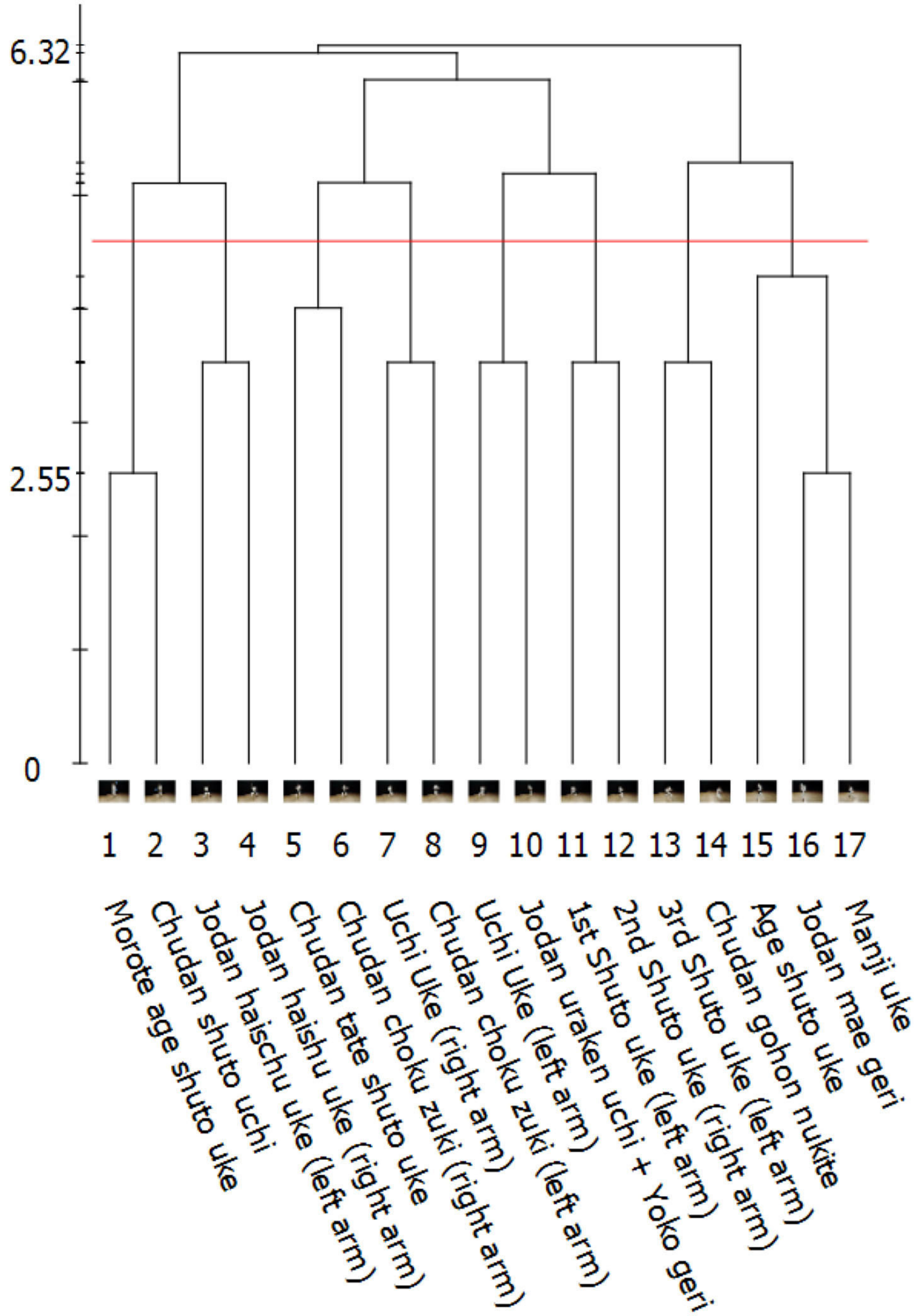

FIGURE 3 | Visualization of a 2nd *dan* (black belt) karate practitioner's individual mental representation structure related to the *Kanku-dai* movement sequence by an SDA-M dendrogram. Numbers below the dendrogram indicate designated positions of each action within the sequence.

the SDA-M-based analyses in this study were concerned with action selection mechanisms for choosing between different karate techniques within the *kata* sequence. The cognitive action architecture model allocates these mechanisms primarily to the level of "mental control" and the associated "basic action concepts" (BACs) as mental representation units (Schack, 2004). The corresponding SDA-M-based analyses in this study were inherently and deliberately indifferent to the quality of individual karate techniques. Therefore, participants had to know and apply these BACs, i.e., execute karate techniques, sufficiently well to allow the experimenter to properly and unambiguously recognize and distinguish them. **Table 1** shows the exact distribution of participant numbers across formal ranks of expertise. They were reimbursed for their time with 5 Euros in cash.

FIGURE 4 | Participant executing a *shuto uke* ("knife hand block") karate technique from the *kanku-dai kata* movement sequence.

TABLE 1 | Formal expertise of participants in karate.

Rank	6th *kyu*	5th	4th	3rd	2nd	1st *kyu*	1st *dan*	2nd *dan*
No. of participants	2	3	1	1	1	0	1	3

Note that expertise increases from left to right, because kyu ranks traditionally decrement from eighth (beginner) to first kyu (advanced student), whereas the subsequent dan ranks ("master level") are counted upwards from 1st dan.

3.3. Procedure

First, participants were welcomed, asked to give informed consent to participation, and provide demographic data, as well as their degree of formal expertise in karate. The following proceedings of each trial could be divided into three consecutive phases:

3.3.1. Recapitulation and Learning Phase

A brief recapitulation of preliminary *katas* served both as a physical warm-up and cognitive trigger for activating relevant memory structures. This included the *Heian Nidan*, *Heian Yondan*, and *Bassai-dai*, which contain similar or identical parts as *Kanku-dai*, each from beginning until the first occurrence of a *kiai*.[1] Participants who had already been tested in a given *kata* as part of an official examination for their *kyu* or *dan* grade were merely asked to demonstrate it once, in a calm and serene manner, without further guidance. The remaining preliminary *katas* were at least once roughly synchronously executed by the participant and the experimenter as an instructor. If participants made mistakes or struggled noticeably the execution was repeated up to two times. Afterwards, a video was shown of the *Kanku-dai* sequence performed by Master Hirokazu Kanazawa (10th *dan* black belt; †8 December 2019). Participants were then rudimentarily taught to execute this sequence by following the moves in rough synchrony with the experimenter. The number of repetitions depended on formal expertise ranks: Relative beginners (eighth to fifth *kyu*) executed the sequence twice, advanced students (fourth to first *kyu*) executed it once, and black belts did no physical execution at all. The video of the *Kanku-dai* sequence was then shown a second time to finalize the learning phase.

[1]The *kiai* is a short shout that is uttered when performing distinct moves in karate. The correct execution of *katas* usually requires *kiais* at certain specified points.

3.3.2. SDA-M Introduction and Split Procedure

The SDA-M split procedure was explained by showing participants a special tutorial video included in the QSplit software, which specifies the instructions as follows (translated from German to English):

> *The software shows representations of two action steps. You shall judge whether these action steps are sequentially "directly associated" during task execution or not, i.e., whether they are executed immediately before or after one another. It does not matter which action step is shown on the left or on the right side of the screen.*

The tutorial video continues to illustrate the implications of these instructions using, as a simple example from daily life, an action sequence for toasting white bread slices and the respective decisions in a corresponding split procedure. Participants were asked to confirm whether they had understood these general instructions. After this, they were subjected to an SDA-M split procedure, which incorporated still images of the first 17 techniques of the *Kanku-dai kata* and corresponding textual descriptions. As usual in karate, Japanese terms were used to denote the techniques. These can be seen in correct order at the bottom of **Figure 3**. The user interface of QSplit, which was used for this split procedure, is shown in **Figure 1**.

3.3.3. Movement Sequence Execution Test

Lastly, participants' capability to freely execute the *Kanku-dai* movement sequence was tested. Participants started the *kata* with their back toward the experimenter, so they could not see the experimenter during the movement sequence execution. The experimenter observed the execution and intervened when errors occurred. In this case the experimenter told participants to freeze in their current position, walked in front of them, and demonstrated the correct technique. Participants should then reverse their previous (wrong) action and continue with the correct execution. This intervention procedure was beforehand explained and demonstrated. Importantly, merely slightly inaccurate action executions were ignored as long as the correct technique was still clearly recognizable. Only wrongly chosen techniques were counted as errors and corrected.

3.4. Data Analysis

All SDA-M procedures were executed with the *QSplit SDA-M Suite* v1.6 for Windows. This included the split procedure and the usual data normalization, scaling, clustering and invariance analysis steps as described by Schack (2012), as well as advanced analyses using the AMPA and CASPA algorithms (see Strenge et al., 2019). Generally, the available data were analyzed on two different levels:

First, on the level of individual karate techniques, the algorithmic predictions by AMPA, CASPA_d and CASPA_i for each action of every participant were compared with the corresponding outcomes during actual execution. For this purpose, several standard metrics for the evaluation of binary classifiers were used. In this context a "true positive" case was counted when the algorithmic analysis predicted an error and this error actually occurred.

Second, participants overall performances, i.e., total numbers of correct actions, and their formal expertise ranks were compared with different SDA-M-based measures, which aim to reflect the overall suitability of individual mental representation structures for the movement task. One of these measures stemmed from Lander and Lange (1992) and Schack (2012), who proposed the structural invariance measure λ. Let the sets S_a and S_b represent the outcomes of SDA-M's hierarchical agglomerative average-linkage clustering for participant a and participant b, which contain the clusters $C_i \in S_a$ and $C_j \in S_b$ of BACs (here: karate techniques). The invariance of the mental representation structures of participants a and b is then defined as follows:

$$\lambda_{a,b} = \sqrt{\frac{\min(|S_a|, |S_b|)}{\max(|S_a|, |S_b|)} \cdot \frac{\sum_{i=1}^{|S_a|} \sum_{j=1}^{|S_b|} |C_i \cap C_j|}{\sum_{i=1}^{|S_a|} \sum_{j=1}^{|S_b|} \sqrt{|C_i| \cdot |C_j|}}}; \lambda_{a,b} \in [0, 1] \tag{1}$$

More recently, the Adjusted Rand Index (ARI) gained popularity among SDA-M researchers for measuring the similarity of two participants' mental representation structures (see e.g., Frank et al., 2013, 2014, 2016; Land et al., 2014; Jeraj et al., 2017; Kim et al., 2017; Meier et al., 2020). The ARI is bounded above by a maximum of 1 and takes on negative values (with no well-defined lower bound) when similarity falls below the expected value from random clustering with the same number of clusters and elements in each (Hubert and Arabie, 1985). Note that both the invariance measure λ and the ARI are based on SDA-M clustering results. This implies they require a reference structure for comparison, e.g., from one or multiple domain experts. In the present study, an ideal reference structure for this purpose was established by perfectly associating the action representations that exactly precede or follow each other in the movement sequence.

In addition to these two established measures (λ and ARI), CASPA_m is newly introduced as an advanced alternative. It represents the arithmetic mean over all likelihoods of successful action selection during the whole sequence of movements as predicted for an individual by the CASPA algorithm. Formally, if n is the number of actions in the designated action sequence (here: $n = 17$) and p_i the probability of correct action selection for a given participant after executing a previous action a_i as estimated by CASPA, then CASPA_m is defined as follows:

$$\text{CASPA}_m := \frac{1}{n-1} \sum_{i=1}^{n-1} p_i; \text{CASPA}_m \in [0, 1] \tag{2}$$

This value can also be interpreted as an overall estimate of the expected probability of correct action selection for a randomly chosen situation within the sequence.[2] CASPA_m has the advantage over previous alternatives (the invariance λ and ARI) that it does not require an explicit reference structure.

[2]This formulation assumes that there is only one correct action sequence to achieve the goal. The case of multiple different correct sequences would require slightly more complex calculations, involving a weighted arithmetic mean that weights the estimated probabilities of correct action selection for each situation with the joint probabilities of having previously chosen exactly the actions needed to get into that situation.

$CASPA_m$ also inherits a notable limitation of the CASPA algorithm though: It is only applicable to SDA-M data sets related to action sequences that have no temporal overlap between the actions. Therefore, it cannot generally replace λ and ARI for arbitrary SDA-M application scenarios if this condition is not satisfied.

4. RESULTS

Substantial, albeit imperfect, matches between algorithmic analyses of participants' mental representation structures and their actual accomplishments while executing the movement sequence were found.

Detailed metrics for the performance of AMPA, $CASPA_d$ (using the default threshold of 0.5), and $CASPA_i$ (using an empirically informed threshold of 0.6207), with respect to predicting participants individual likelihood of making mistakes at the level of each individual action (i.e., discrete karate techniques) are shown in **Table 2**. An overall relatively low prevalence of errors (31 errors in a total of 192 actions $\Rightarrow$ $P(\text{error}) \approx 16\%$) caused a salient discrepancy between positive and negative predictive values (PPVs $\in [0.29, 0.31]$ and NPVs $\in [0.90, 0.94]$). However, the prevalence-independent measures of sensitivity (values $\in [0.55, 0.77]$) and specificity (values $\in [0.67, 0.75]$) were rather close to each other. From an applied perspective sensitivity matters for recognizing as many of the practitioners weak points as possible, whereas specificity helps focusing on these issues instead of unnecessarily practicing parts they already mastered. The $CASPA_i$ algorithm achieved the best results among the different algorithmic variants in terms of balanced accuracy (value 0.72; see **Figure 5**), which represents the arithmetic mean of sensitivity and specificity values (Brodersen et al., 2010).

Table 3 shows the correlations (using Spearman's rank-order correlation coefficient ρ) between participants' formal expertise (*kyu*/*dan* rank), their actual performance in the *Kanku-dai kata* execution test (i.e., number of correctly chosen techniques), the conventional SDA-M measures for assessing the invariance and similarity of individual mental representation structures to an ideal reference structure (Lander's λ and ARI)[3], and the newly proposed $CASPA_m$ measure. All three SDA-M-based assessment metrics showed significant and strong positive correlations with participants actual performances ($CASPA_m : \rho = 0.88, p < 0.001$; Lander's $\lambda : \rho = 0.79, p < 0.01$, ARI: $\rho = 0.65, p < 0.05$). $CASPA_m$ and Lander's λ also correlated significantly and strongly with formal expertise ranks ($CASPA_m : \rho = 0.80, p < 0.01$; Lander's $\lambda : \rho = 0.66, p < 0.05$). The differences between Lander's λ, ARI, and $CASPA_m$'s correlations with performance and expertise were statistically insignificant (using Fisher z-transformation for comparison of the correlation coefficients). However, $CASPA_m$ descriptively showed the strongest correlations with performance and expertise among all three SDA-M-based metrics, and also showed higher correlations with actual performance than formal expertise ranks did ($CASPA_m : \rho = 0.88$ vs. formal expertise: $\rho = 0.84$).

On a sidenote, we recognized that the two most error-prone steps occured when transitioning from the 9th action (*uchi uke* block with left arm) to the 10th action (*jodan uraken uchi* back fist strike with *yoko geri* sidekick) with 5/12 errors and from the 16th action (*jodan mae geri* front kick) to the 17th action (*manji uke* block) with 6/12 errors. In both of these cases, the preceding techniques marked the end of corresponding movement sub-sequences known to some participants from preliminary katas, which at the respective point would go on with other techniques than the tested *kanku-dai* kata demands. Participants up to the 3rd *kyu* grade (brown belt) made errors at these points, which corroborates the supposition that memory interference effects may play an important role in learning and distinguishing these katas.

5. DISCUSSION AND ETHICAL CONSIDERATIONS

Deliberate practice has generally been accepted as an important factor for developing expertise, especially in sports, even though the specific extent of its impact on performance remains a subject of debate (cf. Ericsson, 2008; Anderson, 2010; Macnamara et al., 2016). By definition, deliberate practice requires that a coach or trainer sets specific individual training goals and provides feedback to practitioners. This may constitute a blocking obstacle when no coach is available, e.g., during travel or exercise at home. Motivated by prior research results and applications of the SDA-M method the present study investigated whether automatized SDA-M-based assessments could serve as an approximate technical substitute for the role that human coaches fulfill in deliberate practice. This included identifying potential issues and assessing a practitioner's overall competency with respect to specific movement sequences to derive feasible training goals.

The present study focused on choosing correct movements, not on improving individual actions' execution quality. Arguably, assisting deliberate practice on the level of basic action selection rather than the level of atomic action features seems especially helpful for intermediate and advanced practitioners, since Ericsson (2008, p. 991) noted that after sufficient practice *"the aspiring expert performers become able to monitor their performance so they can start taking over the evaluative activity of the teacher and coach. They acquire and refine mechanisms that permit increased control, which allow them to monitor performance in representative situations to identify errors as well as improvable aspects."* While this kind of self-monitoring might work well for recurring basic actions, like well-known karate techniques, it cannot prevent mistakes in insufficiently practiced action sequences.

Albeit preliminary due to a limited sample size, the empiric results are highly promising: SDA-M-based algorithms reached accuracy values that were highly significant above chance level and correctly predicted up to 77% of all actual errors in action selection during the tested karate

[3]Lander's λ and ARI were both calculated from SDA-M clusterings with a significance level of $\alpha = 1$ %.

TABLE 2 | Detailed results of SDA-M-based error prediction in the *Kanku-dai* sequence.

Algorithm	Accuracy	Sensitivity	Specificity	PPV	NPV	Balanced accuracy
AMPA	0.71***	0.55	0.75	0.29	0.90	0.65
$CASPA_d$	0.69***	0.74	0.68	0.31	0.93	0.71
$CASPA_i$	0.69***	0.77	0.67	0.31	0.94	0.72

***$p \ll 10^{-5}$.

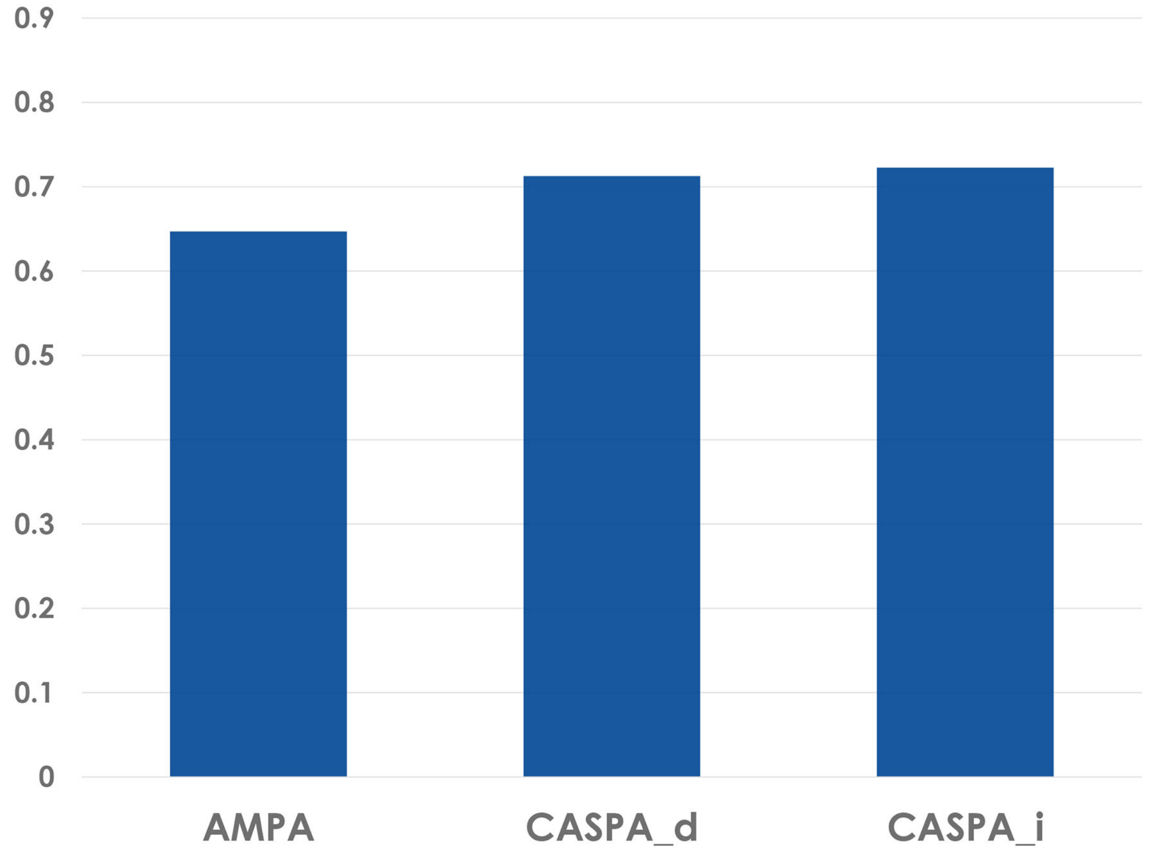

FIGURE 5 | Balanced accuracies of different SDA-M-based algorithms for error prediction in the *Kanku-dai* sequence.

movement sequence. In deliberate practice, this information could be used by coaches, practitioners, and CIT-based training assistance systems to focus on practicing corresponding subsequences including (at least) the preceding and subsequent techniques surrounding the practitioner's most error-prone action steps in order to strengthen associations between these actions.

