©Journal of Sports Science and Medicine (2020) **19,** 149-157
http://www.jssm.org

Review article

Physical Load and Referees' Decision-Making in Sports Games: A Scoping Review

Nicolas Bloß ✉, Jörg Schorer, Florian Loffing and Dirk Büsch
Institute of Sport Science, Carl von Ossietzky University Oldenburg, Oldenburg, Germany

Abstract

Referees in sports games have a high level of responsibility as they have to make correct and appropriate decisions at any point during a match. Regarding referees' decision-making (RDM) as a perceptual-cognitive process, evidence suggests that physical load might reduce cognitive performance and thus might reduce RDM performance as well. In consideration of increasing game dynamics, referees have to cope with high physical load, but they have to make correct and appropriate decisions further on. Here, we review the current state of research on the relationship between physical load and RDM. A scoping review was performed, in accordance with the PRISMA guidelines, using the following databases: Scopus, PubMed, Web of Science (all databases), SURF and SPONET (both with an English and German search). Only primary studies written in English or German that investigated the relationship between physical load and RDM in sports games were included. Eleven studies included in the review investigated six physical parameters and RDM: match period, velocity, blood lactate, running time, heart rate, distance covered. Most findings of the studies showed no relationship between physical load and RDM (n = 18). Thirteen findings suggest a negative relationship and three findings indicate a positive relationship between physical load and RDM. Results of the scoping review show contradictory evidence across and within investigated different physical parameters. As RDM consists of multi-factorial components, it is recommended to conduct systematic research programs – field as well as experimental studies – to resolve the missing control of potential confounding variables and to consider the difference of internal and external load.

Key words: Performance, accuracy, stress, exertion, umpire, officiating.

Introduction

Referees in sports games have the responsibility to ensure that teams and players always abide by the sport-specific rules and intervene when a team or a player infringed a rule (Reilly and Gregson, 2006; Rullang et al., 2017). Therefore, a central task in officiating sports games is to make decisions about technical, offensive and defensive infringements (Bar-Eli et al., 2011; Plessner et al., 2009). As referees' decisions can ultimately impact a game's outcome (Larkin et al., 2014; MacMahon et al., 2015), it is indispensable that referees' decision-making (RDM) is on a high level (Helsen et al., 2019; MacMahon et al., 2015). This includes that referees are able to apply the sport-specific rules during the entire match; that is, they need to have a well-developed physical fitness in order to keep up with the match dynamics when officiating (Kittel et al., 2019; MacMahon et al., 2015; Plessner and MacMahon, 2013; Reilly and Gregson, 2006). Mascarenhas et al. (2005) emphasize the relevance of both RDM and physical fitness for appropriate officiating when identifying five cornerstones of referees' performance: (1) personality and game management, (2) physical fitness, positioning and mechanics, (3) knowledge and application of the law, (4) contextual judgement and (5) psychological characteristics of excellence. However, albeit highlighting the important role of both RDM and physical fitness, the authors' model does not take into account potential relationships between cornerstones, RDM and physical fitness in particular whereby the central parameter of RDM is the correctness of the decisions made by the referees (Schweizer et al., 2011).

From this point of view, the relationship between RDM and physical load is of particular interest. Referees' decisions are a perceptual-cognitive process (Helsen et al., 2019; Plessner and Haar, 2006; Schmidt et al., 2019), which might be impaired through high physical stress (Chang et al., 2012; Schmidt et al., 2019; Tomporowski, 2003). High physical stress can, for example, lead to an excessive enrichment of nor-epinephrine, which can result in a reduction of neuronal activity in the prefrontal cortex and ultimately cause a reduction in attention and executive functions (Arnsten, 2009; Schmidt et al., 2019). Thus, physical load may impair referees' attentional control and, as a consequence, their decisions (Helsen et al., 2019; Schmidt et al., 2019).

Since decision-making is considered as an important task in officiating (Kittel et al., 2019), it is necessary to understand the relationship between physical load and RDM. Here we review the evidence on the relationship between physical load and RDM in sports games.

Methods

The scoping review was performed according to the Preferred Reporting Items for Systematic Reviews guideline (PRISMA; Moher et al., 2009).

Information sources and search strategy

The systematic search was carried out on September 4, 2019, by using the following databases: Web of Science (all databases), Scopus, PubMed, SURF and SPONET. Searches in SURF and SPONET were conducted in English as well as in German. The option 'advanced searches' was chosen for searches in Web of Science, PubMed and Scopus with regard to the articles' title, abstract and keywords. "All Field" searches were executed in SURF and SPONET. The development and selection of the final keywords used was a process consisting of several search

Received: 24 October 2019 / Accepted: 03 January 2019 / Published (online): 24 February 2020

attempts and corresponding adjustments, which were discussed with all authors to improve the quality of the keywords and the search. In the end, the following search terms are used: "(referee* OR umpir* OR officiating) AND ("decision-making" OR performance OR judg*) AND (stress OR strain OR effort OR load OR exertion OR pressure)". The general German search term was: "(Schiedsrichte*) AND (Entscheid* OR Leist* OR Bewert*) AND (Stress OR Belastung OR Anstrengung OR Beanspruchung OR Erschöpfung OR Druck)".

Eligibility criteria and study selection

To be included in the review, articles must strictly fulfil the following inclusion criteria. Articles have to be primary studies and investigate the relationship between physical load and RDM. They must also refer to sports games and be written in English or German. We set no limitation to the year of publication. Besides, articles were excluded when meeting defined exclusion criteria: articles are reviews, abstracts, project descriptions, conference papers, interviews, theoretical papers or dissertations. Studies referring to athletes, studies without reference to sport games or articles not written in English or German were also excluded.