Furthermore, SDA-M-based measures for assessing the overall suitability of participants' individual mental representation structures, especially the newly proposed $CASPA_m$ metric, correlated significantly and strongly with karate practitioners actual performances. After analyzing someone's mental representation structures related to different relevant movement sequences (e.g., a set of several *katas* that may need to be performed in their next belt examination), these metrics could be used to focus deliberate practice on poorly rated sequences, i.e., those that are not yet sufficiently established in practitioners' memory.

TABLE 3 | Correlations between formal expertise, actual performance, and SDA-M-based assessment metrics.

	Expertise	Performance	$CASPA_m$	Invariance λ
Performance	.84***			
$CASPA_m$	.80** b)	.88***		
Invariance λ	.66* a)	.79** c)	.80** a)	
ARI	.44	.65* b)	.49	.75** d)

***$p < 0.001$ **a) $p = 0.0016$ **b) $p = 0.0018$ **c) $p = 0.0022$ **d) $p = 0.0047$
*a) $p = 0.02$ *b) $p = 0.023$.

A notable limitation of the currently available algorithms for automatized SDA-M-based assessments and error predictions

FIGURE 6 | Based on the measurement of mental representation structures, it is possible to learn about the expertise stage of the user and to provide individualized feedback (e.g., in a bakery or golf putting scenario). Photos: CITEC. Reproduced with permission of Thomas Schack.

is that they require a predefined, limited set of correct action sequences in terms of basic actions. This makes them potentially applicable not only to martial arts forms and dance choreographies but also to opening sequences in chess or real-time strategy games (B. Strenge et al., unpublished) and other fixed sequences of basic actions that do not overlap in time. However, they cannot readily be applied to more dynamic, impulsive and spontaneous situations in sports and training that do not satisfy these requirements.

A mobile CIT assistance system, e.g., based on smart glasses, could use the information from SDA-M-based analyses to suggest training goals, provide feedback, and track practitioners' learning curves in terms of developing task-related memory structures over time. Such a system would enable intermediate practitioners to engage in deliberate practice of action sequences anywhere anytime instead of requiring personal contact with their coaches. Arguably, this would entail a broad range of ethically relevant aspects:

- Greater independence from organizational structures like sports clubs,
- Less time spent and environmental damage due to regular traveling,
- More flexible training schedules,
- Better opportunities for independent adjustment of repetitions in deliberate practice, and
- Prevention of potential embarrassment due to the observation of ones mistakes by other people.

With respect to the current situation concerning the ongoing COVID-19 pandemic and impending climate catastrophe, one might add that special circumstances make many of these aspects all the more relevant and pressing issues.

In another research direction, which could be interesting for anticipation in sport and medicine, researchers tried to support activities by seeing the world through assistive glasses. This project, called ADAMAAS (*Adaptive and Mobile Action Assistance in Daily Living Activities*), focused on the development of a mobile adaptive assistance system in the form of intelligent glasses, which provide unobtrusive, anticipatory, and intuitive support in everyday situations (Essig et al., 2016). The system is able to identify problems in ongoing action processes, react to mistakes, and provide context-related assistance via textual, pictorial, or three-dimensional virtual elements superimposed on a transparent display (see **Figure 6**). This project investigated the integration of mental representation analysis, eye tracking, physiological measures (e.g., heart rate), computer vision (i.e., object and action recognition), and augmented reality with modern diagnostics and corrective intervention techniques. The major perspectives that distinguish ADAMAAS from stationary diagnostic systems and conventional head-mounted displays include its ability to react to errors in real-time, provide individualized feedback for action support, and learn from the individual behavior of the user. Such intelligent AR glasses could be used to provide athlete- and sport-sensitive feedback, for remote observation or assistance (e.g., transferring the video to the trainer; or the trainer can use a salient pointer to help the athlete to focus on the relevant cue), as well as new forms of training, such as displaying distracting stimuli in the glasses in order to simulate different training conditions or environments (Schack, 2020).

Despite all the new possibilities opened up by the application of new technologies in sport science there are also many challenges that have to be considered: New technologies allow the recording and storage of detailed user-specific data.

Therefore, privacy issues and other ethical, legal and social implications (ELSI) are becoming more and more important and are seen as essential considerations with respect to technological developments.

A worthwhile CIT system would need to be developed with the aforementioned and other ethical aspects in mind to ensure that the potential benefits actually come into effect. Therefore, the technical development process should adhere to specific rules regarding the inclusion of ethical issues. This is especially important in contemporary agile development settings that are characterized by transient requirements definitions and short-term prioritization of features. Specialized system design methodologies like *"Value-Sensitive Design"* (Friedman et al., 2008) or the *"Ethical System Design Lifecycle"* (Spiekermann, 2015) define methods and processes for this purpose. In a similar vein, Strenge and Schack (2019) proposed an innovative approach to incorporate ethically relevant criteria during agile development processes through a flexibly applicable methodology called *Agile Worth-Oriented Systems Engineering* (AWOSE). First, a predefined model for the ethical evaluation of sociotechnical systems is used to assess ethical issues according to different dimensions. To ensure that ethical issues are not only identified but also systematically considered during system design, the second part of AWOSE integrates the findings with approaches from worth-centered development into a process model compatible with agile methodologies. Improved artifacts of worth-centered development called Worth Maps guide the prioritization of development tasks as well as choices among design alternatives with respect to ethical implications. Furthermore, the improved Worth Maps facilitate the identification of suitable criteria for system evaluations in association to ethical concerns and desired positive outcomes of system usage.

Future research could focus not only on replicating the current study's findings with more extensive and heterogeneous participant samples and other sample applications but also investigate the long-term applicability and usefulness of the automatized assessment approaches. A major research and development objective could be to build an assistance system and empirically test its impact on the quality and efficacy of deliberate practice compared to unassisted training and/or traditional coach interaction. Finally, a long-term study could verify which (if any) ethically relevant benefits actually arise from using such a system.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the ethics committee of Bielefeld University. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

BS was either solely accountable for, or involved in, all aspects of this work. DK contributed to the experimental design and execution, data preparation, and writing. TS provided the theoretical framework for analyzing mental representation structures with SDA-M, supervised the research strategy, and contributed to the writing. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Cluster of Excellence in Cognitive Interaction Technology CITEC (EXC 277) at Bielefeld University, which was funded by the German Research Foundation (DFG). It was partially related to project ADAMAAS, which was funded by the German Federal Ministry of Education and Research (BMBF). We acknowledge support for the Article Processing Charge by the DFG and the Open Access Publication Fund of Bielefeld University.

ACKNOWLEDGMENTS

We would like to thank Aya Ahmaro for assisting preliminary investigations and preparing materials for SDA-M applications in karate that pioneered the work reported in this article.

REFERENCES

Anderson, J. R. (2010). *Cognitive Psychology and Its Implications, 7th Edn*. New York, NY: Worth Publication.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., and Qin, Y. (2004). An integrated theory of the mind. *Psychol. Rev*. 111, 1036–1060. doi: 10.1037/0033-295X.111.4.1036

Anderson, J. R., and Lebiere, C. (1998). *The Atomic Components of Thought*. Mahwah, NJ: Lawrence Erlbaum Associates.

Brodersen, K. H., Ong, C. S., Stephan, K. E., and Buhmann, J. M. (2010). "The balanced accuracy and its posterior distribution," in *20th International Conference on Pattern Recognition (ICPR)* (Istanbul), 2010, 3121–3124. doi: 10.1109/ICPR.2010.764

Ericsson, K. A. (2007). Deliberate practice and the modifiability of body and mind: toward a science of the structure and acquisition of expert and elite performance. *Int. J. Sport Psychol*.38, 4–34.

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: a general overview. *Acad. Emerg. Med*. 15, 988–994. doi: 10.1111/j.1553-2712.2008.00227.x

Essig, K., Strenge, B., and Schack, T. (2016). "ADAMAAS: towards smart glasses for mobile and personalized action assistance," in *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '16), PETRA '16* (New York, NY: ACM), 46:1–46:4. doi: 10.1145/2910674.2910727

Fitts, P. M., and Posner, M. I. (1967). *Human Performance*. Belmont, CA: Brooks/Cole.

Frank, C., Land, W. M., Popp, C., and Schack, T. (2014). Mental representation and mental practice: Experimental investigation on the functional links between motor memory and motor imagery. *PLoS ONE* 9:e95175. doi: 10.1371/journal.pone.0095175

Frank, C., Land, W. M., and Schack, T. (2013). Mental representation and learning: the influence of practice on the development of mental representation structure in complex action. *Psychol. Sport Exerc*. 14, 353–361. doi: 10.1016/j.psychsport.2012.12.001

Frank, C., Land, W. M., and Schack, T. (2016). Perceptual-cognitive changes during motor learning: the influence of mental and physical practice on mental representation, gaze behavior, and performance of a complex action. *Front. Psychol*. 6:1981. doi: 10.3389/fpsyg.2015.01981

Friedman, B., Kahn, P. H., and Borning, A. (2008). "Value sensitive design and information systems," in *The Handbook of Information and Computer Ethics*, eds K. E. Himma and H. T. Tavani (Hoboken, NJ: John Wiley & Sons, Inc.), 69–101. doi: 10.1002/9780470281819.ch4

Hagan, J. E. Jr., Schack, T., and Koester, D. (2018). Passion play. Embracing new scientific perspectives for improved sport psychology consulting. *SOJ Psychol*. 4, 1–5. doi: 10.15226/2374-6874/5/1/00143

Heinen, T., and Schack, T. (2004). *Bewegungsgedächtnis und Bewegungsausführung-Optimierung von Rotationsbewegungen im Gerätturnen*. Lehren und Lernen im Turnen. Veröffentlichungsband zur Jahrestagung, 85–95.

Heinen, T., and Schwaiger, J. (2002). "Optimierung des Trainingsprozesses im Kunstturnen durch kognitive Verfahren," in *Expertise im Sport: Lehren, Lernen, Leisten*, eds B. Strauß, M. Tietjens, N. Hagemann and A. Stachelhaus (Cologne: bps) 67–68.

Heinen, T., Schwaiger, J., and Schack, T. (2002). "Optimising gymnastics training with cognitive methods," in *Proceedings of 7th annual Congress of the European College of Sport Science* (Athens), 608.

Hodges, N. J., Huys, R., and Starkes, J. L. (2007). "Methodological review and evaluation of research in expert performance in sport," in *Handbook of Sport Psychology*, eds G. Tenenbaum and R. C. Eklund (Hoboken, NJ: John Wiley & Sons), 161–183. doi: 10.1002/9781118270011.ch7

Hossner, E.-J., Schiebl, F., and Göhner, U. (2015). A functional approach to movement analysis and error identification in sports and physical education. *Front. Psychol*. 6:1339. doi: 10.3389/fpsyg.2015.01339

Hubert, L., and Arabie, P. (1985). Comparing partitions. *J. Classif*. 2, 193–218. doi: 10.1007/BF01908075

Jeraj, D., Musculus, L., and Lobinger, B. (2017). Body image and mental representation in table tennis players who do versus do not use a prosthesis. *Probl. Psychol. 21 Cent*. 11, 22–30.

Kim, T., Frank, C., and Schack, T. (2017). A systematic investigation of the effect of action observation training and motor imagery training on the development of mental representation structure and skill performance. *Front. Hum. Neurosci*. 11:499. doi: 10.3389/fnhum.2017.00499

Land, W., Frank, C., and Schack, T. (2014). The influence of attentional focus on the development of skill representation in a complex action. *Psychol. Sport Exerc*. 15, 30–38. doi: 10.1016/j.psychsport.2013.09.006

Lander, H.-J., and Lange, K. (1992). Eine differentialpsychologische Analyse begrifflich-strukturierten Wissens. *Zeitschrift für Psychologie mit Zeitschrift für angewandte Psychologie* 200, 181–197.

Macnamara, B. N., Moreau, D., and Hambrick, D. Z. (2016). The relationship between deliberate practice and performance in sports: a meta-analysis. *Perspect. Psychol. Sci*. 11, 333–350. doi: 10.1177/1745691616635591

Meier, C., Frank, C., Gröben, B., and Schack, T. (2020). Verbal instructions and motor learning: how analogy and explicit instructions influence the development of mental representations and tennis serve performance. *Front. Psychol*. 11:2. doi: 10.3389/fpsyg.2020.00002

Ritter, H. (2010). Cognitive interaction technology. *Künstl. Intell*. 24, 319–322. doi: 10.1007/s13218-010-0063-x

Ritter, H., and Sagerer, G. (2009). Excellence cluster "cognitive interaction technology"-cognition as a basis for natural interaction with technical systems. *Inform. Technol*. 51, 112–118. doi: 10.1524/itit.2009.0532

Schack, T. (2004). The cognitive architecture of complex movement. *Int. J. Sport Exerc. Psychol*. 2, 403–438. doi: 10.1080/1612197X.2004.9671753

Schack, T. (2012). "Measuring mental representations," in *Measurement in Sport and Exercise Psychology*, eds G. Tenenbaum, R. C. Eklund, and A. Kamata (Champaign, IL: Human Kinetics), 203–214. doi: 10.5040/9781492596332.ch-019

Schack, T. (2020). "Mental representation in action," in *Handbook of Sport Psychology*, eds G. Tenenbaum and R. C. Eklund (Hoboken, NJ: John Wiley & Sons, Ltd.), 513–534. doi: 10.1002/9781119568124.ch24

Schack, T., Bertollo, M., Koester, D., Maycock, J., and Essig, K. (2014a). "Technological advancements in sport psychology," in *Companion to Sport and Exercise Psychology: Global Perspectives and Fundamental Concepts*, eds A. G. Papaioannou and D. Hackfort (Routledge), 953–965.

Schack, T., Essig, K., Frank, C., and Koester, D. (2014b). Mental representation and motor imagery training. *Front. Hum. Neurosci*. 8:328. doi: 10.3389/fnhum.2014.00328

Schack, T., and Hackfort, D. (2007). An action theory approach to applied sport psychology. *Handb. Sport Psychol*. 3, 332–351. doi: 10.1002/9781118270011.ch15

Schack, T., Hagan, J. E. Jr., and Essig, K. (2020). "New technologies in sport psychology practice," in *The Routledge International Encyclopedia of Sport and Exercise Psychology. Volume 2: Applied and Practical Measures*, eds D. Hackfort and R. J. Schinke (New York, NY: Routledge), 14.

Schack, T., and Mechsner, F. (2006). Representation of motor skills in human long-term memory. *Neurosci. Lett*. 391, 77–81. doi: 10.1016/j.neulet.2005.10.009

Schack, T., and Ritter, H. (2013). Representation and learning in motor action-bridges between experimental research and cognitive robotics. *N. Ideas Psychol*. 31, 258–269. doi: 10.1016/j.newideapsych.2013.04.003

Spiekermann, S. (2015). *Ethical IT Innovation*. Boca Raton, FL: Auerbach Publications. doi: 10.1201/b19060

Strenge, B., and Schack, T. (2019). Awose – A process model for incorporating ethical analyses in agile systems engineering. *Sci. Eng. Ethics* 26, 851–870. doi: 10.1007/s11948-019-00133-z

Strenge, B., Vogel, L., and Schack, T. (2019). Computational assessment of long-term memory structures from SDA-M related to action sequences. *PLoS ONE* 14:e0212414. doi: 10.1371/journal.pone.0212414

Tenenbaum, G., Hatfield, B. D., Eklund, R. C., Land, W. M., Calmeiro, L., Razon, S., et al. (2009). "A conceptual framework for studying emotions-cognitions-performance linkage under conditions that vary in perceived pressure," in *Mind and Motion: The Bidirectional Link Between Thought and Action*, Vol. 174, eds M. Raab, J. G. Johnson, and H. R. Heekeren (Elsevier), 159–178. doi: 10.1016/S0079-6123(09)01314-4

Wachsmuth, I. (2008). "Cognitive interaction technology: humans, robots, and max," in *International Conference on Informatics Education and Research for Knowledge-Circulating Society (ICKS 2008)* (Kyoto), 4–5. doi: 10.1109/ICKS.2008.34

Wachsmuth, S., Schulz, S., Lier, F., Siepmann, F., and Lütkebohle, I. (2012). "The robot head "Flobi": a research platform for cognitive interaction technology," in *German Conference on Artificial Intelligence*, ed S. Wöfl (Saarbrücken: Deutsches Forschungszentrum für Künstliche Intelligenz), 3–7.

Wrede, S., Leichsenring, C., Holthaus, P., Hermann, T., Wachsmuth, S., and Team, T. C. (2017). The cognitive service robotics apartment: a versatile environment for human-machine interaction research. *Künstl. Intell*. 31, 299–304. doi: 10.1007/s13218-017-0492-x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINAL RESEARCH
published: 15 January 2021
doi: 10.3389/fpsyg.2020.570534

Exploring the Implicit Link Between Red and Aggressiveness as Well as Blue and Agreeableness

Lu Geng[1], Xiaobin Hong[1,2]* and Yulan Zhou[1]

[1] Department of Psychology, Wuhan Sports University, Wuhan, China, [2] Hubei Key Laboratory of Exercise Training and Monitoring, College of Health Science, Wuhan Sports University, Wuhan, China

Previous studies have found a link between red and aggressive behavior. For example, athletes who wear red uniforms in sports are considered to have a competitive advantage. So far, most previous studies have adopted self-report methods, which have low face validity and were easily influenced by the social expectations. Therefore, the study used two implicit methods to further explore the association between red and aggressiveness. A modified Stroop task was used in Experiment 1 to probe college students' differences between "congruent" tasks (i.e., red–aggressiveness and blue–agreeableness) and "incongruent" tasks (i.e., red–agreeableness and blue–aggressiveness). Result showed that participants responded more quickly to the congruent tasks than the incongruent tasks. Then, in order to adapt to the competitive context, Experiment 2 used an implicit association test with photos of athletes as the stimulus to college students and athletes to evaluate "congruent" tasks (i.e., red uniform photo-aggressiveness and blue uniform photo-agreeableness) as well as "incongruent" tasks (i.e., red uniform photo-agreeableness and blue uniform photo-aggressiveness), respectively. According to the results, both college students and athletes respond faster to congruent tasks than to incongruent tasks. Besides, athletes' reactions to the red–aggressiveness association are faster than college students, which may relate to the athletes' professional experience. The athletes may be more aggressive and impulsive. Overall, the study has attempted to examine the association between red and aggressiveness through implicit methods, but in the future, researches are need to find a deep association from brain mechanism aspect.