In accordance with the PRISMA approach (Moher et al., 2009), a dditional articles were appended in the selection process after the database search was conducted. Afterwards, duplicates were excluded and titles, abstracts and the written languages were screened and excluded when not meeting inclusion criteria. To ensure that no relevant article was excluded, critical articles were kept up and were discussed by the authors. Finally, full-texts were analysed and discussed before exclusion. The remaining articles were integrated into the scoping review (see PRISMA flowchart in Figure 1).

Data extraction

For each study selected for inclusion in the review, the following parameters were extracted:

- Study characteristics: sport, sample, referee type, expertise (level of officiating, experience), sex, age and a brief report of the measurement
- Study results categorized by physical parameters and physical load (external or internal load), a short description of the results and its statistical key figures

As outcomes and reports of the studies were not uniform, data were processed for a scoping review. This type of review article aims at a first examination of the existing evidence on a research subject especially for elaborating future research (Tod, 2019).

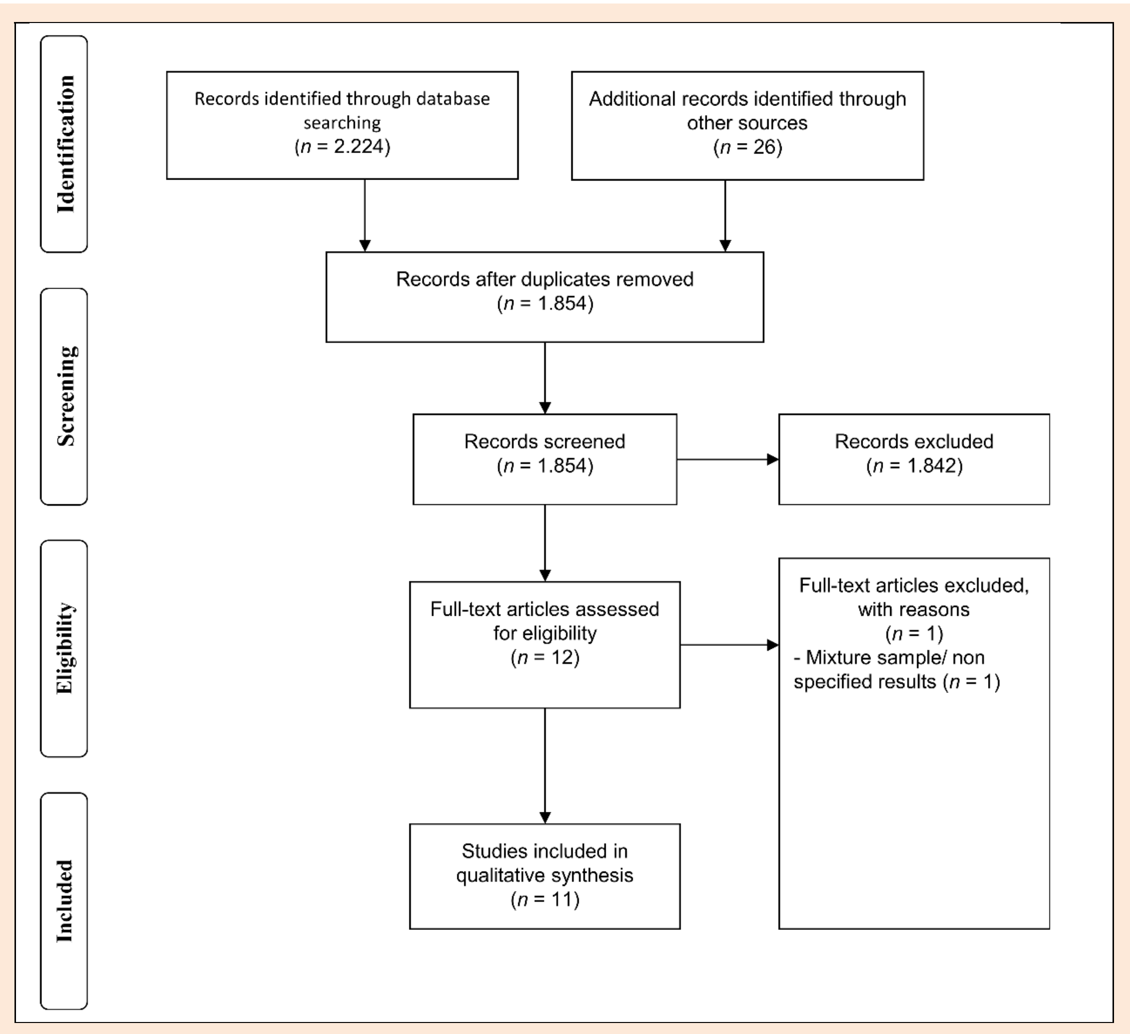

Figure 1. Flow diagram (Moher et al., 2009).

Risk of bias

All studies included in this review were assessed with regard to risk of bias by the authors. To this end, we used the Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018) to assess the methodological quality of included studies. Criteria and the assessment of the risk of bias of included studies are illustrated in Table 1.

With reference to the quality assessment with the MMAT, we follow the MMAT publishers' recommendation to not report a numeric score to better inform about the quality of the studies and to avoid misunderstandings by discussing meanings of the ratings (Hong et al., 2018).