Keywords: color, stroop task, implicit association test, red, aggressiveness-agreeableness, blue

Edited by:
Roy David Samuel,
Interdisciplinary Center Herzliya, Israel

Reviewed by:
Nataniel Boiangin,
Barry University, United States
Bjoern Krenn,
University of Vienna, Austria

****Correspondence:***
Xiaobin Hong
471780693@qq.com

Specialty section:
This article was submitted to Movement Science and Sport Psychology, a section of the journal Frontiers in Psychology

Received: *08 June 2020*
Accepted: *16 December 2020*
Published: *15 January 2021*

Citation:
Geng L, Hong X and Zhou Y (2021) Exploring the Implicit Link Between Red and Aggressiveness as Well as Blue and Agreeableness. Front. Psychol. 11:570534. doi: 10.3389/fpsyg.2020.570534

INTRODUCTION

Color perception is a common sensation in daily lives, not only as an esthetic awareness but also in terms of its wider impact on human psychology and behaviors. Theories about color and psychological functioning have been proposed since the early 19th century (Goethe, 1975, p.115), but only in the last decade, an increasing body of empirical researches began to document the influence of color on human consciousness and conduct. Aside from a few studies exploring the effect of black (Frank and Gilovich, 1988; Şefik Tiryaki, 2005; Caldwell and Burger, 2011), the majority of prior researches have examined the effect of red on psychological functioning, as cognitive performance (Mehta and Zhu, 2009), sexual attractiveness (Elliot and Niesta, 2008), financial prediction (Jiang et al., 2014), competitive sports (Hill and Barton, 2005), and food consumption (Bruno et al., 2013). Researchers have been interested in discerning advantage for competitors wearing red. In their analysis of the Olympic Games in 2004, Hill and Barton (2005) reported an advantage for athletes wearing red (vs. blue) in combat sports such as boxing, Taekwondo, and wrestling. They stated that the advantage conferred by red uniforms may be because of an evolutionary, engrained, and social learning association of red with dominance and aggression.

Hill and Barton (2005) provoked additional academic interest in testing the advantages of wearing red in competition, but inconsistent results have been obtained (Attrill et al., 2008; Sorokowski and Szmajke, 2011; Piatti et al., 2012; Curby, 2016), which indicated that various factors may affect athletes' performance in competition, and red is not the only one. From an implicit perspective, most of the above results are archival research. Thus, the association between red and aggressiveness is possible. Other studies have focused on the benefits originated from the association between red and aggression. Little and Hill (2007) found that, compared to blue shapes, red shapes were regarded to be more aggressive and dominant and more likely to win in physical competitions.

Researchers also examined the influence of red on social perceptions of dominance and aggressiveness in non-competitive contexts. In one study, men in red were rated as more aggressive and more dominant than those in blue or gray (Wiedemann et al., 2015). Briki and Hue (2016) studied how red, blue, and green were judged in relation to conceptualizations of dominance, arousal, and pleasure and found that red is strongly associated with dominance. In competitive contexts, Feltman and Elliot (2011) confirmed that wearing red can enhance perceptions of the dominance and threat, both of the opponents' and of one's own. The factors that affect the performance include not only the players and opponents but also the referee (Plessner and Haar, 2006). Some studies explored the advantages of red uniform from the perspective of referee and found that red is more dominant and aggressive (Hagemann et al., 2008; Krenn, 2014, 2015).

Color-in-Context Theory holds that there are two sources in color meanings: learning and biology (Elliot and Maier, 2012). For example, in competitive context, red-headed birds were found to be more likely to win than black-headed or yellow-headed birds (Pryke and Griffith, 2006). In addition, red in a male mandrill or baboon's face and genitalia is a symbol of status: the brighter the red is, the stronger the male's attack power will be (Setchell and Wickings, 2005; Bergman et al., 2009). Human beings, as their closest primate relatives, hold the same color-meaning pairings as those animals have. Shi et al. (2015) indicated that viewing red can impair participants' performance during a challenging cognitive task. This suggests that red is always associated with warning in the study context due to repetition of wrong pairings (e.g., teachers often mark errors with red pen). Social learning of red can also be extended to the national level. More recently, a study found that red was the most frequently used color in national flags across the world. This may because red was often attached with an aggressive connotation, and red can better reflect the competitiveness of the country. Compared to blue, red is rarely used in collaborative organizations (Zhang et al., 2018).

Associated with clear sky and clean water, blue is often used as the opposite color of red in competitive context. Blue represents peace and quiet in many cultures. It is noteworthy that blue also has different meanings in different cultures. Studies suggested that participants from the United States, Germany, and Turkey associate blue with positive meanings, thereby triggering positive emotions (Garth and Collado, 1921; Gesche, 1927; Choungourian, 1968). However, participants from Japan, the Philippines, Rome, Kuwait, and American India associate blue with negative connotations, which trigger negative emotions (Valdez and Mehrabian, 1994; Camgöz et al., 2002; Wogalter et al., 2002; Hurlbert and Ling, 2007; Frühholz et al., 2009; Gerend and Tricia, 2009). Associated with clear sky and clean water, blue is often used as the opposite color of red in the competitive context. Blue represents peace and quiet in many cultures. It is noteworthy that blue also has different meanings in different cultures. Studies suggested that participants from the United States, Germany, and Turkey associate blue with positive meanings, thereby triggering positive emotions (Garth and Collado, 1921; Gesche, 1927; Choungourian, 1968). However, participants from Japan, the Philippines, Rome, Kuwait, and American India associate blue with negative connotations, which trigger negative emotions (Valdez and Mehrabian, 1994; Camgöz et al., 2002; Wogalter et al., 2002; Hurlbert and Ling, 2007; Frühholz et al., 2009; Gerend and Tricia, 2009).

Above researches, it is proposed that there may be an implicit association between color and meanings. However, most studies adopted self-report methods, which have low ostensible validity and can easily be influenced by social expectation effect. Thus, the study used the implicit methods to explore the issue. Moreover, Jiang et al. (2014) found a "red up and green down" effect among Chinese mainland but a "green up and red down" among Hong Kong, which suggests that social culture might affect color and its associations. Does an association between red and aggressiveness exists among Chinese people? Is it consistent with the research carried out in the West? Experiment 1 assumed that, according to response competition logic (Klinger et al., 2000), if red is associated with aggressiveness, then the simultaneous presentation of red and aggressiveness-related words will be classified more quickly. Same trend applied for agreeableness-related words shown in blue.

Only a few studies conducted experiments from an implicit perspective (Moller et al., 2009; Soriano and Valenzuela, 2009; Fetterman and Meier, 2012; Pravossoudovitch et al., 2014). Mentzel et al. (2017) used a modified Stroop task to testify the red-dominance association, with lexical stimuli only. However, it is not specific in the competition context. Thus, Experiment 2 used the photo of athletes in red as the stimulus to examine the implicit association between red and aggressiveness. Based on the Five-Factor Model (FFM) of personality (Pervin and John, 1997); Trninić et al. (2008) found that the athletes have a significant correlation between aggressiveness and extraversion, and agreeableness and emotional stability. There are significant negative relations between emotional stability and aggressiveness.

From the aspect of their development, adolescence is a turbulent life period; the athletes in the present study are young men in the adulthood late stage, and they are vulnerable to experience strong emotions. Experiment 2 postulated that the athletes will attain a shorter response time than college students in the association between red and aggressiveness.

In summary, the present study applied two implicit methods—a modified Stroop task and an implicit association test (IAT)—to explore red advantage and blue meaning among Chinese people.

EXPERIMENT 1

In this experiment, aggressiveness-related words and agreeableness-related words are presented in red or blue; the aim of the experiment was to examine whether aggressiveness-related words presented in red would be categorized more quickly.

Methods

Participants

Among 80 college students (43 females; mean age = 18.81 ± 0.86), all right-handed and not red–green colorblind, none took part in a similar psychological experiment before. To enhance the response rate, pre-paid envelopes were provided along with a small gift.

Stimuli and Pilot Test

In the modified Stroop task, ten Chinese words were used as lexical stimuli, five denoting aggressiveness (*bullying*, *offense*, *murder*, *aggression*, and *war*) and five denoting agreeableness (*close*, *caring*, *gentle*, *friendly*, and *harmonious*). All ten Chinese words consisted of two characters with the same word length. The words were rated in a pilot test by 97 college students according to the degree of aggressiveness, agreeableness, and familiarity, on a scale of 1 (*not at all*) to 5 (*extremely*). The aggressiveness-related words were rated as more aggressive ($M = 4.58$, SD = 0.71) than the agreeableness-related words ($M = 1.24$, SD = 0.49), $t = -82.89$, $p < 0.001$; the agreeableness-related words were rated as more agreeable ($M = 4.48$, SD = 0.77) than the aggressiveness-related words ($M = 1.18$, SD = 0.46), $t = 79.185$, $p < 0.001$. Moreover, all the words were rated as equal in terms of familiarity, $t = 0.636$, $p > 0.05$. The RGB criteria of red (255, 0, and 0) and blue (0, 0, and 255) were applied to the words. The HSL (hue, saturation, and lightness) criteria are as follows: red (0, 240, and 120) and blue (160, 240, and 120). The words constituted a 2 (Valence: aggressiveness vs. agreeableness) × 2 (Color: red vs. blue) lexical stimulus.

Design and Procedure

Special software E-prime 2.0 was used with a 12.1-inch screen Lenovo Think pad X200 for the stimulus presentation and data logging. After the experiment, the subjects were asked whether they could guess the purpose of the experiment, while all the subjects said no.

The experiment had a 2 (Valence: aggressiveness vs. agreeableness) × 2 (Color: red vs. blue) repeated-measures design. Each word was separately presented in red and blue on a black computer screen in random order. Participants were asked to press a key labeled to judge whether the word was aggressiveness-related or agreeableness-related. There were 20 trials in each practice block and 60 trials in the experiment block.

Before each word was displayed, a fixation cross appeared for 500 ms in the center of the computer screen. The word rendering time was 3,000 ms, and participants who did not make a response within the 3,000 ms went directly to the next trial. Participants whose response times were longer than 3,000 ms or shorter than 300 ms were eliminated from the analyses.

Results and Discussion

A 2 (Color: red vs. blue) × 2 (Valence: aggressiveness vs. agreeableness) repeated-measures analysis of variance (ANOVA) regarding reactions times revealed a significant main effect of Color, but not Valence, $F(1,75) = 27.63$, $p = 0.007$ ($p < 0.01$), $\eta^2 = 0.03$, with participants found to be faster in categorizing red words ($M = 599.59$ ms, SD = 124.13) than blue words ($M = 602.26$ ms, SD = 148.25). More importantly, a significant Valence × Color interaction, $F(1,75) = 42.27$, $p = 0.041$ ($p < 0.05$), $\eta^2 = 0.17$, indicated that participants were faster in categorizing aggressiveness-related words presented in red ($M = 595.16$ ms, SD = 191.60) than in blue ($M = 611.19$ ms, SD = 202.19), $t(76) = -1.46$, $p = 0.038$ ($p < 0.05$), and were faster in categorizing agreeableness-related words presented in blue ($M = 593.33$ ms, SD = 116.82) than in red ($M = 604.01$ ms, SD = 191.54), $t(76) = 3.92$, $p = 0.043$ ($p < 0.05$) (see **Figure 1**).

Thus, the results indicated that red is positively associated with aggressiveness, because the aggressiveness-related words presented in red were categorized more quickly than in blue. In addition, blue was found to be positively associated with agreeableness, as agreeableness-related words presented in blue were categorized more quickly than in red.

EXPERIMENT 2

In the experiment, the IAT (Greenwald et al., 1998) was used to test whether the college students and the athletes have differences in response time between congruent tasks and incongruent tasks.

The IAT included two tasks: congruent task and incongruent task. When the relationship between a concept and an attribute was consistent with participants' implicit attitudes, it was considered congruent task. Participants typically follow their original attitude when making such judgments, which results in shorter response times. Conversely, when the relationship between a concept and an attribute was inconsistent with participants' implicit attitudes, it was deemed incongruent task. Participants tend to experience cognitive conflict when making this type of judgment, and so their reaction time would likely be longer. In both cases, the speed of reaction was taken to be a measure of the associative strength between the concept and the attribute (Williams and Themanson, 2011).

The advantages of the IAT are as follows: First, it minimizes the issue of social expectations that often affects self-report research (Fazio et al., 1995; Wittenbrink et al., 1997). Second, the IAT can facilitate the analysis of implicit associations

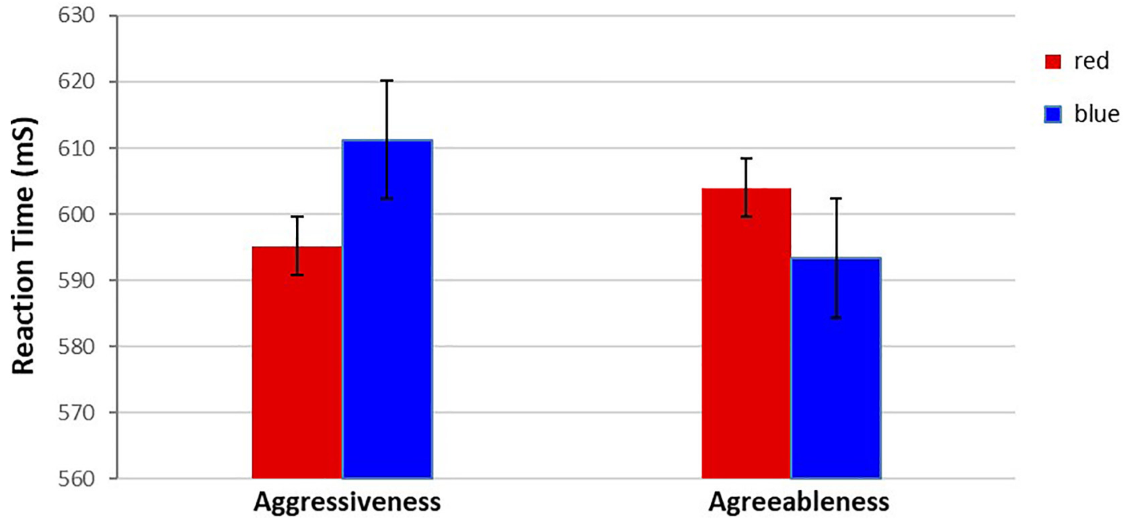

FIGURE 1 | I Mean response times for aggressiveness and agreeableness words presented in red and blue.

by effectively excluding explicit familiarity with the concepts (Dasgupta et al., 2003).

Method

Participants

Among 84 college students (mean age = 18.61 $\pm$ 0.98), all right-handed and not red-green colorblind, none took part in a similar psychological experiment before. To enhance the response rate, pre-paid envelopes were provided along with a small gift. 40 Taekwondo athletes (29 of whom were male) and 44 other college students (37 males) took part in the study. There are 16 national second-grade athletes and 24 national first-grade athletes. Their professional sports training age ranges from seven to ten years. Moreover, they have good physical fitness, strong professional skills, and rich competition experience.

Materials and Procedure

The materials consisted of two dimensions: a concept dimension (including picture material) and an attribute dimension (including text material). These concept materials consisted of 20 photos of Taekwondo athletes wearing red uniforms and blue uniforms. At first, ten athletes were selected to take photos in red uniforms. After that, the RGB criterion of red (255, 0, 0) was applied to the photos. Next, through Photoshop software, the RGB criterion of blue (0, 0, 255) with the same brightness and saturation was used to fill the red uniform area in the photos. The HSL (hue, saturation, and lightness) criteria are as follows: Red (0, 240, and 120) and blue (160, 240, and 120). Finally, the photos were carefully revised; the image format is JPG file with a size of 129 kb. The purpose is to prevent athletes from random movement and changing expression. Besides, the attribute material was made up of the ten words used in Experiment 1, and all words were repeated twice.

The experiment used special software E-prime 2.0 with a 12.1-inch screen Lenovo Think pad X200 for the stimulus presentation and data logging. A fixation cross appeared for 500 ms in the center of the computer screen, and the image stimulus-rendering time was 3,000 ms. Within the 3,000 ms, participants who did not make a response went directly to the next trial. A 500-ms blank screen was presented between each trial. The test was divided into seven blocks, and it consisted of 110 trials. Each participant completed the test on their own according to the instructions, and all were tested separately. Blocks 1, 2, 3, 5, and 6 served as practice rounds, and Block 4 and Block 7 were the test rounds; Block 4 featured the congruent tasks. When the picture stimulus appeared as an athlete in a red uniform and the word appeared as an aggressive-related word, the participants would press the "F" key to react; when the picture stimulus appeared as an athlete in a blue uniform and the word appeared as an agreeable-related word, the participants would press the "J" key to react. In the congruent tasks, participants were asked to classify an athlete in red uniform and an aggressive-related word as well as an athlete in blue uniform and an agreeable-related word. In Block 7 (i.e., an incongruent task), on the contrary, participants were asked to classify an athlete in red uniform and agreeable-related words as well as an athlete in blue uniform and aggressive-related words. After the experiment, a color blindness test was performed. The subjects were asked whether they could guess the purpose of the experiment, while all the subjects said no. Finally, in order to control a sequence effect, Blocks 2, 3, and 4 were, respectively, switched with Blocks 5, 6, and 7 for half of the participants.

Results and Discussion

Experiment 2's data were analyzed according to standard procedures (Greenwald et al., 1998). We recorded reaction times below 300–300 ms and those above 3,000–3,000 ms, and we also discarded incorrect responses. Following the preliminary processing of the data, there were 84 participants in the experiment and 79 valid data sets were obtained.

The experiment had a Groups (college students vs. Taekwondo athletes) × Combination (compatible vs. incompatible) ANOVA to explore whether college students and Taekwondo athletes would make different associations between Color and Valence. The results showed a significant main effect of Groups but not Combination, $F(1,77) = 11.62$, $p = 0.007$ ($p < 0.01$), $\eta^2 = 0.06$, which indicated that athletes were found to be faster in categorizing compatible task than college students. In addition, a significant interaction effect between Groups and Combination was significant, $F(1,77) = 4.03$, $p = 0.041$ ($p < 0.05$), $\eta^2 = 0.02$, which indicated that athletes were faster in categorizing congruent task ($M = 578.43$ ms, SD = 128.03) than incongruent task ($M = 601.24$ ms, SD = 162.19), $t(36) = -3.302$, $p = 0.001$ ($p < 0.01$), and college students were faster in categorizing congruent task ($M = 593.27$ ms, SD = 142.75) than incongruent task ($M = 612.15$ ms, SD = 149.74), $t(41) = 2.211$, $p = 0.027$ ($p < 0.05$). The results showed that the athletes responded faster to compatible task compared with the college students, while the response of incompatible task was not different between the athletes and the college students.

The results not only confirmed the previous hypothesis but also revealed that athletes responded faster on congruent tasks than college students. The reason why athletes reacted faster was as follows: in a competitive context, athletes wearing red uniforms were regarded as more aggressive, which suggested that the Taekwondo athletes had a stronger awareness of the association of red–aggressiveness. Overall, in the IAT, both the Taekwondo athletes and the college students responded significantly slower during the incongruent tasks compared to the congruent tasks.

DISCUSSION

Results from the above two experiments supported the hypothesis that there is an implicit link between red (or blue) and aggressiveness (or agreeableness), which is in line with the findings of previous studies using relatively straightforward methodologies (e.g., Krenn, 2014, 2015). Implicit measures were applied in this study to explore the advantage of red. This methodological improvement made the red–aggressiveness association clearer. Another important contribution of the present study is that photos of athletes are used in the IAT experiment, which is closer to a real competitive context.

Experiment 1 used a modified Stroop task to present textual stimuli and found an implicit association between red and aggressiveness. This is in keeping with the views of Briki and Hue (2016), who found a strong association between red and dominance. There may be the following explanations. Firstly, it could be related to biological factors. In the animal world, red is often related to danger. Likewise, human beings possess a biologically engrained predisposition to associate red with aggressiveness. Secondly, although having a positive meaning for Chinese people in general, red is usually used to signify negative connotations in an educational context. For example, teachers use red pen to mark incorrect answer. Since the participants are all college students, red may arouse their negative association with aggressiveness. In addition, participants have no idea about the purpose of the experiment; thus, they showed little awareness of these red effects. This suggests the red–aggressiveness association appears to take place outside of participants' conscious awareness (Elliot and Maier, 2007). Little and Hill (2007) suggested that the dominance of red can be influenced by hue information in the stimuli. Therefore, the present study left the question of which type of red is most strongly connected to aggressiveness. To find the exact connection between red and aggressiveness, it would be helpful to conduct various experiments in the future. For instance, hue is held constant while lightness and chroma are systematically varied in different ways. In line with previous studies (Pravossoudovitch et al., 2014; Mentzel et al., 2017), no gender difference was found in this study. From an evolutionary point of view, humans are born to relate red with the danger. The implicit cognition of the red–aggressiveness association was strong for both male and female. Therefore, it can be inferred that aggressiveness is a characteristic not only of male but also of female, and there is no significant difference to some extent. Future research can continue to explore whether there is gender difference in implicit cognition between male and female.