Results

Study selection

A total of *n* = 2,250 potential articles were identified through the literature search. Of those, *n* = 2,224 articles were found in the database search and *n* = 26 additional articles were included in the selection process, as required by PRISMA. Additional articles were obtained by screening references from relevant articles. After excluding

Table 1. Characteristics of the included studies.

Study	Sport	Sample (n)	Referee type	Expertise *Level*	Expertise *Experience*	*Sex*	Age	Measurement
Ahmed et al. (2017)	Futsal	18	R	Experts	> 5y	Male	34.1±3.2y	- R recorded in nine competitive matches - R decisions assessed by three futsal experts
Catteeuw et al. (2010)	Soccer	10 38	AR AR	Experts	FIFA: 2.6 ± 3.3y EPL: 4.8 ± 5.1y 6.6 ± 5.7y	NR	NR	- Video assessment of offside situations with ProZone®
Elsworthy et al. (2014)	Australian football	29	Umpire	Experts	NR	NR	32.4±6.1y	- Movement demands measured with GPS - Decisions assessed by AFL coaching staff (n = 3) - Video analysis by the authors
Emmonds et al. (2015)	Rugby	8	R	Experts	7.3 ± 4.3y	NR	35.3±7.1y	- Movement demands measured with GPS - Video-assessment of penalty accuracy in 10-minutes periods
Gomez-Carmona and Pino Ortega (2016)	Soccer	5 1	AR R	Advanced Advanced	NR	NR	21.2±0.98y	- Recorded/ monitored in four games - R decisions assessed by five experts
Larkin et al. (2014)	Australian football	15	Umpire	Experts	235.2 ± 151.3 matches	NR	36±13.5y	- Blood lactate and video test after each quarter - GPS monitoring for movement demands
Mallo et al. (2012)	Soccer	10 20	R AR	Experts	FIFA license: 9.2 ± 4.1y 5.8 ± 3.8y	NR	39.4±2.4y 37.3±4.0y	- Matches recorded during FIFA confederations Cup 2009 in South Africa - Decisions assessed by top-class referees (n = 3; experience: 9.7 ± 1.5y)
Mascarenhas et al. (2009)	Soccer	5	R	Experts	> 4y	Male	38.2±5.89y	- Recorded/ monitored in seven games - Assessment by an authors
Oudejans et al. (2005)	Soccer	4	AR	Experts	> 5y	Male	42.2y	- Four matches recorded - Video analysis of offside situations
Paradis et al. (2015)	Australian football	10 8	Umpire Umpire	Sub-elite Junior	8.86 ± 2.04y NR	NR	2.86±3.39y 22.38±10.5y	- Physical exercise (10x300m runs) with a video test each round
Samuel et al. (2019)	Soccer	*22*	R	Experts	5-23y	Male	27.14±6.4	- Video-based decision test on a treadmill with variable speed

R = referee, AR = assistant referee, EPL: English Premier League, y = years, m = matches, NR = not reported, ± = standard deviation duplicates, $n = 1{,}854$ articles remained for further analyses. Subsequently, these articles were screened with reference to their written language, title and abstract. As a result, $n = 1{,}842$ articles were excluded, because they did not fulfil the inclusion criteria. The remaining $n = 12$ articles underwent a full-text analysis against the eligibility criteria. This led to further exclusion of $n = 1$ article. Finally, $n = 11$ articles were included in the qualitative analysis (see Figure 1).

Sports games and referee type

Most of the included studies were carried out in soccer ($n = 6$) followed by Australian football ($n = 3$), rugby ($n = 1$) and futsal ($n = 1$). The participants' referee type comprised umpires (Australian football), referees (soccer, futsal and rugby) and assistant referees (soccer). Participants officiated at least on an advanced level ($n = 24$), but most participants were experts ($n = 179$). Detailed study characteristics are listed in Table 1.

Measurement

Included studies differ in the selected approach of analysing the relationship between physical load and RDM. In eight studies, video-analysis of recorded matches were conducted and thus RDM was analysed on-field (Ahmed et al., 2017; Catteeuw et al., 2010; Elsworthy et al., 2014a; Emmonds et al., 2015; Gomez-Carmona and Pino-Ortega, 2016; Mallo et al., 2012; Mascarenhas et al., 2009; Oudejans et al., 2005). With the exception of the study by Catteeuw et al. (2010), selected video-sequences were assessed by experts (Ahmed et al., 2017; Elsworthy et al., 2014a; Emmonds et al., 2015; Gomez-Carmona and Pino-Ortega, 2016; Mallo et al., 2012; Mascarenhas et al., 2009; Oudejans et al., 2005).

Three studies used video-based decision tests conducted in the laboratory (Paradis et al., 2015; Samuel et al., 2019) or in the half times of real matches (Larkin et al., 2014). The video-sequences were assessed by experts too (Larkin et al., 2014; Paradis et al., 2015; Samuel et al., 2019).

Quality assessment

Four included studies were categorized as quantitative non-randomized studies (Larkin et al., 2014; Mascarenhas et al., 2009; Paradis et al., 2015; Samuel et al., 2019) and seven studies as quantitative descriptive studies (Ahmed et al., 2017; Catteeuw et al., 2010; Elsworthy et al., 2014a; Emmonds et al., 2015; Gomez-Carmona and Pino-Ortega, 2016; Mallo et al., 2012; Oudejans et al., 2005). Regarding the four quantitative non-randomized studies, two studies (Larkin et al., 2014; Mascarenhas et al., 2009) fulfil four of five criteria while the other two studies met three criteria (Paradis et al., 2015; Samuel et al., 2019). We could not assess criterion number 5 of the quantitative non-randomized studies (if the studies are administered as intended), as it was not explicitly reported. Assessing the other seven studies, all studies met three of five criteria (Ahmed et al., 2017; Catteeuw et al., 2010; Elsworthy et al., 2014a; Emmonds et al., 2015; Gomez-Carmona and Pino-Ortega, 2016; Mallo et al., 2012; Oudejans et al., 2005). Meanings of and reasons for this assessment will be reconsidered at a later stage of this review. An overview on the assessment of the risk of bias is given in Table 2.