Meanwhile, similar to the results from the United States, Germany, and Turkey's participants (Garth and Collado, 1921; Gesche, 1927; Choungourian, 1968), the present study also confirms that participants are faster in response to the blue–agreeableness association, which suggests that blue is associated with the notion of positive and agreeable in Chinese culture. More recently, Mentzel et al. (2017) also found an implicit link between red and dominance through a Stroop task, but they failed to find an implicit connection between blue and rest. A reason for this difference may be that the concepts of agreeableness and aggressiveness used in the present study have opposite meanings, while the contrasting meanings between dominance and rest are not obvious in Mentzel's study. What is more, Curby (2016) found that in men's freestyle wrestling, wearing a blue uniform was significantly associated with winning. This led the doubt whether there is blue uniform effect. Further studies are needed to verify it.

In Experiment 2, Taekwondo athletes felt a stronger implicit association between red and aggressiveness than college students. Taekwondo athletes often participated in sport competitions, whereas college students are seldom exposed to competitive sports. Thus, they may react to competitive context differently. Cooper (1969) showed that athletes have a higher achievement motivation as well as a higher self-confidence and aggressiveness than others. Kirkcaldy (1982) also found that elite athletes are more extroverted and aggressive. From this point of view, athletes may have a more aggressive and impulsive personality than college students. Sometimes, red is commonly linked to love and romance in the affiliation context. Future research may explore red's meaning among athletes in other contexts.

Moreover, Dreiskaemper et al. (2013) found that athletes in red jersey would have significantly higher heart rates and significantly higher pre-contest values on the strength test, but it did not influence the results. This suggested that the influence of red on psychological functioning is as pervasive as it is subtle and provocative. Continuous researches are needed to clarify the impact of red on athletes and opponents. There is no difference between athletes and college students in the incongruent task. This indicates that little close association exists between agreeableness-related words and athletes wearing red. This provided indirect support for the red–aggressiveness association. Participants were more likely to think that athletes wearing red uniforms were more aggressive. As mentioned above, social cultural elements might affect red and its associations. The present study confirmed the association between red and aggressiveness among Chinese people, which is in accord with the research carried out in the West.

The study also offers a wide range of implications for daily life. For example, red, be it associated with dangers and mistakes, could activate an avoidance motivation. It has been shown to make people more vigilant and risk-averse; hence, red is not the main color used in hospital wards. What then is the most attractive color for advertisement? There are two cases. If the advertisements aim at warning, red is preferable. On the contrary, if concerning environmental protection, blue is a better choice. Furthermore, associated with peace and tranquility, blue is likely to activate an approach motivation and thus is usually used to encourage people to create new things.

The meaning of color is used not only in daily life but also in the competitive context. If there is an association between red and aggressiveness, the influence of the association on the referee's penalty would be automatic and unconscious. Since the referee's penalty will be affected by the red uniform, it will have an adverse effect on the fairness of the competition. When the referees' self-control strength was low, the association of red–aggressiveness would have more influence on the referee's penalty. Conversely, when referees' self-control strength was high, the influence of the red–aggressiveness association on the referee's penalty would be weakened. Bertrams et al. (2015) found that red had a negative effect on performance of participants with ego depletion, but hardly or not at all among control groups. This reminds us that referees should take some measures (e.g., glucose intake, enough rest, and taking exercise) to enhance self-control in order to prevent themselves from a red–aggressiveness association. In doing so, it may contribute to achieving fairness in competition and equal opportunities for each athlete to win, irrespective of uniform color.

In conclusion, although this study examined the association between red and aggressiveness, the effect size of the study is comparatively weak. Future researches are needed to explore the association between red and aggressiveness. In Experiment 2, the athletes were not in the same level; this may also have an impact on the results. Additionally, Krenn (2014) found that red uniform may affect the referee's penalty. However, only college students and athletes are selected as participants in the present study. Future study may consider recruiting referees as participants. In terms of experimental material color, although blue and red were selected, which are commonly used as contrast colors in competitive sports, a more professional spectrophotometer could be considered to adjust brightness and chroma in future study. At present, the mechanism of the red–aggressiveness association is not clear. Therefore, future research should explore the brain mechanism of the association between red and aggressiveness.

CONCLUSION

This study showed an implicit association between red and aggressiveness, as well as an association between blue and agreeableness, yielding an evidence for a link between colors and meanings. What is more, athletes were found to have shorter response times than college students regarding the association of red–aggressiveness.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by WSU Medical Ethics Committee. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

XH was responsible for the design of the study and the supervision, planning, and feedback on the written article. LG was responsible for data collection and analysis and for writing the first draft of the article. YZ assisted in the experiment. All authors have approved the manuscript and given consent for its submission and subsequent publication.

FUNDING

This research was supported by the Chinese National Natural Science Foundation (NSFC81971661), the Hubei Provincial Department of Education, China (18D077), the Research Fund Project (D20194102), and the East Lake Scholar Program of Wuhan Sports University, China (Period: 2018.05–2023.05).

ACKNOWLEDGMENTS

We would like to thank all the participants in these experiments and the members of our research group for their help.

REFERENCES

Attrill, M. J., Gresty, K. A., Hill, R. A., and Barton, R. A. (2008). Red shirt colour is associated with long-term team success in English football. *J. Sports Sci.* 26, 577–582. doi: 10.1080/02640410701736244

Bergman, T. J., Ho, L., and Beehner, J. C. (2009). Chest color and social status in male Geladas (*Theropithecus gelada*). *Int. J. Primatol.* 30, 791–806. doi: 10.1007/s10764-009-9374-x

Bertrams, A., Baumeister, R. F., Englert, C., and Furley, P. (2015). Ego depletion in color priming research: self-control strength moderates the detrimental effect of red on cognitive test performance. *Pers. Soc. Psychol. Bull.* 41, 311–322. doi: 10.1177/0146167214564968

Briki, W., and Hue, O. (2016). How red, blue, and green are affectively judged. *Appl. Cogn. Psychol.* 30, 301–304. doi: 10.1002/acp.3206

Bruno, N., Martani, M., Corsini, C., and Oleari, C. (2013). The effect of the color red on consuming food does not depend on achromatic (Michelson) contrast and extends to rubbing cream on the skin. *Appetite* 71, 307–313. doi: 10.1016/j.appet.2013.08.012

Caldwell, D. F., and Burger, J. M. (2011). On thin ice: does uniform color really affect aggression in professional hockey? *Soc. Psychol. Pers. Sci.* 2, 306–310. doi: 10.1177/1948550610389824

Camgöz, N., Yener, C., and Güvenc, D. (2002). Effects of hue, saturation, and brightness on preference. *Color Res. Appl.* 27, 199–207. doi: 10.1002/col.10051

Choungourian, A. (1968). Color preferences and cultural variation. *Percept. Mot. Skills* 26, 1203–1206. doi: 10.2466/pms.1968.26.3c.1203

Cooper, L. (1969). Athletics, activity and personality: a review of the literature. *Res. Q.* 40, 17–22. doi: 10.1080/10671188.1969.10616637

Curby, D. G. (2016). Effect of uniform color on outcome of match at Senior World Wrestling Championships 2015. *Int. J. Wrestling Sci.* 6, 62–64. doi: 10.1080/21615667.2016.1210266

Dasgupta, N., Greenwald, A., and Banaji, M. (2003). The first ontological challenge to the IAT: attitude or mere familiarity? *Psychol. Inq.* 14, 238–243. doi: 10.2307/1449682

Dreiskaemper, D., Strauss, B., Hagemann, N., and Büsch, D. (2013). Influence of red jersey color on physical parameters in combat sports. *J. Sport Exerc. Psychol.* 35, 44–49. doi: 10.1123/jsep.35.1.44

Elliot, A. J., and Maier, M. A. (2007). Color and psychological functioning. *Curr. Dir. Psychol. Sci.* 16, 250–254. doi: 10.1111/j.1467-8721.2007.00514.x

Elliot, A. J., and Maier, M. A. (2012). Color-in-context theory. *Adv. Exp. Soc. Psychol.* 45, 61–125. doi: 10.1016/B978-0-12-394286-9.00002-0

Elliot, A. J., and Niesta, D. (2008). Romantic red: red enhances men's attraction to women. *J. Pers. Soc. Psychol.* 95, 1150–1164. doi: 10.1037/0022-3514.95.5.1150

Fazio, R. H., Jackson, J. R., Dunton, B. C., and Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial stereotypes: a bona fide pipeline? *J. Pers. Soc. Psychol.* 69, 1013–1027. doi: 10.1037/0022-3514.69.6.1013

Feltman, R., and Elliot, A. J. (2011). The influence of red on perceptions of relative dominance and threat in a competitive context. *J. Sport Exerc. Psychol.* 33, 308–314. doi: 10.1123/jsep.33.2.308

Fetterman, A. K., and Meier, B. P. (2012). Anger as "seeing red": evidence for a perceptual association. *Cogn. Emot.* 26, 1445–1458. doi: 10.1080/02699931.2012.673477

Frank, M. G., and Gilovich, T. (1988). The dark side of self- and social-perception: black uniforms and aggression in professional sports. *J. Pers. Soc. Psychol.* 54, 74–85. doi: 10.1037/0022-3514.54.1.74

Frühholz, S., Fehr, T., and Herrmann, M. (2009). Early and late temporo-spatial effects of contextual interference during perception of facial affect. *Int. J. Psychophysiol.* 74, 1–13. doi: 10.1016/j.ijpsycho.2009.05.010

Garth, T. R., and Collado, I. R. (1921). The color preferences of Filipino children. *J. Comp. Psychol.* 9, 397–404. doi: 10.1037/h0075071

Gerend, M. A., and Tricia, S. (2009). Message framing and color priming: how subtle threat cues affect persuasion. *J. Exp. Soc. Psychol.* 45, 999–1002. doi: 10.1016/j.jesp.2009.04.002

Gesche, I. (1927). The color preferences of one thousand one hundred and fifty-two Mexican children. *J. Comp. Psychol.* 7, 297–311. doi: 10.1037/h0072520

Goethe, J. W. (1975). *Theory of Colors, Trans. C. L. Eastlake*. New York, NY: Gordon Press.

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. (1998). Measuring individual difference in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 74, 1464–1480. doi: 10.1037/0022-3514.74.6.1464

Hagemann, N., Strauss, B., and Leissing, J. (2008). When the referee sees red.... *Psychol. Sci.* 19, 769–771. doi: 10.1111/j.1467-9280.2008.02155.x

Hill, R. A., and Barton, R. A. (2005). Red enhances human performance in contests. *Nature* 435:293. doi: 10.1038/435293a

Hurlbert, A. C., and Ling, Y. (2007). Biological components of sex differences in color preference. *Curr. Biol.* 17, 623–625. doi: 10.1016/j.cub.2007.06.022

Jiang, F., Lu, S., Yao, X., Yue, X., and Au, W. (2014). Up or down? How culture and color affect judgments. *J. Behav. Decis. Mak.* 27, 226–234. doi: 10.1002/bdm.1800

Kirkcaldy, B. D. (1982). Personality profiles at various levels of athletic participation. *Pers. Individ. Dif.* 3, 321–326. doi: 10.1016/0191-8869(82)90052-6

Klinger, M. R., Burton, P. C., and Pitts, G. S. (2000). Mechanisms of unconscious priming: I. Response competition, not spreading activation. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 441–455. doi: 10.1037/0278-7393.26.2.441

Krenn, B. (2014). The impact of uniform color on judging tackles in association football. *Psychol. Sport Exerc.* 15, 222–225. doi: 10.1016/j.psychsport.2013.11.007

Krenn, B. (2015). The effect of uniform color on judging athletes' aggressiveness, fairness and chances of winning. *J. Sport Exerc. Psychol.* 37, 207–212. doi: 10.1123/jsep.2014-0274

Little, A. C., and Hill, R. A. (2007). Attribution to red suggests special role in dominance signaling. *J. Evol. Psychol.* 5, 161–168. doi: 10.1556/JEP.2007.1008

Mehta, R., and Zhu, R. (2009). Blue or red? Exploring the effect of color on cognitive task performances. *Science* 323, 1226–1229. doi: 10.1126/science.1169144

Mentzel, S. V., Schücker, L., Hagemann, N., and Strauss, B. (2017). Emotionality of colors: an implicit link between red and dominance. *Front. Psychol.* 8:317. doi: 10.3389/fpsyg.2017.00317

Moller, A. C., Elliot, A. J., and Maier, M. A. (2009). Basic hue-meaning associations. *Emotion* 9, 898–902. doi: 10.1037/a0017811

Pervin, A. L., and John, P. O. (1997). *Personality: Theory and Research*, 7th Edn. New York, NY: John Wiley & Sons.

Piatti, M., Savage, D. A., and Torgler, B. (2012). The red mist? Red shirts, success and team sports. *Sport Soc.* 15, 1–28. doi: 10.1080/17430437.2012.690400

Plessner, H., and Haar, T. (2006). Sports performance judgments from a social cognitiveperspective. *Psychol. Sport Exerc.* 7, 555–575. doi: 10.1016/j.psychsport.2006.03.007

Pravossoudovitch, K., Cury, F., Young, S. G., and Elliot, A. J. (2014). Is red the colour of danger? Testing an implicit red–danger association. *Ergonomics* 57, 503–510. doi: 10.1080/00140139.2014.889220

Pryke, S. R., and Griffith, S. C. (2006). Red dominates black: agonistic signalling among head morphs in the colour polymorphic Gouldian finch. *Proc. R. Soc. Lond. B Biol. Sci.* 273, 949–957. doi: 10.2307/25223391

Şefik Tiryaki, M. (2005). Assessing whether black uniforms affect the decisions of Turkish soccer referees: is finding of Frank and Gilovich's study valid for Turkish culture? *Percept. Mot. Skills* 100, 51–57. doi: 10.2466/pms.100.1.51-57

Setchell, J. M., and Wickings, E. J. (2005). Dominance, status signals and coloration in male *Mandrills* (*Mandrillus sphinx*). *Ethology* 111, 25–50. doi: 10.1111/j.1439-0310.2004.01054.x

Shi, J., Zhang, C., and Jiang, F. (2015). Does red undermine individuals' intellectual performance? A test in China. *Int. J. Psychol.* 50, 81–84. doi: 10.1002/ijop.12076

Soriano, C., and Valenzuela, J. (2009). Emotion and colour across languages: implicit associations in Spanish color terms. *Soc. Sci. Inform.* 48, 421–445. doi: 10.1177/0539018409106199

Sorokowski, P., and Szmajke, A. (2011). The influence of the "red win" effect in sports: a hypothesis of erroneous perception of opponents dressed in red—preliminary test. *Hum. Mov.* 12, 367–373. doi: 10.2478/v10038-011-0043-5

Trninić, V., Barančić, M., and Nazor, M. (2008). The five-factor model of personality and aggressiveness in prisoners and athletes. *Kinesiology* 40, 170–181. doi: 10.2519/jospt.2008.0412

Valdez, P., and Mehrabian, A. (1994). Effects of color on emotions. *J. Exp. Psychol.* 123, 394–409. doi: 10.1037//0096-3445.123.4.394

Wiedemann, D., Burt, D. M., Hill, R. A., and Barton, R. A. (2015). Red clothing increases perceived dominance, aggression and anger. *Biol. Lett.* 11:20150166. doi: 10.1098/rsbl.2015.016

Williams, J. K., and Themanson, J. R. (2011). Neural correlates of the implicit association test: evidence for semantic and emotional processing. *Soc. Cogn. Affect. Neurosci*. 6, 468–476. doi: 10.1093/scan/nsq065

Wittenbrink, B., Judd, C. M., and Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *J. Pers. Soc. Psychol*. 72, 262–274. doi: 10.1037//0022-3514.72.2.262

Wogalter, M. S., Conzola, V. C., and Smith-Jackson, T. L. (2002). Research-based guidelines forwarning design and evaluation. *Appl. Ergon.* 33, 219–230. doi: 10.1016/S0003-6870(02)00009-1

Zhang, T., Feng, S., Han, B., and Sun, S. (2018). Red color in flags: a signal for competition. *Color Res. Appl.* 43, 114–118. doi: 10.1002/col.2216

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINAL RESEARCH
published: 10 February 2021
doi: 10.3389/fpsyg.2021.618507

Morality and Values in Sports Among Young Athletes: The Role of Sport Type and Parenting Styles – A Pilot Study

Yosi Yaffe[1†], Orr Levental[2†], Dalit Lev Arey[3] and Assaf Lev[4]*

[1]Department of Education, Tel-Hai Academic College, Qiryat Shemona, Israel, [2]Department of Phsyical Education, Tel-Hai Academic College, Qiryat Shemona, Israel, [3]Department of Psycology, Tel-Aviv-Yaffo Academic College, Tel Aviv-Yaffo, Israel, [4]Department of Sports Therapy, Ono Academic College, Kiryat Ono, Israel

OPEN ACCESS

Edited by:
Yair Galily,
Interdisciplinary Center Herzliya, Israel

Reviewed by:
Sharon Stoll,
University of Idaho, United States
Michael Leitner,
California State University, Chico

***Correspondence:**
Orr Levental
levental@ohalo.ac.il;
levental@gmail.com

†These authors share first authorship

Specialty section:
This article was submitted to Movement Science and Sport Psychology, a section of the journal Frontiers in Psychology

Received: 17 October 2020
Accepted: 11 January 2021
Published: 10 February 2021

Citation:
Yaffe Y, Levental O, Arey DL and Lev A (2021) Morality and Values in Sports Among Young Athletes: The Role of Sport Type and Parenting Styles – A Pilot Study. Front. Psychol. 12:618507. doi: 10.3389/fpsyg.2021.618507

Given the great importance of morality and values in modern sports, especially among young athletes, in this pilot study, we sought to broaden the exploration of the factors that may play role in these contexts, which have not been widely researched to date. Accordingly, the study tested the relationships between sport type (team or individual) and parenting styles (authoritative vs. non-authoritative), and moral decision-making in sport and sport values among 110 adolescent athletes whose age ranges from 11 to 22 (M = 16.04, SD = 2.86). The findings indicated that participants with authoritative parents, as compared to those with non-authoritative parents, are significantly less accepting of cheating in sport, while they also tend more to keep winning in proportion and hold significantly stronger moral values toward sports. Moreover, participants whose main sport is a team sport type tend to accept more cheating and gamesmanship than participants whose main sport is an individualistic sport type. While no differences were recorded between these groups in moral values, team athletes tend to value status in sport more than individual athletes, while the latter tend to value competence regarding their sport. The implications of the findings are discussed in light of no interaction between the effects of parenting styles and sport type on moral and sport values.

Keywords: parenting style, morality and values, Israel, adolescents, athleths

INTRODUCTION

The professionalization and commercialization of sports emphasize the need for wins over participation. This combination of sportsmanship and competitiveness has created a conflict between values and functionality (Levental, 2020). Thus, a moral dilemma arises, related to the desire and need to win against the importance of participation and following the rules of the game, or as Jones and David Howe (2005) suggest, a contest between fairness and merit (Jones and David Howe, 2005). According to the social learning theory, behavioral change is a product of reinforcement and punishment along with behavioral imitation of other meaningful individuals. Thus, there is importance in examining the internalization of moral standards among young athletes (Whitehead et al., 2013). Furthermore, the distinction made between the competition-game context and life itself should also be addressed.

According to Shields and Bredemeier (1994, 2007), the rules of some competitions often allow for greater egocentrism and moral flexibility than life outside of sports. That is, the environmental context influences ethical attitudes and behavior. Ethical perceptions among young athletes should, therefore, be examined in two aspects. The first is identifying values and their behavioral expressions (see Kavussanu, 2008), and the second is the effects of different variables, personal and environmental, on the formation of these perceptions. For example, the presence of other individuals or rewards and penalties (Lee and Cockmen, 2013). One aspect related to both environmental influences and behavior is the culture in which the sport is performed. A study by Whitehead and Goncalves (2013) showed that different sports values can be found in different countries. Demographic aspects may also have an impact on adopted values. More than half a century ago, Webb (1969) found that increasing age indicates a higher commitment to victory than fair play. In addition, with increasing age, there is a wider acceptance of the use of aggressive behavior during sports (Lee and Williams, 1989). The age factor was also examined in an article by Lee et al. (2013) that found that values associated with sport become less significant with increasing age. Another study by Lee et al. (2013) found that older athletes (14–16) more than young athletes (11–13) reported higher Acceptance of Cheating as well as Acceptance of Gamesmanship.