Aspects of physical load

Results of the included studies were summarized in Table 3. Eight studies investigated the relationship between match period and RDM (Ahmed et al., 2017; Catteeuw et al., 2010; Emmonds et al., 2015; Gomez-Carmona and Pino-Ortega, 2016; Larkin et al., 2014; Mallo et al., 2012; Mascarenhas et al., 2009), followed by six studies that examined the relationship between referees' velocity and RDM (EL; Catteeuw et al., 2010; Elsworthy et al., 2014a; Emmonds et al., 2015; Gomez-Carmona and Pino-Ortega, 2016; Mascarenhas et al., 2009; Oudejans et al., 2005). Three studies investigated the relationship between the referees' heart rate and RDM (Emmonds et al., 2015; Gomez-Carmona and Pino-Ortega, 2016; Mascarenhas et al., 2009) and two studies assessed the relationship between the total distance covered and RDM (Emmonds et al., 2015;

Table 2. Assessment of the risk of bias (appraised with the MMAT).

Quantitative descriptive studies	Ahmed et al. (2017)	Catteeuw et al. (2010)	Elsworthy et al. (2014)	Emmonds et al. (2015)	Gomez-Carmona and Pino Ortega (2016)	Mallo et al. (2012)	Oudejans et al. (2005)
Is the sampling strategy relevant to address the research question?	1	1	1	1	1	1	1
Is the sample representative of the target population?	1	1	1	1	1	1	1
Are the measurements appropriate?	0	0	0	0	0	0	0
Is the risk of non-response bias low?	0	0	0	0	0	0	0
Is the statistical analysis appropriate to answer the research question?	1	1	1	1	1	1	1
Quantitative non-randomized studies	**Larkin et al. (2014)**	**Mascarenhas et al. (2009)**	**Paradis et al. (2015)**	**Samuel et al. (2019)**			
Are the participants representative of the target population?	1	1	1	1			
Are measurements appropriate regarding both the outcome and intervention (or exposure)?	1	1	1	1			
Are there complete outcome data?	1	1	1	1			
Are the confounders accounted for in the design and analysis?	1	1	0	0			
During the study period, is the intervention administered (or exposure occurred) as intended?	c	c	c	c			

*Legend: 1 = yes, 0 = no; c = can't tell

Table 3. Results of the included studies.

Load	Study	Description	p	d	r
A	Ahmed et al. (2017)	1st half vs. 2nd half: 91.1 ± 14.9% vs 73.3 ± 17.4% correct decisions 1.7 ± 5.3% vs. 8.8% contentious decisions 7.2 ± 14.4% vs. 17.9 ± 17.1% incorrect decisions 24.8 ± 18.4% vs. 75.2 ± 18.4% missed fouls	 0.007 0.068 0.041 0.002	 *NR* *NR* *NR* *NR*	 *NR* *NR* *NR* *NR*
	Catteeuw et al. (2010)	Uniform distribution of correct and incorrect decisions over six 15-minutes periods	0.91	*NR*	*NR*
	Emmonds et al. (2015)	Moderate increase of accuracy from minutes 40-50 to 50-60 Moderate decrease of accuracy in the last 10 minutes of the game No difference in decision-making accuracy in 10-min periods	*NR* *NR* 0.557	0.7 0.86 *NR*	*NR* *NR* *NR*
	Gomez-Carmona and Pino Ortega (2016)	69% mistakes in 1st half vs. 31% in 2nd half	*NR*	*NR*	*NR*
	Larkin et al. (2014)	Improvement of decision-making performance from quarter two to quarter four Improvement of decision-making performance from quarter three to quarter four	0.002 0.001	*NR* *NR*	*NR* *NR*
	Mallo (2012)	R: no uniform distribution of errors for six 15-min periods. Higher error percentage in last 15-min: 1st 15-min= 7%| 2nd 15-min= 9%| 3rd 15-min= 12%| 4th 15 min= 17%| 5th 15-min= 9%| 6th 15-min= 23% AR: no uniform distribution of errors for six 15-min periods. Higher error percentage in last 15- min: 1st 15-min= 6%| 2nd 15-min= 10%| 3rd 15-min= 10%| 4th 15 min= 19%| 5th 15-min= 25%| 6th 15-min= 21%	0.013 0.002	*NR* *NR*	*NR* *NR*
	Mascarenhas et al. (2009)	Less accurate in the first 15-min in each half, but not statistically significant: 1st 15-min= 51%| 2nd 15-min= 69%| 3rd 15-min= 70%	*NR*	*NR*	*NR*
	Samuel et al. (2019)	Lower accuracy in the second quarter compared to the third quarter Lower accuracy in the fourth quarter compared to the third quarter No difference between first and second section (both lasts 30 minutes)	*NR* *NR* 0.21	0.91 0.66 *NR*	*NR* *NR* *NR*
B	Catteeuw et al. (2010)	The error rate was uniformly distributed over six 15-min periods regarding movement speed of AR	0.24	*NR*	*NR*
	Elsworthy et al. (2014)	No relationship between correct and incorrect decisions at instantaneous velocity at the time of the decision Lower running speed at incorrect decisions over the prior 5 seconds compared to correct decisions	< 0.05 0.018	*NR* *NR*	*NR* *NR*
	Emmonds et al. (2015)	No relationship between the referees' velocity and decision-making accuracy	*NR*	*NR*	0.36
	Gomez-Carmona and Pino Ortega (2016)	Most errors at slow velocity of the R (35.2%): >18 km/h ~ 4% | 13-18 km/h ~ 8% | 7.2-13 km/h ~ 19% | 3.6-7.2 km/h ~ 29%	*NR*	*NR*	*NR*
	Mascarenhas et al. (2009)	No relationship between referees velocity and decision-making accuracy	0.9	*0.02*	*NR*
	Oudejans et al. (2005)	R make more errors when running or sprinting (>8 km/h) compared when jogging or walking	0.05	*NR*	*NR*
C	Paradis et al. (2015)	No relationship between running time and decision-making performance	> 0.05	*NR*	*NR*
D	Emmonds et al. (2015)	No correlation between distance covered and decision-making accuracy	*NR*	*NR*	-0.02
E	Emmonds et al. (2015)	No correlation between heart rate (mean) and decision-making accuracy	*NR*	*NR*	0.2
	Gomez-Carmona and Pino Ortega (2016)	Errors occur when heart rate above 75%, especially when HR is between 85-95% (67.7%) and over 95% (26.7%)	*NR*	*NR*	*NR*
	Mascarenhas et al. (2009)	No difference between heart rate at correct (165.5 ± 12.5 bpm) and incorrect decisions (165.6 ± 13.3 bpm)	0.9	0.01	*NR*
F	Larkin et al. (2014)	No relationship between physical exertion (blood lactate) and decision-making performance for each quarter: physical exertion 1 – quarter one physical exertion 2 – quarter two physical exertion 3 – quarter three physical exertion 4 – quarter four	 0.568 0.847 0.771 0.579	 *NR* *NR* *NR* *NR*	 *NR* *NR* *NR* *NR*

*Load: A = match period, B = velocity, C = running time, D = distance covered, E = heart rate, F = blood lactate.
†Abbreviations: R = referees, AR = assistant referees, AFO = Australian field officials, NR= not reported, p = statistical significance, d = effect size, r = correlation, ± = standard deviation

Mascarenhas et al., 2009). Finally, one study examined the relationship between blood lactate (Larkin et al., 2014) respectively running time and RDM (Paradis et al., 2015).

Match period. Five studies found a negative relationship between match period and RDM, which means that with increasing playing time RDM decreased (Ahmed et al., 2017; Emmonds et al., 2015; Gomez-Carmona and Pino-Ortega, 2016; Mallo et al., 2012; Samuel et al., 2019). Another five studies did not find evidence of a negative or positive relationship between match period and RDM (Ahmed et al., 2017; Catteeuw et al., 2010; Emmonds et al., 2015; Mascarenhas et al., 2009; Samuel et al., 2019).

Two studies revealed an increase of RDM with increasing playing time (Emmonds et al., 2015; Larkin et al., 2014).

Samuel et al. (2019; soccer) found a drop in RDM from the second to the third quarter (effect size: d = 0.91) as well as from the third to the fourth quarter (d = 0.66). Two studies, carried out in futsal (Ahmed et al., 2017) and soccer (Mallo et al., 2012), showed that RDM dropped in the second half of a game. In contrast, Samuel et al. (2019) did not find a difference between half times. Likewise, Ahmed et al. (2017) did not find a drop in RDM at contentious decisions either. Three studies did not reveal a relationship between 10-minute or 15-minute periods and RDM (soccer: Catteeuw et al., 2010; rugby: Emmonds et al., 2015; soccer: Mascarenhas et al., 2009). However, two studies found a decrease in RDM in the last 10 minutes (rugby: d = 0.86, Emmonds et al., 2015) respectively 15 minutes (soccer: Mallo et al., 2012) of a match. At last, two studies revealed a positive relationship between playing time and RDM (rugby: Emmonds et al., 2015; Australian football: Larkin et al., 2014).

Velocity. The Relationship between referees' velocity and RDM was analysed with the instantaneous referees' velocity at the time of the decision (Catteeuw et al., 2010; Elsworthy et al., 2014a; Emmonds et al., 2015; Gomez-Carmona and Pino-Ortega, 2016; soccer: d = 0.02, Mascarenhas et al., 2009; Oudejans et al., 2005) and with the running speed prior to the decision (Elsworthy et al., 2014a).

Examining first the relationship between the referees' instantaneous velocity and RDM, two studies revealed a decrease of RDM with increasing velocity (soccer: Gomez-Carmona and Pino-Ortega, 2016; soccer: Oudejans et al., 2005) while four studies did not indicate a positive or negative relationship between the referees' velocity and RDM (soccer: Catteeuw et al., 2010; Australian football: Elsworthy et al., 2014a; rugby: Emmonds et al., 2015; soccer: d = 0.02, Mascarenhas et al., 2009). A closer look at the two studies that reveal a negative relationship between the referees' velocity and RDM indicates contradictory findings: Gomez-Carmona and Pino-Ortega (2016) found that most errors occur at slow velocity (<8 km/h; soccer), while Oudejans et al. (2005) found that errors occur more often at higher velocities (> 8km/h; soccer). As mentioned, four studies show no change of RDM in terms of the distribution of errors according to six different categories of speed (standing: 0.0 – 0.1 m/s, walking: 0.2 – 1.9 m/s, jogging: 2.0 – 3.9 m/s, running: 4.0 – 5.4 m/s, high-intensity running: 5.5 – 6.9 m/s, sprinting: >7.0 m/s; soccer: Catteeuw et al., 2010) nor to the instantaneous velocity at the time of the decision (Australian football: Elsworthy et al., 2014a; rugby: Emmonds et al., 2015; soccer: Mascarenhas et al., 2009).