The present article focuses mainly on the implications of the type of sport – individual or team – on the ethical attitudes of young athletes. This is based on the understanding that structural aspects in both types of sport affect the athletes' environment and consequently their ethical perception. One example of the difference between athletes from a team sport vs. an individual sport can be found in the study of Boardley et al. (2015) about athletes' perceptions regarding the use of performance-enhancing drugs. They found that due to the team sharing responsibility for winning, athletes who are part of a team develop a more lenient approach regarding the possible benefit of using drugs to gain a competitive advantage. This is due to the fact that the illegal advantage gained by one athlete has a relatively marginal effect on the result of the game. This comparative aspect was also reflected in a study by Woolf and Mazanov (2017). However, while a difference was found concerning doping between athletes at different competitive levels, no differences were found in the perceptions of athletes from individual and team sports. Also, another study found no differences in terms of ranking sports values among young athletes from individual and team sports (Lee et al., 2013). Whereas regarding ethical decisions in youth sport, it was found that there are differences between athletes from team sports and individual sports in two out of three categories (Lee et al., 2013). Young athletes participating in team sports received a higher score on Acceptance of Cheating and Acceptance of Gamesmanship but were no different from individual sport athletes in Keeping Winning in Proportion. The difference between team and individual sports in terms of ethical perception may stem from the different approaches of coaches in different fields and emphasizing personal development vs. sporting success (Peláez et al., 2016). The moral reasoning of athletes in a team is greatly influenced both by its moral climate (Kavussanu et al., 2002) and by the orientation of team performance vs. mastery, which directs the behavioral norms of individuals (Ommundsen et al., 2003). Naturally, a major part of the moral perception of the sport depends on rivalry. Therefore, a moral-ethical difference could be the result of different perceptions of the individual opponent or rival team. Vallerand et al. (1997) found that team athletes tend to express less concern about their opponent than athletes in an individual sport. Moreover, the findings of Calmeiro et al. (2015) show that athletes from team sports exhibit moral reasoning at lower levels than athletes from individual sports. One of the explanations offered by the authors is that in team sports, due to the direct encounter with an opponent, self-improvement contributes to victory as well as impairing the opponent's abilities. Another explanation for this may be the collective moral responsibility and the reduction of the role of the individual athlete within the team (Rudd and Stoll, 2004; Boardley and Kavussanu, 2009).

When discussing young athletes' morality and values system in sport, the youngsters' individual differences in fundamental morality should also be considered. A central factor affecting children's and adolescents' development of morality is the family environment, with parents specifically playing the most important role in this issue (Smetana, 2015). Through socialization processes, parents impart the habits, values, and norms congruent with adaptation to their culture (Baumrind, 1980) which shape and construct their moral knowledge. As the primary caregivers and raisers who spend most time with the child and closest to them, parents normally serve as the main socialization agents whose role in the child's moral development is crucial. Parents are also central in the context of moral development due to their affective relationship with their children and their responsibility to educate and discipline them while instilling moral values (Wainryb and Recchia, 2014; Smetana, 2015). They convey moral values *via* both cognitive and affective components (Smetana, 1999) that underlie the family's social interactions and parent-child relationships. In the context of sports, the role of socialization is key to understanding aspects of an athlete's sensations, ideals, and values (see: Lev, 2019). Regarding family and the process through which young athletes learn the values and norms of sports culture, children's experiences can be molded based on their exposure to motivational, cognitive, and affective responses (Harwood et al., 2019). Family influences also affect young athletes' embrace of positive values and satisfaction, and their approach and comportment, in times of stress, toward sport and physical activity (Tamminen et al., 2016; Lev et al., 2020). In this regard, a recent study (Danioni et al., 2017) showed that adolescent athletes whose parents endorsed core values in sport gave great importance to competence values (e.g., skillsets and accomplishments) and moral values (e.g., living up to obligations and fairness), but attributed little importance to status values (e.g., titles and winning). In other words, adolescent athletes valued sportsmanship above gamesmanship. However, families and parents vary in a manner in which they socialize their children, which, according to a large body of research, may be related to various differences in children's moral, behavioral, and emotional development (Pinquart, 2017a,b; Pinquart and Gerke, 2019; Fatima et al., 2020). While both parents play a significant role in their child's sporting activity, there are differences in the roles that the genders play in conveying values and norms through the socialization of children regarding sports (Palomo-Nieto et al., 2011). In this context, socialization theorists and researchers have focused on different types of parental behaviors, practices, and styles within the family that influence the moral internalization in children, and which are assumed to subsequently effect the development of the children's broader personality and emotional characteristics. Parenting is a broad construct that comprises stable and durable attitudes and behaviors toward child-rearing (Smetana, 2017), while the theoretically widely-used terminology in the literature to describe its substance and types is parenting styles (Baumrind, 1966). The concept of parenting styles refers to certain types of parenting that are characterized by distinct attitudes and behaviors toward the child and constitute a form of marital climate (Yaffe, 2020). Three discerned styles of parenting are known in the literature as authoritative, indulgent (or permissive), and authoritarian (Maccoby and Martin, 1983; Baumrind, 1991), which generally differ by the practices they use and the type of control they exert in raising their children. Authoritative parents consistently incorporate behavioral control with providing warmth and emotional support and closeness (Baumrind, 1966; Maccoby and Martin, 1983). They establish their offspring's socialization on reasoning, negotiation, and shared decision-making, while setting consistent limits and rules and encouraging autonomy (Baumrind, 1966, 1968, 2005; Yaffe, 2020). Authoritarian parents display a discerned type of parental authority, as they exert strict and intrusive control over the child, while avoiding negotiation, using punishment, and maintaining an emotional distance. Finally, permissive parents tend to practice lax control, to avoid punishing, and to maintain emotional closeness (Baumrind, 1966, 1968, 2005).

Indeed, as compared to an offspring from non-authoritative families, adolescents who were raised by authoritative parents exhibit several developmental socio-emotional advantages: they academically outperform their counterparts at school (Spera, 2005; Pinquart and Kauser, 2018), manifest better psychological adjustment in terms of lower depression and anxiety (Pinquart, 2017a; Yaffe, 2018a), tend to be more morally developed while engaging with less behavior and externalized problems (Freeze et al., 2014; Pinquart, 2017b; Fatima et al., 2020), have higher self-esteem (Pinquart and Gerke, 2019; more), and are more likely to be health behaviorally oriented (Vollmer and Mobley, 2013; Yaffe, 2018b). Accordingly, Darling and Steinberg (1993) submitted that since offspring of authoritative parents tend to endorse their parents' parental authority and feel obligated to obey their rules, they are more convenient for parental socialization and, consequently, tend to internalize more intensively their parents' social and moral values.

The Current Study

This research is a pilot study aiming to explore the parenting and sport factors associated with young athletes' moral values in sports. In light of the inconsistencies reflected from the body of research on the moral-ethical values of young athletes from different sport types, the current study sought to broaden that scope by also accounting for the youngsters' familial background as an additional factor explaining their ethical values in sport. As demonstrated by the above literature review, parental socialization of children and adolescents plays a substantial role in their moral development, which could also specifically apply to young athletes' moral values in the context of sport. Yet, the research in the area of morality and values in sports among young athletes has focused on sport variables and has failed to consider adequately the role of interpersonal differences in familial background, such as parenting styles. In this respect, incorporating determinants of general morality (i.e., parenting) with those of sport morality (i.e., sport type) could be promising in advancing our understanding regarding athletes' moral and values system in sport, which becomes crucial with the growing popularity of sports in the modern era. In this regard, the following hypotheses are tested in the current study:

a. Young individual-sports athletes will differ from young team-sports athletes in moral decision-making in sport and sport values, so that the former would express lower acceptance of cheating and gamesmanship, a stronger tendency to keep winning in proportion, and stronger moral values in sport.
b. Young athletes who perceive their parents as authoritative will differ from young athletes who perceive their parents as either authoritarian or permissive in moral decision-making in sport and sport values, so that the former would express lower acceptance of cheating and gamesmanship, a stronger tendency to keep winning in proportion, and stronger moral values in sport.

MATERIALS AND METHODS

Participants and Procedure

The sample included 110 adolescent athletes (97 males and 13 females) whose age ranges from 11 to 22 (M = 16.04, SD = 2.86), with the majority of them Jewish students (84.5%) and the rest Arab students. They belonged to families whose sizes ranged from 2 to 10, with a mean size of 4.95 ± 1.25. The participants were all engaged in regular competitive and professional sports activities, which included cycling, aerobic sports, motor sports, tennis, soccer, basketball, fencing, gymnastics, triathlon, climbing, and swimming. As part of our research purposes, these sports were classified into two groups of sport type: individual sports and team sports. Of the sample of students, approximately 24% reported having some diagnosed learning disability. Since their scores on the dependent variables (i.e., the moral decision-making and system values in sport) did not statistically differ from those of the sample of students without learning disabilities, the former were equally included in the general sample as part of the analyses of the research hypotheses (apart from this reason, separate analyses for this group could not be applied due to its small sample size). The data collection procedure was carried out using a snowball sampling method, where the participants were recruited due to their regular engagement in professional sport *via* public and personal announcements to take part in the current pilot study. Participants (from organized sport groups or individually) were initially informed about a pilot study dealing with ethical perceptions in sport and were asked whether they would be willing to take part in completing anonymous questionnaires. Those students who expressed their interest were given an online link where they could read a detailed explanation about the study's objectives and the participation terms. Filling out the research forms was conditioned on signing an online informed consent in advance (for participants ages over 18) and, for minor participants, in giving their assent and obtaining parental permission for taking part in the study. An *ad-hoc* institutional ethics committee reviewed and verified the data collection program prior to applying its procedure.

Measures

Parental Authority Questionnaire

The Parental Authority Questionnaire (PAQ; Buri, 1991) contains 30 items and is used to classify parents into one of Baumrind's three parenting styles (Baumrind, 1971), based on the adolescent's self-report: *Authoritative* (10 items, e.g., "As I was growing up, once family policy had been established, my parents discussed the reasoning behind the policy with the children in the family"), *Authoritarian* (10 items, e.g., "As I was growing up my parents did not allow me to question any decision they had made"), and *Permissive* (10 items, e.g., "As I was growing up my parents seldom gave me expectations and guidelines for my behavior"). The response scales for an item range from 1 (strongly disagree) to 5 (strongly agree). The index for each parenting style is the sum of the relevant items of each scale. Thus, the total score for each parenting scale ranges from 10 to 50, with a higher score reflecting a higher specification of the style. PAQ is a valid questionnaire with relatively high internal consistency and test-retest reliabilities (0.74–0.78; see: Buri, 1991; Smetana, 1995), widely used in Israel (e.g., Enten and Golan, 2009; Yaffe, 2018a) and around the world to measure Baumrind's (1971) three basic styles of parenting. It was originally developed and validated in English using high-school and college students (Buri, 1991), and has previously shown supporting evidence of validity and reliability in its Hebrew version with early adolescents (Yaffe, 2018a). The current study recorded alpha coefficients for the permissive, authoritarian, and authoritative scales of 0.64, 0.88, and 0.79 (respectively). The scores obtained in the current sample for the instrument's scales appear in **Table 1**.

Attitudes to Moral Decision-Making in Youth Sport Questionnaire

This is 9-item questionnaire with three scales designed to measure attitudes toward *Acceptance of Gamesmanship* (e.g., "Sometimes I waste time to unsettle the opposition"), Acceptance of Cheating (e.g., "I would cheat if I thought it would help me win"), and *Keeping Winning in Proportion* (e.g., "It's OK to lose sometimes because in life you do not win everything"). In the current study, we recorded Cronbach's alpha values of 0.80, 0.83, and 0.72 for these three scales, respectively, which are adequate reliability indexes of internal consistency given the small number of items included in the scales. Each scale comprises three items whose responses are given on a 1–5 Likert response scale. Responses should be averaged to produce the scales' *mean score*. The Attitudes to Moral Decision-Making in Youth Sport Questionnaire (AMDYSQ) was concurrently validated against several external

TABLE 1 | Descriptive statistics and correlation matrix for the research variables.

S. No.	Variable	1	2	3	4	5	6	7	8	9
Parenting styles										
1.	Permissive	-								
2.	Authoritarian	−42***	-							
3.	Authoritative	0.38***	−0.42***	-						
Moral decision-making in sport										
4.	Acceptance of cheating	0.05	0.26***	−0.31***	-					
5.	Acceptance of gamesmanship	−0.01	0.17	−0.15	0.48***	-				
6.	Keeping winning in proportion	0.01	0.03	0.21*	−0.22*	−0.03	-			
Sport values										
7.	Moral values	0.05	−0.14	0.30**	−0.56***	−0.40***	0.50***	-		
8.	Competence values	0.15	−0.17	0.11	−0.18	−0.08	0.25**	0.53***	-	
9.	Status values	−0.03	0.22*	−0.25**	0.33***	0.41***	0.04	0.03	0.41***	-
	Mean	27.48	26.93	35.0	2.06	2.79	4.20	3.50	4.02	2.90
	SD	5.42	8.08	6.48	0.90	1.04	0.74	0.87	0.76	1.06

*$p \leq 0.05$.
**$p \leq 0.01$.
***$p \leq 0.001$.

variables (such as commitment to sport, respect for conventions, respect for opponents, and respect for rules) and its measurement demonstrated gender invariance (Lee et al., 2013). The AMDYSQ's subscales have also been shown to have negligible associations with social desirability scale. Normally, the *Acceptance of Gamesmanship* and *Acceptance of Cheating* scales correlate positively, while each of these scales correlates negatively with the *Keeping Winning in Proportion* scale (Lee et al., 2007).

Youth Sport Values Questionnaire

This 13-item questionnaire is a modified version of the original Youth Sport Values Questionnaire (YSVQ; (Lee et al., 2000, 2008), designed to measure young athletes' values system in sports. Its three subscales contain higher-order measure *moral values* (five items; e.g., "I try to be fair"), *competence values* (four items; e.g., "I improve my performance"), and *status values* (four items; e.g., "I win or beat others"; Lee et al., 2013). The mean scores are obtained by averaging the item scores for each scale. Normally, *Competence* values correlate more highly with *moral* values than with *status* values and the correlation between *moral* and *status* values is lowest of all (Lee et al., 2013). The English YSVQ-2 was translated and used in studies across, at least, eight different countries outside the United Kingdom, which their findings underpinning the suitability of the questionnaire for different cultural contexts. The YSVQ-2's subscales have been shown in previous studies to relate logically to the AMDYSQ's subscales, especially the moral and competence values, which were negatively correlated with the acceptance of cheating and gamesmanship (see Whitehead et al., 2013). The developers reported a good 4-week test-retest reliability indication for the questionnaire's scales. In the current study, we obtained acceptable indices for the scales' internal consistency reliability (that is, the *moral* values, the *status* values, and the *status* values scales), with their Cronbach's alpha coefficients ranging between 0.71 and 0.80. For the purposes of the current study, the authors adapted and translated the English forms of the AMDYSQ and the YSVQ-2 using the back-forward translation method. The scores obtained for these scales in the current study are presented in **Table 1**.

Both sport questionnaires (i.e., the AMDYSQ and the YSVQ-2) were developed with the 12–16 years age group of young athletes, implying that their content could fit close and older age populations.

Ethics

The study involving human participants (as described above) was reviewed and approved by an institutional ethics committee of Ohalo Academic College, Israel.

RESULTS

At first, we tested the scales' scores and the zero-order correlations between the study variables, to identify the general tendencies of the results. The study's operational variables included the parenting styles scales, the moral decision-making, and the sport values scales, which each contains three subscales. As appears in **Table 1**, the parenting scales are significantly intercorrelated in accordance with the expected directions, with the authoritative and permissive scales being reversely associated with the authoritarian scale. The parenting scales also exhibited numerous significant connections with the moral decision-making and sport values subscales, as the authoritative parenting was negatively correlated with acceptance of cheating and status values and positively correlated with moral values and keeping winning in proportion. Relative to the authoritative scale, the authoritarian scale was inversely correlated with most of the moral decision-making and sport values subscales. That is to say that young athletes who perceive their parents as more authoritative (rather than authoritarian) are more likely to embrace moral values in sport, to keep winning in proportion, and to reject cheating.

We also observed some significant correlations between the moral decision-making and the sport values subscales, which partially reflect the correspondence and agreement between the measurements. Hence, acceptance of cheating was considerably correlated with the accepting of gamesmanship, while these two variables were simultaneously associated with lower levels of moral values and higher levels of status values in sport. And, while the sample of athletes who hold high moral values less favored cheating and gamesmanship, they did value competence in their sport and tended more to keep winning in proportion.

Testing the Research Hypotheses

We generally hypothesized that young athletes' morality and values in sport would differ by their perceptions of their parenting styles and their main sport type. As mentioned, the dependent variables were measured in the current study *via* the youth "moral decision making" in sport and the "youth sport values" scales. Parenting styles were classified into groups using the highest score among the three continuous scales (i.e., permissive, authoritarian, and authoritative). This parenting categorization resulted in considerably unbalanced group sizes, with the vast majority classified as authoritative, forcing us to use only two groups of parenting (that is, authoritative and non-authoritative, which includes permissive and authoritarian parenting). Although parents who use demandingness (that is, authoritarian) and parents who do not use demandingness (that is, permissive) are essentially different in their styles, classifying them as non-authoritative parents is theoretically logical and has been operationally utilized in previous studies for various comparison-based purposes against authoritative parents (Steinberg et al., 1991; Zuquetto et al., 2019). To test the research hypotheses, we conducted a multivariate ANOVA (MANOVA) for the differences in these scales' scores by parenting style and sport type (displayed in **Table 2**). Since the study's demographics had merely negligible effects (for the most part not statistically significant) on the dependent variables, they were not taken into account as part of the analysis.

The multivariate tests indicated significant general effects of parenting styles (Wilks' lambda = 0.84, $p < 0.01$) and sport type (Wilks' lambda = 0.69, $p < 0.001$) on the dependent variables. An inspection of the individual effects using the between-subjects tests (**Table 2**) revealed a significant main effect of parenting styles on the cheating, proportion, and the moral values subscales, with the largest effect size observed on cheating. This indicates that participants with authoritative parents, as compared to those with non-authoritative parents, are significantly less accepting of cheating in sport, while they also tend more to keep winning in proportion. These participants (i.e., youngster athletes with authoritative parents) also hold significantly stronger moral values toward sports, but they do not differ from their counterparts (i.e., youngster athletes with non-authoritative parents) in the acceptance of gamesmanship and in competence and status values.

Furthermore, we recorded main effects for the sport type on the subscales of cheating, gamesmanship, competence, and status values, with the largest effect size observed on gamesmanship. This means that participants whose main sport is team sport type tend to accept more cheating and gamesmanship than participants whose main sport is an individualistic sport type. While no differences were recorded between these groups in moral values, they did significantly differ in their values of competence and of status. Thus, team athletes tend to value status in sport more than individual athletes, while the latter tend to value more competence regarding their sport. These differences in sport values by sport type were of the smallest size as compared to the other effects observed here, yet they were statistically significant. Finally, whereas both parenting styles and sport type affected several variables of moral decision-making and the sport values simultaneously, their effects in either of these contexts did not interact. This generally denotes that differences between the sample of participants with authoritative parents and those with non-authoritative parents in moral decision-making and sport values apply similarly for both team athletes and individual athletes (and vice versa).

DISCUSSION

The study sought to examine the relationships between young athletes' sport type and parenting styles and their moral and values system in sport, in an attempt to determine whether and how young athletes vary in their moral decision making in sport and sport values by the type of sport they specialize in (i.e., individual vs. team sport) and by their parents' background (i.e., authoritative vs. non-authoritative parenting style). The first research hypothesis held that young athletes engaged in individual sports express lower acceptance of cheating and gamesmanship, a stronger tendency to keep winning in proportion, and stronger moral values in sport. The findings of the study show that there is indeed a significant difference between them and athletes from team sports. These findings are consistent with those of Boardley et al. (2015), according to which group effort overshadows the individual and, therefore, forms of cheating, such as taking drugs, do not greatly impact the result. Similar results were also found by Lee et al. (2013), who found that athletes in individual sports received lower scores on Acceptance of Cheating and Acceptance of Gamesmanship. However, in contrast to the current study, no difference was found regarding Keeping Winning in Proportion. A possible reason for this is that cultural differences in moral standards in different societies (Whitehead and Goncalves, 2013). In contrast to the attitude toward cheating, which relies on shared responsibilities, the importance of winning is a product of social perceptions. Thus, local norms of competitive behavior shape the perceptions of young athletes. Furthermore, the competitive level may have an impact, as found in the studies of Woolf and Mazanov (2017). According to their study, the differences in ethical perceptions between the team and individual athletes decrease when the level of play is non-competitive. This is partially similar to the sports in which the current study's participants take part.

The findings of the present study show no differences between the two groups in their moral values. These findings are

TABLE 2 | Means, SDs, and the results of multivariate ANOVA (MANOVA) analysis for the differences between youth's moral decision making in sport and sport values by parenting style and sport type.

	Parenting style		F (1,106)	Partial $\eta 2$	Sport		F (1,106)	Partial $\eta 2$
	Authoritative (n = 67)	Non-authoritative (n = 43)			Team (n = 65)	Individual (n = 45)		
Moral decision-making								
Cheating	1.82 (0.79)	2.44 (0.94)	13.07***	0.110	2.27 (0.88)	1.76 (0.85)	6.79**	0.060
Gamesmanship	2.70 (1.10)	2.93 (0.93)	0.83	0.001	3.14 (0.94)	2.28 (0.97)	17.64***	0.143
Proportion	4.29 (0.72)	4.05 (0.73)	4.28*	0.040	4.26 (0.65)	4.11 (0.84)	2.19	0.020
Sport values								
Moral values	3.69 (0.82)	3.20 (0.88)	9.78**	0.084	3.45 (0.69)	3.58 (1.09)	0.03	0.000
Competence values	4.09 (0.75)	3.91 (0.77)	0.51	0.001	3.89 (0.63)	4.19 (0.88)	5.02*	0.045
Status values	2.74 (1.15)	3.14 (0.85)	2.85	0.026	3.08 (0.94)	2.62 (1.06)	4.41*	0.040

Partial $\eta 2$ is a measure of effect size.
$^{*}p \leq 0.05$.
$^{**}p \leq 0.01$.
$^{***}p \leq 0.001$.

consistent with previous studies, such as those of Lee et al. (2013). As previously argued, earlier studies found that athletes from team sports exhibit moral reasoning at lower levels than athletes from individual sports. This is because team athletes express less concern for their opponent. The proposed rationale is that in a direct encounter, impairment of the opponent's abilities is tantamount to self-improvement, or alternatively, the collective responsibility that diminishes the influence of the individual in the group (Boardley and Kavussanu, 2009). These aspects were expressed in the fact that the young athletes from the team sports emphasized the importance of status while the athletes in the individual sports preferred competence. An explanation for this can be found in the effect of group climate on moral reasoning (Kavussanu et al., 2002) and by the orientation of team performance (Ommundsen et al., 2003). Because the teams' goal is collective success, self-improvement is not necessarily measurable and, therefore, receives less attention.

Our second hypothesis dealt with the relationships between young athletes' parents' parenting styles and their moral and values system in sport, assuming that offspring who were raised by different types of parents may vary in their moral development and their fundamental morality. The current study's findings are in line with previous findings suggesting that young athletes raised by parents who communicate high morals and values as part of socialization thus carry those traits into their sport activity (Danioni et al., 2017). Partially consistent with our hypothesis, we found that the sample of youngsters who perceived their parents as authoritative, compared to their peers who perceived their parents as non-authoritative, expressed lower acceptance of cheating and a stronger tendency to keep winning in proportion. Along with the formers' stronger tendency to hold stronger moral values in sport, our findings generally suggest that young athletes from authoritative families are more morally oriented in sports than those from non-authoritative families. Assuming that these moral tendencies in sport reflect the youngsters' wider moral value system, this conclusion principally corroborates the premise that adolescents whose parents are authoritative are more prone to parental socialization and, therefore, more intensively internalize and apply their parents' social and moral values in various areas of their lives (Darling and Steinberg, 1993). Indeed, the current study's findings are consistent with previous findings suggesting that children and adolescents who were raised by authoritative parents might be more morally developed than those who were raised by non-authoritative parents (Freeze et al., 2014; Pinquart, 2017b; Fatima et al., 2020). Interestingly, the effects of parenting styles on moral and sport values discovered here did not interact with the athletes' sport type, which means that the differences between athletes from authoritative and non-authoritative families in these variables apply equally to both individual and team sports.

Moreover, alongside other social and emotional differences observed between offspring with a discerned parental background (Spera, 2005; Pinquart, 2017a,b; Pinquart and Kauser, 2018; Yaffe, 2018a), this study initially reveals another unique outcome variable to which parenting styles could be related in adolescence. While studies dealing with the consequences of parent-child relationships on adolescents' development tend to focus primarily on central psychological, educational, and mental health variables (e.g., academic achievements, anxiety, depression and self-esteem, pro-social behavior, etc.; Steinberg, 2001), the current findings are the first to demonstrate the exclusive link between parenting styles and the outcomes of young athletes' morality in sport. Apart from the significance these findings may bear in terms of the importance of parenting with relation to child-socialization, they also can have implications for the origins of sport morality among professional athletes. In this context, perhaps our findings imply that athletes' moral values in sports are rooted in the fabric of their relationships with their parents, whose parenting styles primarily shaped their more general moral thinking. Since the current study design is cross-sectional in nature, a causal conclusion cannot be drawn from its findings regarding the relationship between parenting styles and young athletes' morality in sport. Therefore, this assumption must be further inspected in studies using a more advanced research design suitable to determine the causal direction in the associations between these variables. Given that most studies about moral and values in sport have focused exclusively on sport variables, broadening the inspection of the origins of morality in sports to family background among young athletes might be promising.

The study findings are limited in several respects. First and foremost, with the sample of athletes composed mainly of men, the study's conclusions regarding both effects sport type and parenting styles on moral values in sports, could not be generalized equally for women athletes. Further study with more balanced gender-distribution is required to determine how young men and women athletes differ on the basis of the study variables. Also, with a relatively small sample size which is unrepresentative of the young athletes' population in Israel, the findings of this pilot study should be treated cautiously and predominantly serve to cultivate further topic-related research.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the institutional ethics committee of Ohalo Academic College, Israel. Participants provided their written informed consent to participate in the study.

AUTHOR CONTRIBUTIONS

DLA and AL contributed to analyzing the data and interpreted the findings. They also expanded the literature review following the reviewers' comments. DLA and AL, along with OL and YY wrote and approved the final revision of the article.

REFERENCES

Baumrind, D. (1966). Effects of authoritative parental control on child behavior. *Child Dev.* 37, 887–907.

Baumrind, D. (1968). Authoritarian vs. authoritative parental control. *Adolescence* 3, 255–272.

Baumrind, D. (1971). Current patterns of parental authority. *Dev. Psychol. Monogr.* 4, 1–103. doi: 10.1037/h0030372

Baumrind, D. (1980). New directions in socialization research. *Am. Psychol.* 35:639. doi: 10.1037/0003-066X.35.7.639

Baumrind, D. (1991). The influence of parenting style on adolescent competence and substance use. *J. Early Adolesc.* 11, 56–95. doi: 10.1177/0272431691111004

Baumrind, D. (2005). Patterns of parental authority and adolescent autonomy. *New Dir. Child Adolesc. Dev.* 108, 61–69. doi: 10.1002/cd.128

Boardley, I. D., Grix, J., and Harkin, J. (2015). Doping in team and individual sports: a qualitative investigation of moral disengagement and associated processes. *Qua. Res. Sport Exercise Health* 7, 698–717. doi: 10.1080/2159676X.2014.992039

Boardley, I. D., and Kavussanu, M. (2009). Effects of goal orientation and perceived value of toughness on antisocial behavior in soccer: the mediating role of moral disengagement. *J. Sport Exercise Psy.* 32, 176–192. doi: 10.1123/jsep.32.2.176

Buri, J. R. (1991). Parental authority questionnaire. *J. Pers. Assess.* 57, 110–119. doi: 10.1207/s15327752jpa5701_13

Calmeiro, L., Stoll, S. K., and Davis, P. (2015). Moral reasoning in sport: validation of the Portuguese version of the RSBH value-judgement inventory in adolescents. *Sport Sci. Rev.* 24, 285–304. doi: 10.1515/ssr-2015-0020

Danioni, F., Barni, D., and Rosnati, R. (2017). Transmitting sport values: the importance of parental involvement in children's sport activity. *Eur. J. Psychol.* 13:75. doi: 10.5964/ejop.v13i1.1265

Darling, N., and Steinberg, L. (1993). Parenting style as context: an integrative model. *Psychol. Bull.* 113, 487–496. doi: 10.1037/0033-2909.113.3.487

Enten, R. S., and Golan, M. (2009). Parenting styles and eating disorder pathology. *Appetite* 52, 784–787. doi: 10.1016/j.appet.2009.02.013

Fatima, S., Dawood, S., and Munir, M. (2020). Parenting styles, moral identity and prosocial behaviors in adolescents. *Curr. Psychol.*, 1–9. doi: 10.1007/s12144-020-00609-3

Freeze, M. K., Burke, A., and Vorster, A. C. (2014). The role of parental style in the conduct disorders: a comparison between adolescent boys with and without conduct disorder. *J. Child Adolesc. Ment. Health* 26, 63–73. doi: 10.2989/17280583.2013.865627

Harwood, C. G., Knight, C. J., Thrower, S. N., and Berrow, S. R. (2019). Advancing the study of parental involvement to optimise the psychosocial development and experiences of young athletes. *Psychol. Sport Exerc.* 42, 66–73. doi: 10.1016/j.psychsport.2019.01.007

Jones, C., and David Howe, P. (2005). The conceptual boundaries of sport for the disabled: classification and athletic performance. *J. Philos. Sport* 32, 133–146. doi: 10.1080/00948705.2005.9714678

Kavussanu, M. (2008). Moral behaviour in sport: a critical review of the literature. *Int. Rev. Sport Exerc. Psychol.* 1, 124–138. doi: 10.1080/17509840802277417

Kavussanu, M., Roberts, G. C., and Ntoumanis, N. (2002). Contextual influences on moral functioning of college basketball players. *J. Sport Psychol.* 16, 347–367. doi: 10.1123/tsp.16.4.347

Lee, M. J., and Cockman, M. (2013). "What sport values do young people hold?" in *Values in youth sport and physical education*. eds. J. Whitehead, H. Telfer and J. Lambert (London: Routledge), 37–48.

Lee, M. J., Whitehead, J., and Balchin, N. (2000). The measurement of values in youth sport: development of the youth sport values questionnaire. *J. Sport. Exerc. Psychol.* 22, 307–326. doi: 10.1123/jsep.22.4.307

Lee, M. J., Whitehead, J., and Ntoumanis, N. (2007). Development of the attitudes to moral decision-making in youth sport questionnaire (AMDYSQ). *J. Sport. Exerc. Psychol.* 8, 369–392. doi: 10.1016/j.psychsport.2006.12.002

Lee, M. J., Whitehead, J., and Ntoumanis, N. (2008). Relationships among values, achievement orientations, and attitudes in youth sport. *J. Sport. Exerc. Psychol.* 30, 588–610. doi: 10.1123/jsep.30.5.588

Lee, M. J., Whitehead, J., and Ntoumanis, N. (2013). "How important are ethical attitudes? Development of the attitudes to moral decision-making in youth sport questionnaire" in *Values in youth sport and physical education*. eds. J. Whitehead, H. Telfer and J. Lambert (London: Routledge), 86–104.

Lee, M. J., and Williams, V. (1989). Over the top. Sport and Leisure, March–April, 27–28.

Lev, A. (2019). Becoming a long-distance runner–deriving pleasure and contentment in times of pain and bodily distress. *Leis. Stud.* 38, 790–803. doi: 10.1080/02614367.2019.1640776

Lev, A., Bichman, A., Moyal, A., Brenner, S., Fass, N., and Been, E. (2020). No cutting corners: the effect of parental involvement on youth basketball players in Israel. *Front. Psychol.* 11:607000. doi: 10.3389/fpsyg.2020.607000

Levental, O. (2020). Built heritage or lost nostalgia: Israeli fans and the conservation of sports venues. *Israel Stu.* 26, 1–16. doi: 10.1080/13537121.2020.1775938

Maccoby, E. E., and Martin, J. A. (1983). "Socialization in the context of the family: parent-child interaction" in *Handbook of child psychology,* Socialization, personality, and social development. *Vol. 4*. eds. P. H. Mussen and E. M. Hetherington (New York: Wiley), 1–101.

Ommundsen, Y., Roberts, G. C., Lemyre, P. N., and Treasure, D. (2003). Perceived motivational climate in male youth soccer: relations to social–moral functioning, sportspersonship and team norm perceptions. *Psychol. Sport Exerc.* 4, 397–413. doi: 10.1016/S1469-0292(02)00038-9

Palomo-Nieto, M., Ruiz-Perez, L. M., Sanchez-Sanchez, M., and García-Coll, V. (2011). Young athletes' perceptions of the role of their mothers in their sport career. *Mot. Eur. J. Hum. Mov.* 26, 141–157.

Peláez, S., Aulls, M. W., and Bacon, S. L. (2016). Morality in sport: the coach's perspective. *Int. J. Sports Sci. Coach.* 11, 237–249. doi: 10.1177/1747954116637494

Pinquart, M. (2017a). Associations of parenting dimensions and styles with externalizing problems of children and adolescents: an updated meta-analysis. *Dev. Psychol.* 53, 873–932. doi: 10.1037/dev0000295

Pinquart, M. (2017b). Associations of parenting dimensions and styles with internalizing symptoms in children and adolescents: a meta-analysis. *Marriage Fam. Rev.* 53, 613–640. doi: 10.1080/01494929.2016.1247761

Pinquart, M., and Gerke, D. C. (2019). Associations of parenting styles with self-esteem in children and adolescents: a meta-analysis. *J. Child Fam. Stud.* 28, 1–19. doi: 10.1007/s10826-019-01417-5

Pinquart, M., and Kauser, R. (2018). Do the associations of parenting styles with behavior problems and academic achievement vary by culture? Results from a meta-analysis. *Cult. Divers. Ethn. Minor. Psychol.* 24, 75–100. doi: 10.1037/cdp0000149

Rudd, A., and Stoll, S. (2004). What type of character do athletes possess? An empirical examination of college athletes versus college non athletes with the RSBH value judgment inventory. *Sport J.* 7, 1–10.

Shields, D. L. L., and Bredemeier, B. J. L. (1994). *Character development and physical activity*. Champaign, IL: Human Kinetics.

Shields, D. L. L., and Bredemeier, B. J. L. (2007). "Advances in sport morality research" in *Handbook of sport psychology*. *3rd Edn*. eds. G. Tenenbaum and R. C. Ecklund. (Hoboken, NJ: Wiley and Son), 662–684.

Smetana, J. G. (1995). Parenting styles and conceptions of parental authority during adolescence. *Child Dev.* 66, 299–316. doi: 10.2307/1131579

Smetana, J. G. (1999). The role of parents in moral development: a social domain analysis. *J. Moral Educ.* 28, 311–321. doi: 10.1080/030572499103106

Smetana, J. G. (2015). Talking the talk and walking the walk: conversational pathways to moral development. *Hum. Dev.* 58, 301–307. doi: 10.1159/000439012

Smetana, J. G. (2017). Current research on parenting styles, dimensions, and beliefs. *Curr. Opin. Psychol.* 15, 19–25. doi: 10.1016/j.copsyc.2017.02.012

Spera, C. (2005). A review of the relationship among parenting practices, parenting styles, and adolescent school achievement. *Educ. Psychol. Rev.* 17, 125–146. doi: 10.1007/s10648-005-3950-1

Steinberg, L. (2001). We know some things: parent–adolescent relationships in retrospect and prospect. *J. Res. Adolesc.* 11, 1–19. doi: 10.1111/1532-7795.00001

Steinberg, L., Mounts, N. S., Lamborn, S. D., and Dornbusch, S. M. (1991). Authoritative parenting and adolescent adjustment across varied ecological niches. *J. Res. Adolesc.* 1, 19–36.

Tamminen, K. A., McEwen, C. E., and Crocker, P. R. (2016). Perceived parental support, pressure, and the socialization of adolescent athletes' coping. *Int. J. Sport Psychol.* 47, 335–354. doi: 10.7352/IJSP2016.47.335

Vallerand, R. L., Deshaies, P., and Cuerrier, J. P. (1997). On the effects of the social context on behavioral intentions of sportsmanship. *Int. J. Sport Psychol.* 28, 126–140.

Vollmer, R. L., and Mobley, A. R. (2013). Parenting styles, feeding styles, and their influence on child obesogenic behaviors and body weight. A review. *Appetite* 71, 232–241. doi: 10.1016/j.appet.2013.08.015

Wainryb, C., and Recchia, H. E. (2014). *Talking about right and wrong: Parent-child conversations as contexts for moral development*. New York, NY: Cambridge University Press.

Webb, H. (1969). "Professionalization of attitudes towards play among adolescents" in *Aspects of contemporary sport sociology*. ed. G. S. Kenyon (Athletic Institute: Chicago, IL), 161–128.

Whitehead, J., and Goncalves, C. E. (2013). "Are sport values similar in other nations? Exploring cross-cultural value system" in *Values in youth sport and physical education*. eds. J. Whitehead, H. Telfer and J. Lambert (New York: Routledge), 104–118.

Whitehead, J., Telfer, H., and Lambert, J. (2013). *Values in youth sport and physical education*. New York: Routledge.

Woolf, J., and Mazanov, J. (2017). How athletes conceptualise doping, winning, and consequences: insights from using the cognitive interviewing technique with the Goldman dilemma. *Qua. Res. Sport Exercise Health* 9, 303–320. doi: 10.1080/2159676X.2016.1272480

Yaffe, Y. (2018a). Establishing specific links between parenting styles and the S-anxieties in children: separation, social, and school. *J. Fam. Issues* 39, 1419–1437. doi: 10.1177/0192513X17710286

Yaffe, Y. (2018b). Physical activity among Israeli-Arab adolescent males: how do parenting styles matter? *Am. J. Mens Health* 12, 2037–2043. doi: 10.1177/1557988318790881

Yaffe, Y. (2020). "Parental authority: a contemporary integrative-theoretical conceptualization" in *A closer look at parenting styles and practices*. ed. N. Roman (New York: Nova Science Publishers, Inc.), 73–96.

Zuquetto, C. R., Opaleye, E. S., Feijó, M. R., Amato, T. C., Ferri, C. P., and Noto, A. R. (2019). Contributions of parenting styles and parental drunkenness to adolescent drinking. *Braz. J. Psychiatr.* 41, 511–517. doi: 10.1590/1516-4446-2018-0041

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 16 November 2020
doi: 10.3389/fpsyg.2020.607000

No Cutting Corners: The Effect of Parental Involvement on Youth Basketball Players in Israel

Assaf Lev[1†], Adi Bichman[2], Avi Moyal[2], Shmulik Brenner[2], Nir Fass[2] and Ella Been[1,3†]*

[1] Department of Sports Therapy, Faculty of Health Professions, Ono Academic College, Kiryat Ono, Tel Aviv-Yafo, Israel, [2] School of Education, Ono Academic College, Kiryat Ono, Israel, [3] Department of Anatomy and Anthropology, Sackler Faculty of Medicine, Tel Aviv University, Israel

OPEN ACCESS

Edited by:
Edson Filho,
Boston University, United States

Reviewed by:
Simon Pack,
St. John's University, United States
Michael Leitner,
California State University, Chico, United States

***Correspondence:**
Assaf Lev
anthropolev@gmail.com

†ORCID:
Assaf Lev
orcid.org/0000-0002-6254-1326
Ella Been
orcid.org/0000-0002-6279-6935

Specialty section:
This article was submitted to Movement Science and Sport Psychology, a section of the journal Frontiers in Psychology

Received: *16 September 2020*
Accepted: *15 October 2020*
Published: *16 November 2020*

Citation:
Lev A, Bichman A, Moyal A, Brenner S, Fass N and Been E (2020) No Cutting Corners: The Effect of Parental Involvement on Youth Basketball Players in Israel. Front. Psychol. 11:607000. doi: 10.3389/fpsyg.2020.607000

This paper explores the nature of parental involvement in youth basketball in Israel with regard to parenting style and in the context of dilemmas and ethical issues. It is well established that parental involvement in their child's sporting activity has vast implications on the child's motivation and enjoyment. With reference to Israeli society, only a few studies have focused on this subject. In order to address this lacuna, we used two questionnaires, given to 173 youth basketball players (child questionnaire) and their parents (parent questionnaire). Key findings illustrate three main themes. First, a higher level of satisfaction and contentment among basketball players whose parents demonstrated greater involvement; second, that parental emotional involvement is the most important variable for young athletes' satisfaction; and finally, differences in gender roles reveal that fathers are more involved with logistics, while mothers are more dominant in emotional involvement. Moreover, the findings demonstrate that parents should mainly place emphasis on emotional involvement. However, we suggest that parents do not bypass logistical care as this may create opportunities for greater emotional support and therefore greater child satisfaction.