At least, Elsworthy et al. (2014a) investigated the relationship between RDM and the referees' velocity five seconds, thirty seconds, one minute and five minutes prior to the decision. Results indicate that the referees' velocity five seconds prior to incorrect decisions was significantly higher compared to correct decisions.

Running time. Paradis et al. (2015) investigated the relationship between running time and RDM in Australian football (Australian football: Paradis et al., 2015), but they did neither find a negative nor positive relationship.

Distance covered. No relationship was found between the total distance covered and RDM (rugby: Emmonds et al., 2015; soccer: d = 0.27, Mascarenhas et al., 2009).

Physiological parameters. Two studies did not find a relationship between *heart rate* and RDM (rugby: Emmonds et al., 2015; soccer: d = 0.01, Mascarenhas et al., 2009), while one study revealed a drop in RDM especially when the heart rate was above 85% (soccer: Gomez-Carmona and Pino-Ortega, 2016). One study examined physical exertion with blood lactate did not find evidence for a relationship between physical exertion and RDM (Australian football: Larkin et al., 2014).

Overall, most findings in this review did not reveal a negative or positive relationship between physical load and RDM (n = 18; Ahmed et al., 2017; Catteeuw et al., 2010; Elsworthy et al., 2014a; Emmonds et al., 2015; Larkin et al., 2014; Mascarenhas et al., 2009; Paradis et al., 2015; Samuel et al., 2019). However, n = 13 findings point to a negative relationship between physical load and RDM (Ahmed et al., 2017; Elsworthy et al., 2014a; Emmonds et al., 2015; Gomez-Carmona and Pino-Ortega, 2016; Mallo et al., 2012; Oudejans et al., 2005; Samuel et al., 2019), while only two studies suggest an increase of RDM with increasing load (match period; Emmonds et al., 2015; Larkin et al., 2014).

Discussion

Empirical evidence

Recent studies examined the physical demands imposed on referees across various sports games, especially in soccer (Castagna et al., 2007) and Australian football (Elsworthy et al., 2014b). Current trends in competitive sports show increasing game dynamics and thus altered physical load not only for players but also referees (Bilge, 2012; Dolański et al., 2017; Mascarenhas et al., 2005; Weston et al., 2007). Hence, referees are required to cope with physical load (Brightmore et al., 2016; Dolański et al., 2017; Leicht, 2008; Martin et al., 2001; Weston et al., 2007) when making decisions in a match or day by day in a tournament (Helsen and Bultynck, 2004). Thus, it is important to focus on referees' physical fitness and the relationship between physical load and RDM, as RDM might decrease with increasing physical load (Helsen et al., 2019; Schmidt et al., 2019). Therefore, the aim of this review was to provide an overview of studies examining the relationship between physical load and RDM. Our findings suggest that empirical evidence is unclear and contradictory.

Methodological assessment

With regard to the risk of bias, an advantage of the seven studies conducting video-analyses is that the referees made decisions in real-world situations. Thus, these studies probably have high external validity. However, such approach does not enable to control and vary the independent variables nor to identify and control confounding variables. Consequently, the above seven studies' conclusions on the relationship between physical load and RDM have to be treated carefully as the results might be impaired due to confounding variables, for example crowd noise (e.g., Unkelbach and Memmert, 2010). Lab-based studies, in turn, may ensure high internal validity by identifying and controlling confounding variables, systematically varying the independent variable(s) and minimizing error variance (Aziz, 2017). Hence, lab-based studies can provide first understanding of the suggested systematic relationship between physical load and RDM as well as about confounding variables. Overall, to comprehensively add to both theory and practice, and reduce the risk of bias, we recommend a combination of a lab-based and field-based. Both have specific advantages, in systematic research programs comprising several studies which build upon each other (Büsch, 2019). Research programs should base on theoretical models that allow deriving testable hypotheses.

Furthermore, most studies examined the relationship between external load and RDM rather than internal load and RDM. This may be problematic because external load is more a kind of description of a physical task (e.g., distance). Internal load, however, is an individual's psychophysiological response to external load (Impellizzeri et al., 2019). Thus, different individuals (e.g., referees) can experience the same external load (e.g., operationalized by match period) as differently exhausting (Impellizzeri et al., 2019). Therefore, we recommend that future studies should test the hypothesized relationship between physical load and RDM with consideration of external and internal load. By controlling the internal load, it can be ensured that an individualized external load leads to a comparable internal load, e.g. relative heart rate (Impellerizi et al., 2019), in order to examine the relationship between physical load and RDM inter-individually and to make intra-individual comparisons more reliable.

Limitations

The present review does not provide a clear picture of evidence regarding the relationship between physical load and RDM. First of all, this might be due to the low number of studies included in this review. Furthermore, studies were conducted in different sports and this may render comparisons across studies potentially difficult, although it also remains open whether there are any sport-specific differences at all. Likewise, consideration of different referee types (e.g., referee or assistant referee in soccer) across studies might have enhanced heterogeneity in findings, assuming that different referee types are faced with different physical as well as decision-making demands (Mallo et al., 2012).

Conclusion

Based on this scoping review, we conclude that the current state of evidence on the relationship between physical load and RDM is contradictory and thus inconclusive. This refers to evidence across and within physical parameters. Finally, in view of the risk of bias described above, we suggest a combination of lab- and field-based approaches as part of systematic research programs aiming to determine the impact of selected physical load parameters on RDM under naturalistic conditions.

Acknowledgements

This research was funded from the Federal Institute of Sport Science based on a resolution of the German Bundestag (official file number: ZMVI4-071001/19/20). The authors have no conflict of interest directly relevant to the content of the study. The authors declare the present study comply with the current laws of the country in which they were performed.

References

Ahmed, H., Davison, G. and Dixon, D. (2017) Analysis of activity patterns, physiological demands and decision-making performance of elite Futsal referees during matches. *International Journal of Performance Analysis in Sport* **17,** 737-751.

Arnsten, A.F. (2009) Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience* **10,** 410-22.

Aziz, H. (2017) Comparison between field research and controlled laboratory research. *Archives of Clinical and Biomedical Research* **1,** 101-104.

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

Bilge, M. (2012) Game analysis of Olympic, World and European Championships in men's handball. *Journal of Human Kinetics* **35,** 109-118.

Brightmore, A., O'Hara, J., Till, K., Cobley, S., Hubka, T., Emmonds, S. and Cooke, C. (2016) Movement and physiological demands of Australasian National Rugby League referees. *International Journal of Sports Physiology and Performance* **11,** 1080-1087.

Büsch, D. (2019) Wandel zwischen den erfahrungswelten. Fehler und fallen einer anwendungsnahen sportwissenschaft [Changing between experiential worlds. Mistakes and pitfalls of an application-oriented sports science]. *German Journal of Exercise and Sport Research* **49,** 201-203.

Castagna, C., Abt, G. and D'Ottavio, S. (2007) Physiological aspects of soccer refereeing performance and training. *Sports Medicine* **37,** 625-646.

Catteeuw, P., Gilis, B., Wagemans, J. and Helsen, W. (2010) Offside decision making of assistant referees in the English Premier League: impact of physical and perceptual-cognitive factors on match performance. *Journal of Sports Sciences* **28,** 471-481.

Chang, Y.K., Labban, J.D., Gapin, J.I. and Etnier, J.L. (2012) The effects of acute exercise on cognitive performance: a meta-analysis. *Brain Research* **1453,** 87-101.

Dolański, B., Szwarc, A., Heinig, B. and Sitek, M. (2017) Physical activity profile of the referee and the assistant referee during official football matches. *Baltic Journal of Health and Physical Activity* **9,** 97-105.

Elsworthy, N., Burke, D. and Dascombe, B.J. (2014a) Factors relating to the decision-making performance of Australian football officials. *International Journal of Performance Analysis in Sport* **14,** 401-410.

Elsworthy, N., Burke, D., Scott, B.R., Stevens, C.J. and Dascombe, B.J. (2014b) Physical and decision-making demands of Australian football umpires during competitive matches. *Journal of Strength and Conditioning Research* **28,** 3502-3507.

Emmonds, S., O'Hara, J., Till, K., Jones, B., Brightmore, A. and Cooke, C. (2015) Physiological and movement demands of rugby league referees: influence on penalty accuracy. *Journal of Strength and Conditioning Research* **29,** 3367-3374.

Gomez-Carmona, C.D. and Pino-Ortega, J. (2016) Kinematic and physiological analysis of the performance of the football referee and its relationship with decision making. *Journal of Human Sport and Exercise* **11,** 397-414.

Helsen, W. and Bultynck, J.B. (2004) Physical and perceptual-cognitive demands of top-class refereeing in association football. *Journal of Sports Sciences* **22,** 179-89.

Helsen, W., MacMahon, C. and Spitz, J. (2019) Decision making in match officials and judges. In: *Anticipation and Decision Making in Sport*. Eds: Williams, A.M. and Jackson, R.C. London: Routledge. 250-266.

Hong, Q.N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B.,

O'Cathain, A., Rousseau, M.-C., Vedel, I. and Pluye, P. (2018) Mixed Methods Appraisal Tool (MMAT) version 2018 for infomration professionals and researchers. *Education for Information (Special Issue)*.
Impellizzeri, F.M., Marcora, S.M. and Coutts, A.J. (2019) Internal and external training load: 15 years on. *International Journal of Sports Physiology and Performance* **14,** 270-273.
Kittel, A., Larkin, P., Elsworthy, N. and Spittle, M. (2019) Identification of key performance characteristics of elite Australian football umpires. *International Journal of Sports Science & Coaching*.
Larkin, P., O'Brien, B., Mesagno, C., Berry, J., Harvey, J. and Spittle, M. (2014) Assessment of decision-making performance and in-game physical exertion of Australian football umpires. *Journal of Sports Sciences* **32,** 1446-1453.
Leicht, A.S. (2008) Physiological demands of basketball refereeing during international competition. *Journal of Science and Medicine in Sport* **11,** 357-60.
MacMahon, C., Mascarenhas, D., Plessner, H., Pizzera, A., Oudejans, R. and Raab, M. (2015) *Sports officials and officiating: science and practice*. Abingon, UK: Routledge.
Mallo, J., Frutos, P.G., Juarez, 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. *Journal of Sports Sciences* **30,** 1437-1445.
Martin, J., Smith, N.C., Tolfrey, K. and Jones, A.M. (2001) Activity analysis of English Premiership rugby football union refereeing. *Ergonomics* **44,** 1069-1075.
Mascarenhas, D., Button, C., O'Hare., D.O. and Dicks, M. (2009) Physical performance and decision making in association football referees: a naturalistic study. *The Open Sports Sciences Journal* **2,** 1-9.
Mascarenhas, D., Collins, D. and Mortimer, P. (2005) Elite refereeing performance: developing a model for sport science support. *The Sport Psychologist* **19,** 364-379.
Moher, D., Liberati, A., Tetzlaff, J. and Altman, D.G. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of Clinical Epidemiology* **62,** 1006-1212.
Oudejans, R.R.D., Bakker, F.C., Verheijen, R., Gerrits, J.C., Steinbrückner, M. and Beek, P.J. (2005) How position and motion of expert assistant referees in soccer relate to the quality or their offside judgements during actual match play. *International Journal of Sport Psychology* **36,** 3-21.
Paradis, K., Larkin, P. and O'Connor, D. (2015) The effects of physical exertion on decision-making performance of Australian football umpires. *Journal of Sports Sciences* **34,** 1535-1541.
Plessner, H. and Haar, T. (2006) Sports performance judgments from a social cognitive perspective. *Psychology of Sport and Exercise* **7,** 555-575.
Plessner, H. and MacMahon, C. (2013) The sport official in research and practice. In: *Developing sport expertise*. Eds: Farrow, D., Baker, J. and MacMahon, C. London: Routledge. 71-88.
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. *Progress in Brain Research* **174,** 151-158.
Reilly, T. and Gregson, W. (2006) Special populations: the referee and assistant referee. *Journal of Sports Sciences* **24,** 795-801.
Rullang, C., Emrich, E. and Pierdzioch, C. (2017) Why do referees end their careers and which factors determine the duration of a referee's career? *Current Issues in Sport Science* **2,** 1-13.
Samuel, R.D., Galily, Y., Guy, O., Sharoni, E. and Tenenbaum, G. (2019) A decision-making simulator for soccer referees. *International Journal of Sports Science & Coaching* **14,** 480-489.
Schmidt, S.L., Schmidt, G.J., Padilla, C.S., Simoes, E.N., Tolentino, J.C., Barroso, P.R., Narciso, J.H., Godoy, E.S. and Costa Filho, R.L. (2019) Decrease in Attentional Performance After Repeated Bouts of High Intensity Exercise in Association-Football Referees and Assistant Referees. *Frontiers in Psychology* **10,** 2014.
Schweizer, G., Plessner, H., Kahlert, D. and Brand, R. (2011) A Video-Based Training Method for Improving Soccer Referees' Intuitive Decision-Making Skills. *Journal of Applied Sport Psychology* **23,** 429-442.
Tod, D. (2019) *Conducting Systematic Reviews in Sport, Exercise, and Physical Activity*. Cham, Switzerland: Palgrave Macmillan/Springer.
Tomporowski, P.D. (2003) Effects of acute bouts of exercise on cognition. *Acta Psychologica* **112,** 297-324.
Unkelbach, C. and Memmert, D. (2010) Crowd noise as a cue in referee decisions contributes to the home advantage. *Journal of Sport & Exercise Psychology* **32,** 483-498.
Weston, M., Castagna, C., Impellizzeri, F.M., Rampinini, E. and Abt, G. (2007) Analysis of physical match performance in English Premier League soccer referees with particular reference to first half and player work rates. *Journal of Science and Medicine in Sport* **10,** 390-397.

Key points

- Officiating in sport games requires high-level decision-making under various constraints.
- A scoping review was conducted on the relationship between physical load and RDM.
- Most studies suggest that physical load is not or negatively correlated with RDM.
- Conclusions, however, are limited due to the low internal validity of the included studies.
- Systematic research programs with targeted isolated testing of the relationship between physical load and RDM are recommended.

AUTHOR BIOGRAPHY

Nicolas BLOß
Employment
Institute of Sport Science, Carl von Ossietzky University Oldenburg
Degree
PhD Student, M.A.
Research interests
judgement and decision-making in sport, sport performance, cognition, exercise science, sports sciences
E-mail:
nicolas.bloss1@uni-oldenburg.de

Jörg SCHORER
Employment
Institute of Sport Science, Carl von Ossietzky University Oldenburg
Degree
Full-professor, PhD
Research interests
Talent in sport, strength and conditioning, expertise, cognition, perception, motor learning and motor control
E-mail:
joerg.schorer@uni-oldenburg.de

Florian LOFFING
Employment
Institute of Sport Science, Carl von Ossietzky University Oldenburg
Degree
PhD
Research interests
Laterality, expertise, cognition, perception, motor learning and motor control
E-mail:
florian.loffing@uni-oldenburg.de

Dirk BÜSCH
Employment
Institute of Sport Science, Carl von Ossietzky University Oldenburg
Degree
Full-professor, PhD
Research interests
Exercise performance, exercise science, sports sciences, exercise testing, exercise physiology, sport physiology, sport biomechanics, muscle physiology, talent in sport
E-mail: dirk.buesch@uni-oldenburg.de

✉ **Nicolas Bloß**
Institute of Sport Science, Carl von Ossietzky University of Oldenburg, Ammerländer Heerstr. 114-118, 26129 Oldenburg, Germany