Keywords: parent–child relationship, basketball, competitive sports, questionnaires, child's satisfaction

INTRODUCTION

Over the course of the last two generations, there has been growing awareness by parents relating to their child's development in sports (Wheeler and Green, 2014; Harwood et al., 2019; Knight, 2019). While exploring this shift, one must take into account cultural differences (Cheung and Pomerantz, 2011; Knight et al., 2017), parent–teacher or parent–coach factors, as well as historical, demographic, political, and economic issues (Hornby and Lafaele, 2011; Harwood et al., 2019). Parents have become more involved in their child's sports activities, which has implications on the parent–child relationship and the complex socialization process that a child goes through (Dorsch et al., 2015). The reasons parents decide to become involved in their child's sport activity are varied. Some regard it as a model of good parenting, while others use it as a way to develop intimacy with their child (Stefansen et al., 2018). These dynamics can create conflicts such as child frustration due to the pressure of high parental expectations or a feeling of disappointment by the parent if the child is not interested in sport (Sorkkila et al., 2017). The parents and the child may have different perspectives on parental involvement, which can lead to discord. While the parent regards the various forms of more active engagement in the child's sport as encouraging and good parenting, the child may view this engagement as criticism and pressure (Elliott and Drummond, 2013). Considering the social nature of competitive sport and the emphasis placed on achievement in society, the relationship between parents and their child becomes a critical element in the healthy development of young athletes today (Welk et al., 2003).

Much has been written about the positive relationship between active parents and the good physical activity habits of their child, and the negative effect of less active parents (Cleland et al., 2005; Eriksson et al., 2008; Tate et al., 2015; Folle et al., 2018; Rodrigues et al., 2018). However, as not all parents are the same, it is important to recognize the differing attitudes among them and to consider this diversity in an attempt to create more tailored interventions and educational materials (Knight et al., 2017).

Parental Influence on the Youth Player

Parents are usually involved in their child's sport activity from a young age and have an influence on the child's initial involvement and long-term participation in sport (Fraser-Thomas and Côté, 2009). This involvement can have a positive or negative influence on a child's sporting development, experiences, and emotions (Knight and Holt, 2014). The importance of parents' behavior for their child's enjoyment and motivation in sports is broadly described by the Sánchez-Miguel et al. (2013) study, which demonstrates a positive relationship between parental support of the sport and a player's enjoyment and motivation. Moreover, it suggests that in Spanish players who perceive more pressure from their parents, there is a positive association with motivation and a negative association with enjoyment. Further evidence is found in professional female gymnasts in Brazil (Nunomura and Oliveira, 2013), where parental support enables and greatly influences a child's entrance into sports, their access to the practice of a sport, their level of participation, their degree of involvement, and their physical and emotional well-being. However, when this support is perceived negatively, it can result in stress, conflicts between parents and the child, burnout, and eventual dropout. This negative influence is observed in a study, which examines the relationship of young Australian football players and their parents after the match ended (Elliott and Drummond, 2013). While the parents perceive their involvement, including verbally assessing the child's performance as useful, the child regards it as criticism or negativity. The end result is a negative effect on the child's performance—if a child receives criticism for dropping a mark (catch), they may deliberately avoid future attempts to mark the ball. In order to optimize parental involvement, it is important to understand the individual child's preferences and needs in sport and the gap between the parents' and the child's perspectives (Knight et al., 2010; Knight and Holt, 2013; Knight and Holt, 2014; Knight, 2019). It is also relevant to highlight the area of sport-based positive youth-development (PYD) in the context of parental involvement (Holt and Sehn, 2008; Harwood et al., 2019; Holt et al., 2020). In this regard, Harwood et al. (2019) stressed that sport parents' influence on their child's psychosocial development is less explored. Moreover, they highlight the importance of improving the study of parental involvement in relation to the experience and development of young athletes. In this context, the current study aims to address this issue by shining a spotlight on parental involvement and the child's satisfaction and contentment.

Israeli Parental Involvement in Sporting Activity

Parental involvement in their child's sports is a global phenomenon and, for this reason, should be examined within the sociocultural context. Ultimately, the linkage between parental involvement and its effect on a child's subsequent perceptions regarding their sport's field is influenced by broad sociocultural characteristics (Knight et al., 2017; Harwood et al., 2019; Holt et al., 2020). For example, Kang et al. (2015) find significant correlation between the self-esteem of Korean high school athletes and a level of trust and communication with their parents. Knight et al. (2016) show that parental involvement among American and British parents depends on various factors such as the type of sporting activity, observance of other parents, previous experience and knowledge, set of expectations, and social values. McMahon and Penney (2015) find that parents of Australian amateur swimmers tend to discipline their child when they do not meet their ideals. Côté (1999) and Wall et al. (2019) examine parental involvement in Canadian youth sport. While Côté (1999) emphasizes the importance of the family in the context of the development of young athletes, Wall et al. (2019) highlights the complexity of the relationship between the coaches in figure-skating and the athletes' mothers. However, very little has been written in reference to parental involvement in their child's sporting activity in Israel. The studies that do exist mainly focus on soccer, the most popular sport in Israel. For example, Gershgoren et al. (2011) show how parental feedback influences the child's motivation and performance. In their study, they find that feedback, which focuses on abstract, general compliments intending to increase the child's sense of self-worth results in the child's perception of their performance to increase, while their actual skills on the field remain level. On the other hand, parental feedback based on specific, concrete actions by the child on the field during the course of the game does not increase the young player's ego and improves performance over time. In a different study, Gindi et al. (2016) examine the way youth soccer coaches cope with parental involvement. They show that coaches do acknowledge the positive aspects of parental involvement, mainly logistical (e.g., driving their kids to practices), however, the researchers indicate that this involvement results in increased parental attempts of involvement in the coaches' professional decisions. As a result, coaches find themselves frustrated and, faced with growing pressure, increasingly compliant with the parents' demands. Muchtar et al. (2013) emphasize the importance of familial support on the development of young professional tennis players in Israel. Their findings align with multiple studies, which examine this issue (see: Gould et al., 2006; Knight et al., 2010). Broadly speaking, in reference to the Israeli culture, Tesler et al. (2019) claims that the level of physical activity, not necessarily in team sports, among Israeli youth is influenced significantly by family involvement. According to the researchers, parental involvement that includes an active sporting engagement increases the child's level of physical activity and results in a general embrace of a healthy lifestyle. These findings can be relevant given that 28.8% of Israeli adolescents are considered overweight on the BMI (body mass index) scale. Therefore, providing parents with this data can help contribute to their success in implementing long term well-being for their child.

It is important to note that in the Israeli context, one predominant factor is mandatory army service, which plays a significant role in shaping citizens' mentality (Sasson-Levy and Lomsky-Feder, 2018) and influences leisure time physical activity and gender inequality (Lev and Hertzog, 2017).

Gender and Family Roles in Parental Involvement

As parental influence is clearly significant, whether in creating the optimal support structure or in role modeling, it is important to reflect on the differences between parents. This distinction is made in the Palomo-Nieto et al. (2011) study, which identifies the different roles a mother and a father play in their child's sport activity. The mother occupies a much more encouraging role and gives unconditional support; this support takes the form of emotional assistance, regardless of her child's actual sport performances, and is a result of the strong mother–child bond, which exists from early childhood. The father, on the other hand, may have a negative emotional impact on the child as he often dominates the role of second coach, which may lead to unnecessary pressure on the child. It should be noted that these assumptions must consider the child's age, as the mother's sociological and emotional influence is predominant mostly in childhood (Chan et al., 2012). Although the father's emotional influence may lead to burnout and frustration by the child, their level of sport activity is related to positive support (Eriksson et al., 2008). In his study, Eriksson provides evidence to the notion that athletic competence is related to the association a child has of sports based on parental athletic activity and that the father's participation in sport is of particular importance. In other words, a father who is a sporting role model can contribute to his child's self-determined motivation for sport. However, current research indicates that both mothers' and fathers' physical activity are influential for a child's sport participation, and that the influence of parental modeling appears to be stronger in parent–child pairs of the same gender (Cleland et al., 2005; Rodrigues et al., 2018). The family context is further developed in a study investigating young Brazilian female basketball players (Folle et al., 2018), which finds that families for whom practicing basketball plays an important role in their routine, serve to motivate the child to persevere in the basketball club. This reinforces the correlation between autonomy support and child motivation, and asserts that active, involved parents contribute to the construction of an environment that is favorable to the positive development of athletes. This involvement or participation in their child's sport routine can be manifested either by presence at matches, emotional encouragement, financial support, or even moderate pressure during training sessions and matches (ibid).

In light of the literature reviewed and given the paucity of research on Israeli parents' involvement, the aim of this study is to focus on the psychosociological aspects of parental involvement in their child's basketball performance.

METHODOLOGY

A total of 173 couples of parent and child answered the questionnaires. Out of the 173 parents, 51.4% are males and 48.6% are females, of an average age of 45.9 (SD = 6.4) years. Out of the 173 children, 78% are boys and 22% girls of an average age 13.7 (SD = 1.7) years, all of whom are active basketball players practicing an average of 5–10 h per week. The participants hail from various regions throughout Israel ranging from small villages to large cities.

Questionnaire: In this study, the researchers used two questionnaires—"parent questionnaire" and "child questionnaire" (Lev et al., 2021; Supplementary Material). Each questionnaire consisted of 22 valid and reliable questions regarding parents' involvement—logistical and discourse, parental expectations and contentment (parent questionnaire) and child contentment, expectation, motivation, and parental discourse (child questionnaire). Out of each questionnaire, the researchers chose the relevant items to examine for this study. Each item stood independently and was found to be reliable and valid (Lev et al., 2021). Each question was scored on a five-point Likert-type scale ranging from strongly disagree (1) to strongly agree (5).

The researchers are interested, however, in a more encompassing measure that demonstrates three main themes: parental logistical involvement (items P5,6,17–19 from parents' questionnaire); parental emotional involvement (items P7,8,11 from parents' questionnaire); and finally, the child's satisfaction/contentment (items C4–6,21,22 from child's questionnaire).

Statistical analysis was performed using SPSS (v.24). Data were expressed as means ± standard deviation (SD). Student t-test was used to compare between the results of mothers and fathers. Cronbach's α was used to determine internal consistency for the subscale scores "logistic involvement" and "emotional involvement" for parental questions and for "child's satisfaction." Mean scores were computed for each of those subscales. Pearson's correlation analysis was used to determine the relationship between parents' and child's answers. Multiple regression was used to identify which kind of involvement is the best for describing the variation in child's satisfaction. A value of $p < 0.05$ was considered statistically significant.

RESULTS

Descriptive Statistics

The average and standard deviation of the parents' answers are seen in **Table 1** and those of the children in **Table 2**. Israeli parents show high involvement (average score > 3; **Table 1**) with their child's sports activity when it comes to shuttling to practices and games, attending games, and spending time talking with their child about their participation in practices and games,

TABLE 1 | Descriptive statistics for parental (P) answers and the significance of comparison between mothers and fathers (*p* value).

Question/item	All x̄ ± *SD*	Male x̄ ± *SD*	Female x̄ ± *SD*	*p* value Male/female
P5. How involved are you with your child's coaching staff?	2.99 ± 1.29	2.82 ± 1.21	3.18 ± 1.35	0.068
P6. How involved are you with shuttles of your child to practice and games?	3.87 ± 1.22	3.79 ± 1.20	3.95 ± 1.24	ns
P7. To what extent is there a dialog between you and your child regarding practice and games?	3.97 ± 1.10	3.88 ± 1.12	4.07 ± 1.08	ns
P8. To what extent is there a dialog between you and your child regarding his performance as a basketball player?	3.79 ± 1.19	3.70 ± 1.24	3.89 ± 1.13	ns
P11. To what extent do you talk to your child about feelings (joy, stress, anxiety. . .) that your child is experiencing on the basketball court?	3.78 ± 1.07	3.53 ± 1.12	4.04 ± 0.94	**0.002**
P15. To what extent do you have a dialog with the referees or the scoring table during official games?	1.77 ± 1.15	1.97 ± 1.21	1.57 ± 1.04	**0.024**
P16. To what extent is there communication between you and your child during games?	2.25 ± 1.23	2.49 ± 1.26	1.99 ± 1.14	**0.07**
P17. How often do you film team games/practices?	2.66 ± 1.38	2.54 ± 1.31	2.79 ± 1.45	ns
P18. How often do you come to games during the year?	3.84 ± 1.18	4.02 ± 1.01	3.64 ± 1.32	**0.035**
P19. How often do you come to practices during the year?	2.12 ± 1.20	2.37 ± 1.24	1.86 ± 1.11	**0.005**
Parental logistic involvement	3.10 ± 0.84	3.11 ± 0.78	3.08 ± 0.91	ns
Parental emotional involvement	3.84 ± 0.93	3.70 ± 0.94	3.99 ± 0.93	**0.043**

ns, non-significant; bold, p < 0.05.

TABLE 2 | Descriptive statistics for young basketball players' (C) answers.

Question/item	*n*	Average	*SD*
C4. To what extent are you satisfied with your social integration in the team?	173	4.45	0.78
C5. How satisfied are you with your participation in practice?	173	4.39	0.79
C6. How satisfied are you with your participation in games?	172	4.06	1.07
C7. To what extent do you invest in improving your athletic abilities beyond the team practice framework?	172	4.02	0.94
C8. What is your parents' involvement in practice and games shuttles?	173	3.90	1.19
C9. How satisfied are you with your parents' involvements in practice and games shuttles?	173	4.32	1.04
C11. To what extent do you estimate that your parents' involvement has an impact on your athletic performance?	171	3.63	1.12
C12. To what extent would you like your parents to come to your games?	172	4.36	0.92
C13. To what extent would you like your parents to come to your practices?	172	2.77	1.35
C18. To what extent do your parents come to games during the regular season?	173	3.98	1.11
C19. To what extent do your parents come to practices during the regular season?	173	2.02	1.24
C20. To what extent do you think your performances influence your parents' mood?	173	3.22	1.32
C21. How much do you enjoy playing basketball?	173	4.74	0.58
C22. How satisfied are you with your sporting performances?	173	4.10	0.82
Child overall satisfaction	173	4.35	0.61

performance, and feelings (items P6,7,8,11,18). On the other hand, they show moderate involvement (1 < average score < 3; **Table 1**) with the coaching staff, game referees, attending practices, and filming their child during practices/games (items P5,15,16,17,19). Both mothers and fathers show high logistical and emotional involvement with their child's sports activity, but there are some substantial differences between the two. While fathers show higher involvement with the logistical side such as attending practices or games, and talking to the referees or their child during the game (items P15–17,19), mothers show higher involvement with the emotional side, especially talking with their child about their emotions (item P11).

Overall, the children in this study report very high satisfaction and enjoyment (average score > 4) regarding participation in their sports activity, their social integration, and their sporting performances (**Table 2**; items C4–6,21,22), and they invest in improving their athletic abilities beyond the team practice framework (item C7). Children also express that they want their parents to show high involvement in their sports specifically when it comes to shuttling them to practices/games and to attending games (**Table 2**; items C9,12,13). Additionally, children estimate that their parents' involvement has an impact on their athletic performance (item C11), and they believe that their sports performances influence their parents' mood (item C20).

Correlations

The researchers performed Pearson correlations coefficient to explore interaction between the parents' and the children's answers (**Table 3**). In general, higher involvement of the parents in their child's basketball activity (items P5–19) positively

TABLE 3 | Matrix of correlations between parental answers (P) and young basketball players' answers (C).

Question	P5	P6	P7	P8	P11	P15	P16	P17	P18	P19
C4	0.181**	0.191**	0.306**	0.316**	0.344**	0.159*	0.070	0.143*	0.273**	0.047
C5	0.256**	0.210**	0.348**	0.415**	0.323**	0.079	0.095	0.243**	0.277**	-0.032
C6	0.215**	0.024	0.278**	0.310**	0.318**	0.181**	0.139*	0.160*	0.214**	0.085
C7	0.120	0.069	0.156*	0.227**	0.156*	0.196**	0.071	0.039	0.139*	0.164*
C8	0.204**	0.536**	0.220**	0.237**	0.224**	-0.045	0.028	0.141*	0.290**	0.124
C9	0.165*	0.353**	0.325**	0.393**	0.304**	-0.100	-0.050	0.115	0.164*	0.021
C11	0.230**	0.157*	0.272**	0.350**	0.245**	0.233**	0.170*	0.195**	0.199**	0.060
C12	0.223**	0.417**	0.313**	0.332**	0.308**	0.224**	0.219**	0.306**	0.449**	0.098
C13	0.159*	0.078	0.058	0.063	0.029	0.260**	0.197**	0.162*	0.144*	0.358**
C18	0.332**	0.419**	0.261**	0.352**	0.287**	0.160*	0.200**	0.306**	0.651**	0.187**
C19	0.269**	0.029	0.039	0.102	0.021	0.278**	0.187**	0.182**	0.209**	0.650**
C20	0.147*	0.015	0.082	0.161*	0.057	0.356**	0.286**	0.184**	0.164*	0.112
C21	0.138*	0.240**	0.358**	0.482**	0.295**	-0.007	0.023	0.075	0.219**	0.013
C22	0.138*	0.176*	0.191**	0.190**	0.152*	0.300**	0.216**	0.225**	0.232**	0.123

*$p < 0.05$, **$p < 0.01$.

correlated with the child's answers regarding their satisfaction (items C4,5,6,9,21,22). More specifically, the researchers found weak but significant correlations between parents' involvement with the coaching team (item P5) and the child's satisfaction ($0.138 < r < 0.256$); weak but significant correlations between parents' involvement with shuttling the child to practices and games (item P6) and the child's satisfaction ($0.176 < r < 0.353$) in all but two questions (items C6,21); weak to moderate correlations ($0.152 < r < 0.415$) between parental dialog with their child (items P7,8,11) and their child's satisfaction, with the exception of C21; weak but significant correlations between parents' involvement with game referees (item P15) and the child's satisfaction ($0.159 < r < 0.300$) in all but two questions (items C5,9); weak significant correlations between parents' interaction with the child during the game (item P16) and the child's satisfaction are found only in two questions (items C6,22); weak but significant correlations between parents' involvement with filming and photographing the game (item P17) and the child's satisfaction ($0.143 < r < 0.243$) in all but one of the child's questions (item C9); weak but significant correlations between parents' attending the child's games (item P18) and all of the measures of the child's satisfaction ($0.164 < r < 0.277$). There was no correlation between how often parents attended the child's practices (item P19) and the child's satisfaction.

The researchers divided parental involvement into "logistical involvement," which was the average score for parental items 5, 6, 17–19, and to "emotional involvement," which was the average score for parental items 7, 8, and 11 (**Table 1**). They found high internal consistency between the questions that constituted each of the parental involvement themes (Cronbach's $\alpha = 0.688$). They also created a new theme regarding a child's overall satisfaction, which was the average score for items 4–6, 21, 22. They found high internal consistency between the five questions that constituted the child's satisfaction theme (Cronbach's $\alpha = 0.788$). Correlation analysis revealed significant, moderate correlation between parental logistical and emotional involvement ($r = 0.540$), indicating that parents who spent more time and effort in logistics relating to their child's sports activity also showed more emotional involvement. The researchers also found significant positive correlations between a child's satisfaction and parental logistical and emotional involvement ($r = 0.321$, $r = 0.479$, respectively).

Mediation Effect

Multiple linear regression analysis was used to develop a model for predicting the child's satisfaction from the parents' logistical involvement and the parents' emotional involvement. The two-predictor model was able to account for 48.4% of the variance in the child's satisfaction, $F(2,170) = 26.053$, $p < 0.001$. As mentioned previously, each of the predictor variables had a significant ($p < 0.001$) zero-order correlation with the child's satisfaction, but only the parental emotional involvement had significant ($p < 0.001$) partial effects in the full model. Therefore, parental emotional involvement was the most important variable for a child's satisfaction. The zero-order correlations between parental logistical involvement and a child's satisfaction showed that there was a statistically significant, moderate, positive correlation ($r = 0.321$, $p < 0.001$), but while controlling for parental emotional involvement, the correlation was not statistically significant ($r = 0.084$, $p > 0.05$).

DISCUSSION

The aim of this paper is to shed light on parents' level of involvement with their child's sporting activity in Israel. It is generally agreed that parental involvement has vast implications on a child's motivation and enjoyment, however, as many researchers have suggested (Cheung and Pomerantz, 2011; Knight et al., 2017; Harwood et al., 2019; Holt et al., 2020) this phenomenon has to be comprehended in the sociocultural context. Thus, this paper makes a unique contribution by exploring the nature of parental involvement on youth basketball in Israel when it comes to parenting style in the context of dilemmas and ethical issues.

The findings reveal that Israeli parents invest great time in their child's sporting activity; more specifically, parents invest time in shuttling their child to practices and games and in talking with their child about their physical and emotional well-being and the child's performances at practices and games. These findings are in line with the recent Furusa et al. (2020) study from Britain, which demonstrates the child's desire for their parents to not only provide tangible support but also to be actively engaged in the discussions about their sporting performances. However, the researchers find that while regularly attending games, Israeli parents are only semi-attending practices, documenting their child's performances on camera and talking with professional staff. This finding contradicts the Gindi et al. (2016) study, which examines the way youth soccer coaches cope with parental involvement. While the researchers demonstrate a high level of logistical engagement among the Israeli soccer parents in their study, leading to more parental interference with coaches during practices, in the current study, the Israeli basketball parents are less engaged in the child's practices altogether. This situation may present an ethical dilemma to Israeli parents who question whether to increase their involvement or allow their child greater autonomy (Harwood et al., 2019). There are various autonomous domains, which may be allowed for adolescents of a certain age, such as logistical independence (driving oneself to practices/games), boundaries during games (allowing coaches and referees to make decisions without interference, and not directing their child during the course of the game), and encouraging the child to assert their independence (lack of need of "helicopter parenting"). The findings of the current study show that the level of parental, logistical involvement in Israel is high (P6 > 3) contrasted with their involvement with coaches and during games (P5,15,16 < 3). The children in this study display high levels of satisfaction and enjoyment from the sport; therefore, the researchers conclude that the domain in which parents are best to provide greater autonomy to their child lies in the area of non-interference with sporting staff. Thus, as opposed to the parents of youth soccer players (Gindi et al., 2016), parents of adolescent basketball players should consider not becoming overly intrusive with their child's coaches. This aligns with Gould et al. (2008) and Lauer et al. (2010) who stress that when American parents interfere with coaches, it can result in pressure and anxiety for the child. With regard to the differences between the parents of basketball players versus those of soccer players, the researchers suggest the possibility that due to soccer's popularity over basketball in Israel, soccer parents are more inclined to attend and be actively involved in both practices and games.

The current study finds a higher level of satisfaction and contentment among youth basketball players whose parents demonstrate higher levels of involvement, and vice versa. Put simply, the more parents are involved in their child's sporting activity, the higher the child's contentment and satisfaction with the sport. These findings support the Sánchez-Miguel et al. (2013) study, which demonstrates a positive relationship between parental support of the sport and a player's enjoyment and motivation among youth Spanish athletes. However, the findings of this study seem to contradict scholars who suggest that parental support acts as a stressor that diminishes the athlete's physical and emotional well-being (Elliott and Drummond, 2013; Knight and Holt, 2014) among Australian and British athletes. This can be partly explained given that Israeli culture values tight-knit families. Although Israeli families are undergoing changes as part of the postmodern trend of adolescents seeking greater autonomy at a younger age, they still have relative solidarity and emotional closeness (Peer, 2014; Ben-Nun and Tene, 2018).

Further, in the current study, the researchers find a strong correlation between emotional parental involvement and their child's satisfaction, while logistical parental involvement contributes to a lesser extent. This is not to say that the latter is not important or is negligible. After all, there is vast agreement among scholars that parental logistical support demonstrates to the child that their parents value their sporting activity and thus enhances feelings of enjoyment, competence, motivation, and persistence (Knight et al., 2010; Harwood et al., 2019). However, this study emphasizes that logistical support cannot stand alone; it must include emotional support. Although shuttling their child to practices and games can be perceived as strictly technical, it should be stressed that parents do not perform as "taxi drivers" providing impartial engagement. On the contrary, driving their child constitutes active engagement, which can facilitate profound dialog between parents and their child before and after practices/games. Elliott and Drummond (2017) find that verbal comments, for example, those articulated in the car in the after-sport experience are interpreted as negative and thus weaken the enjoyment of sporting experiences for the child. As demonstrated in this particular research, the findings are not in line with those of Elliott and Drummond. The results of the study indicate that verbal comments by the parents in transit during the after-sport experience have the potential to increase a young athlete's satisfaction in the game. The implications are noteworthy. First, this study highlights the significance of being involved in a child's sporting activity not only emotionally but also logistically. Hence, Israeli parents who send their child to practices/games via taxi or another form of public transportation do not impart the emotional component, thus missing the opportunity for emotional engagement with their child. Parents may be torn with the question of the appropriate level of involvement and whether what they are doing is enough for their child's satisfaction and well-being. In modern times, busy parents often face a dilemma when they feel they must choose either logistical or emotional involvement with their child based on limits on their time due to various other responsibilities such as their professional careers. The researchers highly recommend that parents not bypass logistical support opportunities as these lead to an enhancement of a child's contentment and satisfaction in their sporting activity. Thus, the results indicate that the role of the parents in not cutting corners, in other words, providing both emotional and logistical support contributes to the notion of sport-based PYD. Moreover, the analysis of this study also shows a gender variation with regard to the nature of parental involvement. While both parents play a significant role in their child's sporting activity, the findings reveal that fathers are more involved with logistical aspects (such as attending games and practices and speaking with the coach), whereas mothers are more involved with providing emotional support. In this respect, mothers constitute a "sounding board," which helps a child express their emotions. This finding is in line with the Palomo-Nieto et al. (2011) study, which stresses that Spanish mothers offer emotional assistance, assuming a much more encouraging role and giving unconditional support. Nevertheless, the current study's findings contradict the (Palomo-Nieto et al., 2011) conclusion that fathers have a negative emotional impact on the child as they often dominate the role of "second coach" leading to unnecessary pressure on the child. As mentioned above, in this study, a child perceives emotional and logistical involvement from both parents as supportive, which increases their contentment, satisfaction, enjoyment, and motivation with the sport. In order to understand the differing gender roles, the researchers believe it is important to emphasize the discourse of militarization in Israeli society, due to the gender separation of certain units within the Israeli army, which implicitly reinforces domestic roles (Sasson-Levy and Lomsky-Feder, 2018). The clear gender divide in Israel among parents is not surprising due to mandatory army service resulting in gender inequality, which pervades everyday life (ibid). This reality can also take place in Israeli sport leisure activities (see: Lev and Hertzog, 2017; Hertzog and Lev, 2019).

This study examines numerous youth basketball teams in Israel, however, it concentrates mainly on the adolescent, male, and Jewish population. In some regard, their experiences may differ from adolescents of other genders, cultures, and religions. Therefore, the study is limited due to its focus on only one specific ethnic group and only a specific sport. A further study could include research that follows parents' level of involvement with their child's sporting activity among both genders within different sports and different ethnic groups. It is important to stress that this study emphasizes parental involvement at a specific time and, therefore, does not provide any information regarding the previous parent–child relationship. Moreover, while this study is based on a quantitative approach, a further study on parents' level of involvement with their child's sporting activity in Israel using a qualitative approach, such as semi-structured interviews and observations, is suggested. This has the potential to encourage participants to provide their own unique context for events, feelings, and behaviors (see Lev, 2019, 2020).

Further, given that Israel recently placed second among European countries in adolescent BMI (body mass index) rankings with 28.8% of youth overweight, this study can contribute to encouraging youth to continue with sporting activity by helping parents understand the importance of their interactions with their child regarding sports. The paper constitutes an effective tool to aid Israeli parents in comprehending the value of their approach and how this approach might lead to a child's increased satisfaction and enjoyment with sporting activity, therefore increasing adolescent health. Finally, as little has been written in reference to the parent/child relationship in youth basketball in Israel, this paper addresses the lacuna of literature regarding parental involvement in sports. Harwood et al. (2019) remind us that there is a certain homogeneity in the literature in reference to the "best" parenting style within youth sport. In this regard, this study also contributes to the scope of parental involvement in sport by shedding light on Israeli youth basketball players.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ono Academic College, Ethics Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

AL conceived the presented idea, assembled the questionnaires, wrote the manuscript, analyzed the results, and supervised the project. AB assembled the questionnaires and collected the data. AM, SB, and NF assembled the questionnaires. EB performed the analysis, assembled questionnaires, wrote the manuscript, and supervised the project. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We wish to thank the reviewers for their helpful comments on an earlier draft. We also wish to thank Emily Fishman and Shay Leibovich for their support and to all the parents and children who participated in the study. The research was supported by Ono Academic College grant for preliminary studies.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.607000/full#supplementary-material.

REFERENCES

Ben-Nun, E., and Tene, Z. (2018). Mutual parenting motherhood in a hetero-gay family: can we enjoy both worlds? *Psichoactualia* 1, 28–36.

Chan, D. K., Lonsdale, C., and Fung, H. H. (2012). Influences of coaches, parents, and peers on the motivational patterns of child and adolescent athletes. *Scand. J. Med. Sci. Sports* 22, 558–568. doi: 10.1111/j.1600-0838.2010.01277.x

Cheung, C. S. S., and Pomerantz, E. M. (2011). Parents' involvement in children's learning in the United States and China: implications for children's academic and emotional adjustment. *Child Dev.* 82, 932–950. doi: 10.1111/j.1467-8624.2011.01582.x

Cleland, V., Venn, A., Fryer, J., Dwyer, T., and Blizzard, L. (2005). Parental exercise is associated with Australian children's extracurricular sports participation and cardiorespiratory fitness: a cross-sectional study. *Int. J. Behav. Nutr. Phys. Act* 2:3.

Côté, J. (1999). The influence of the family in the development of talent in sport. *Sport Psychol.* 13, 395–417. doi: 10.1123/tsp.13.4.395

Dorsch, T. E., Smith, A. L., and McDonough, M. H. (2015). Early socialization of parents through organized youth sport. *Sport Exerc. Perform. Psychol.* 4, 3–18. doi: 10.1037/spy0000021

Elliott, S. K., and Drummond, M. J. (2013). A socio-cultural exploration of parental involvement in junior Australian Rules football. *Asia Pac. J. Health Sport Phys. Educ.* 4, 35–48. doi: 10.1080/18377122.2013.760426

Elliott, S. K., and Drummond, M. J. (2017). Parents in youth sport: what happens after the game?. *Sport Educ. Soc.* 22, 391–406. doi: 10.1080/13573322.2015.1036233

Eriksson, M., Nordqvist, T., and Rasmussen, F. (2008). Associations between parents' and 12-year-old children's sport and vigorous activity: the role of self-esteem and athletic competence. *J. Phys. Act. Health* 5, 359–373. doi: 10.1123/jpah.5.3.359

Folle, A., Nascimento, J. V. D., Salles, W. D. N., Maciel, L. F. P., and Dallegrave, E. J. (2018). Family involvement in the process of women's basketball sports development. *J. Phys. Educ.* 29:2914. doi: 10.4025/jphyseduc.v29i1.2914

Fraser-Thomas, J., and Côté, J. (2009). Understanding adolescents' positive and negative developmental experiences in sport. *Sport Psychol.* 23, 3–23. doi: 10.1123/tsp.23.1.3

Furusa, M. G., Knight, C. J., and Hill, D. M. (2020). Parental involvement and children's enjoyment in sport. *Qual. Res. Sport Exerc. Health* [Epub ahead of print]. doi: 10.1080/2159676X.2020.1803393

Gershgoren, L., Tenenbaum, G., Gershgoren, A., and Eklund, R. C. (2011). The effect of parental feedback on young athletes' perceived motivational climate, goal involvement, goal orientation, and performance. *Psychol. Sport Exerc.* 12, 481–489. doi: 10.1016/j.psychsport.2011.05.003

Gindi, S., Feldman, A., Honig, O., Rinat, S., and Feldman, A. (2016). Soccer coaches' handling of parental involvement. *Mov. J. Phys. Educ. Sport Sci.* 11, 209–227.

Gould, D., Lauer, L., Rolo, C., Jannes, C., and Pennisi, N. (2006). Understanding the role parents play in tennis success: a national survey of junior tennis coaches. *Br. J. Sports Med.* 40, 632–636.

Gould, D., Lauer, L., Rolo, C., Jannes, C., and Pennisi, N. (2008). The role of parents in tennis success: focus group interviews with junior coaches. *Sport Psychol.* 22, 18–37. doi: 10.1123/tsp.22.1.18

Harwood, C. G., Knight, C. J., Thrower, S. N., and Berrow, S. R. (2019). Advancing the study of parental involvement to optimise the psychosocial development and experiences of young athletes. *Psychol. Sport Exerc.* 42, 66–73. doi: 10.1016/j.psychsport.2019.01.007

Hertzog, E., and Lev, A. (2019). Male dominance under threat: machoism confronts female defiance in israeli gyms. *J. Contemp. Ethnogr.* 48, 836–866. doi: 10.1177/0891241619834662

Holt, N. L., Deal, C. J., and Pankow, K. (2020). "Positive youth development through sport," in *Handbook of Sport Psychology*, eds G. Tenenbaum, and R. C. Eklund (Hoboken, NJ: John Wiley & Sons), 429–446.

Holt, N. L., and Sehn, Z. L. (2008). "Processes associated with positive youth development and participation in competitive youth sport" in *Positive Youth Development Through Sport*, ed. N. L. Holt (New York, NY: Routledge), 38–47.

Hornby, G., and Lafaele, R. (2011). Barriers to parental involvement in education: an explanatory model. *Educ. Res. Rev.* 63, 37–52. doi: 10.1080/00131911.2010.488049

Kang S., Jeon, H., Kwon, S., and Park, S. (2015). Parental attachment as a mediator between parental social support and self-esteem as perceived by Korean sports middle and high school athletes. *Percept. Mot. Skills* 120, 288–303. doi: 10.2466/10.pms.120v11x6

Knight, C. J. (2019). Revealing findings in youth sport parenting research. *Kinesiol. Rev.* 8, 252–259. doi: 10.1123/kr.2019-0023

Knight, C. J., Berrow, S. R., and Harwood, C. G. (2017). Parenting in sport. *Curr. Opin. Psychol.* 16, 93–97. doi: 10.1016/j.copsyc.2017.03.011

Knight, C. J., Boden, C. M., and Holt, N. L. (2010). Junior tennis players' preferences for parental behaviors. *J. Appl. Sport Psychol.* 22, 377–391. doi: 10.1080/10413200.2010.495324

Knight, C. J., Dorsch, T. E., Osai, K. V., Haderlie, K. L., and Sellars, P. A. (2016). Influences on parental involvement in youth sport. *Sport Exerc. Perform. Psychol.* 5, 161–178. doi: 10.1037/spy0000053

Knight, C. J., and Holt, N. L. (2013). Factors that influence parents' experiences at junior tennis tournaments and suggestions for improvement. *Sport Exerc. Perform. Psychol.* 2, 173–189. doi: 10.1037/a0031203

Knight, C. J., and Holt, N. L. (2014). Parenting in youth tennis: understanding and enhancing children's experiences. *Psychol. Sport Exerc.* 15, 155–164. doi: 10.1016/j.psychsport.2013.10.010

Lauer, L., Gould, D., Roman, N., and Pierce, M. (2010). Parental behaviors that affect junior tennis player development. *Psychol. Sport Exerc.* 11, 487–496. doi: 10.1016/j.psychsport.2010.06.008

Lev, A. (2019). Becoming a long-distance runner – deriving pleasure and contentment in times of pain and bodily distress, Leis. *Studies* 38, 790–803. doi: 10.1080/02614367.2019.1640776

Lev, A. (2020). Distance runners in a dys-appearance state – Reconceptualizing the perception of pain and suffering in times of bodily distress. *Qual. Res. Sport Exerc. Health* [Epub ahead of print]. doi: 10.1080/2159676X.2020.1734647

Lev, A., Bichman, A., Moyal, A., Brenner, S., Fass, N., and Been, E. (2021). Research tools for evaluation of parental involvement in team physical activity and its influence on child perception. *Spirit Sport* 7.

Lev, A., and Hertzog, E. (2017). Typology of gender resistance and defiance in Israeli gyms. Sport Soc. 20, 1699–1714. doi: 10.1080/17430437.2017.1310203

McMahon, J. A., and Penney, D. (2015). Sporting parents on the pool deck: living out a sporting culture?. *Qual. Res. Sport Exerc.* 7, 153–169. doi: 10.1080/2159676X.2014.901985

Muchtar, O., Zilber, Z., and Lazar-Shoef (2013). The influence of the family in the development of Tennis athletes. *Soc. Issues Israel* 15, 28–57.

Nunomura, M., and Oliveira, M. S. (2013). Parents' support in the sports career of young gymnasts. *Sci. Gymnast. J.* 5, 5–18.

Palomo-Nieto, M., Ruiz-Perez, L. M., Sanchez-Sanchez, M., and García-Coll, V. (2011). Young athletes' perceptions of the role of their mothers in their sport career. *Motricidad. Eur. J. Hum. Mov.* 26, 141–157. doi: 10.4324/9781315545202-13

Peer, Y. (2014). *The Attitudes of Heterosexuals Toward Different Routes of Gay Parenting*. MA thesis, Ben-Gurion University, Israel.

Rodrigues, D., Padez, C., and Machado-Rodrigues, A. M. (2018). Active parents, active children: the importance of parental organized physical activity in children's extracurricular sport participation. *J. Child Health Care* 22, 159–170. doi: 10.1177/1367493517741686

Sánchez-Miguel, P. A., Leo, F. M., Sánchez-Oliva, D., Amado, D., and García-Calvo, T. (2013). The importance of parents' behavior in their children's enjoyment and amotivation in sports. *J. Hum. Kinet.* 36, 169–177. doi: 10.2478/hukin-2013-0017

Sasson-Levy, O., and Lomsky-Feder, E. (2018). *Gender At The Base: Women, Men and Military Service in Israel*. Tel Aviv: Van Leer Institute Press.

Sorkkila, M., Ryba, T. V., Aunola, K., Selänne, H., and Salmela-Aro, K. (2017). Sport burnout inventory-dual career form for student-athletes: assessing validity and reliability in a Finnish sample of adolescent athletes. *J. Sport Health Sci.* 1–9. doi: 10.1016/j.jshs.2017.10.006

Stefansen, K., Smette, I., and Strandbu, Å. (2018). Understanding the increase in parents' involvement in organized youth sports. *Sport Educ. Soc.* 23, 162–172. doi: 10.1080/13573322.2016.1150834

Tate, E. B., Shah, A., Jones, M., Pentz, M. A., Liao, Y., and Dunton, G. (2015). Toward a better understanding of the link between parent and child physical activity levels: the moderating role of parental encouragement. *J. Phys. Act. Health* 12, 1238–1244. doi: 10.1123/jpah.2014-0126

Tesler, R., Kolobov, T., Ng, K. W., Shapiro, E., Walsh, S. D., Shuval, K., et al. (2019). Ethnic disparities in physical activity among adolescents in Israel. *Am. J. Health Behav.* 43, 337–348. doi: 10.5993/AJHB.43.2.10

Wall, J. M., Baugh, L. M., Pradhan, K., Beauchamp, M. R., Marshall, S. K., and Young, R. A. (2019). The coach-parent relationship in Canadian competitive figure skating: an interpretive description. *Psychol. Sport Exerc.* 45:101577. doi: 10.1016/j.psychsport.2019.101577

Welk, G. J., Wood, K., and Morss, G. (2003). Parental influences on physical activity in children: an exploration of potential mechanisms. *Pediat. Exerc. Sci.* 15, 19–33. doi: 10.1123/pes.15.1.19

Wheeler, S., and Green, K. (2014). Parenting in relation to children's sports participation: generational changes and potential implications. *Leis. Stud.* 33, 267–284. doi: 10.1080/02614367.2012.707227

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership