Psychology of Sport & Exercise 79 (2025) 102841

Contents lists available at ScienceDirect

Psychology of Sport & Exercise

journal homepage: www.elsevier.com/locate/psychsport

Review

The effectiveness of decision-making training in team-sport officials: A systematic review and meta-analysis

Aden Kittel [a,b,*], Riki Lindsay [c], Paul Larkin [d,e], Michael Spittle [d], Ian Cunningham [f]

[a] School of Exercise and Nutrition Sciences, Deakin University, Geelong, Australia
[b] Centre for Sport Research, Institute for Physical Activity and Nutrition, Deakin University, Geelong, Australia
[c] Institute of Education, Arts and Community, Federation University, Ballarat, Australia
[d] Institute for Health and Sport, Victoria University, Melbourne, Australia
[e] Maribyrnong Sport Academy Research Centre, Melbourne, Australia
[f] School of Applied Sciences, Edinburgh Napier University, Scotland, UK

ARTICLE INFO

Keywords:
Officials
Umpire
Referee
Decision making
Video training

ABSTRACT

Purpose: Decision making is a critical skill for sports officials, often directly influencing the flow and fairness of a match. While this topic has received considerable interest in the literature, a synthesis of current evidence to understand the effectiveness of decision-making training interventions remains unexplored. Therefore, the aim of this study was to conduct a systematic review and meta-analysis of decision-making interventions in team sport officials.

Principal results: A total of 14 studies were identified, with a random-effects meta-analysis revealing an overall moderate positive effect of decision-making training on decision-making performance outcomes ($g = 0.68$, $p < .001$) compared to control conditions. Notably, decision-making training was more effective in Soccer ($g = 1.05$), Rugby Union ($g = 0.90$), but not for Australian Football ($g = 0.24$). Video-based (i.e., 2-D footage) showed significant improvements, especially for objective decision-making outcomes like offside identification ($g = 1.48$, $p < .001$). However, our findings indicated that decision-making training tends to be less effective for subjective decision-making outcomes that requires higher levels of interpretation. Furthermore, shorter interventions (4–6 weeks) were found to be most effective, with performance improvements reducing as interventions increased in time.

Major conclusions: Our findings highlight the need for further research to explore alternative technologies such as virtual reality to understand how to better replicate game scenarios and assess the transferability of decision-making training to real-world officiating contexts. Additionally, this review highlights the need to investigate sports beyond Soccer, Rugby, and Australian Rules Football to develop our understanding further into optimising decision-making training in sports officials.

1. Introduction

Officials play a pivotal role in sporting contests, often facing scrutiny as the decisions they make can potentially impact the outcome of games (Larkin et al., 2011). As such, decision making is considered one of, if not the most important skill for sport officials (i.e., referees, umpires, judges) (Kittel et al., 2019b; Morris & O'Connor, 2017). Decision making involves planning, selecting, and executing an action based on both the individual (i.e., the official)'s knowledge, and the information available in a situation (Williams & Ford, 2013). This is a complex process where officials are applying the rules of the sport to determine whether an infringement has occurred (Morris & O'Connor, 2017), while also incorporating elements, of accuracy, fairness, safety and entertainment (Russell et al., 2019). Decision-making processes differ among officials from different sports, due to factors such as the number of the decisions they need to make, and the patterns and complexity of cues needed to make the decision.

Team sport officials are often categorised as 'interactors' MacMahon et al. (2014), where they have high interaction and physical movement demands and large number of perceptual cues to attend to. Most research has explored the performance and development of decision making within a single category, the interactor official, particularly

---

* Corresponding author. Centre for Sport Research, School of Exercise and Nutrition Sciences Deakin University Melbourne, Australia.
E-mail address: Aden.Kittel@deakin.edu.au (A. Kittel).

https://doi.org/10.1016/j.psychsport.2025.102841
Received 13 November 2024; Received in revised form 12 March 2025; Accepted 14 March 2025

Soccer referees (Cunningham et al., 2022; Hancock et al., 2021; Kittel et al., 2021). Within the game, decisions vary in terms of their objectivity (e.g., whether a ball was in or out, or a player was onside or offside), or their subjectivity (e.g., was a player held 'enough' by their opposition, or the 'holding the ball' in Australian Football; see Larkin, Mesagno, Berry, and Spittle (2018). There are also suggested desirable game imperatives to be considered alongside decision accuracy, such as fairness, safety and entertainment (Russell et al., 2019). Although these factors are linked to decision-making performance, it is unknown how they can be incorporated within a training environment effectively. Given the interpretive nature of some decisions and the importance of considering context, Brand et al. (2009) discussed the concept of 'accurate' vs 'adequate' decisions. For example, an 'accurate' decision could be to the letter of the law, without considering factors such as player intent, game context, or prior decisions. Some objective decisions (e.g., whether a ball was in or out, or a player was onside or offside) need to be consistently made accurately and to the letter of the law. 'Adequate' decisions, on the other hand, consider additional factors such as game context (e.g., 'putting the whistle away', 'letting the players play'), severity of the infraction, and more closely align with the model of fairness, safety and entertainment (Russell et al., 2019). Evidently, decision making is a complex process with several (sometimes conflicting) priorities leading to the final decision. Despite this, most research has explored developing decision-making accuracy in team sport officials where the correct decisions are prescribed through specific scenarios in training and rely on evaluation from expert consensus (Kittel et al., 2021). There is limited quantitative evidence, however, to support whether this approach is effective for developing decision-making performance.

Unlike athletes, sport officials are limited in the training methods they can employ to develop their decision-making, resulting in some suggestions about the limited 'practice richness' of officiating training environments (MacMahon et al., 2014). For athletes to develop decision-making skills, coaches can incorporate small-sided games to imitate in-game decision-making demands, which may represent an ecologically valid approach (O'Connor et al., 2017). However, this approach can be more difficult for officials, who either need to attend players' training sessions to officiate these game-based scenarios/simulations or play game-based scenarios to officiate themselves. A limitation of this approach is that this would increase the physical load and injury risk to the officials, plus the skill level of the small-sided game simulations may not be reflective of the player skill they officiate on gameday. To combat the potential issues with on-field decision-making practice, officiating coaches and developers have favoured more off-field training methods to develop their decision-making skill (Cunningham et al., 2022). In certain sporting contexts, officials identify decision-making skills as very to extremely important to their role, however indicated that they only purposefully trained this skill less than once a month (McEwan et al., 2024).

Officiating coaches commonly use video-based training to develop decision-making performance (Kittel et al., 2021). In video-based training, officials are usually presented with video footage of in-game decision-making scenarios and are then required to officiate the footage like they would in a game (which can occur individually of in group discussion formats). Also, for video-based training sometimes the video presented is manipulated such as blurring images (van Biemen et al., 2018) or altering the video speed (Put et al., 2016) to understand varied effects to normal presentation. The premise of this training modality is to create engaging visual representations to help officials build more elaborate knowledge structures about decision situations (Macarenhas et al., 2005) and introduce common constraints of the performance environment to assist in the development of decision-making accuracy (Pinder et al., 2011). Reviewing and deciding on video scenarios, however, might be a more passive training approach that has been critiqued by O'Brien and Rynne (2021) as 'missing the mark' due to the limited representativeness of this practice activity to the actual performance environment. Following the observations of the field of video-based training by O'Brien and Rynne (2021), immersive learning and representative approaches have now emerged and been tested, namely Virtual Reality (VR) training (van Biemen et al., 2023). Furthermore, video-based training programs are generally reliant on isolated performance and assessment (with some studies exercising follow-up retention tests to evaluate performance on the same video task; Larkin et al., 2018), with researchers highlighting the limited understanding of the transfer of skills to in-game performance (Kittel et al., 2019c). Despite this, officials such as Soccer referees view decision-making training using video clips to be very important but spend very little time completing this form of training (McEwan et al., 2023). The effectiveness of these various programs therefore warrants further investigation, particularly if officiating coaches and sporting organisations are going to embed these practices into their training of sport officials.

With increasing research into the training of decision making in team sports officials (Kittel et al., 2021), there is a need for a meta-analysis on the outcomes of decision-making training to guide potential recommendations. This information would assist in understanding the effectiveness of these methods to aid practitioners in their design of decision-making training tools, and researchers in designing and evaluating evidence-based training interventions. Therefore, the aim of the current study was to conduct a systematic review and meta-analysis of the team sports officials decision-making training literature to provide an evidence-based synthesis of the current body of knowledge. In particular, the study will address the following research questions:

I. Does decision-making training improve team sports officials' overall decision-making skills and if so, by what magnitude of change?
II. How do different conditions of decision-making training (training volume, type, length) influence relative changes in decision-making performance?
III. How are relative changes in decision-making performance as a result of training influenced by sport-type and expertise?

2. Methods

2.1. Search strategy

The present review was conducted according to PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols guidelines (Page et al., 2021). A literature search was conducted in the following electronic databases: MEDLINE, Web of Science, SPORTDiscus, and PsychINFO. This search was conducted by the first author (AK), and databases were searched up to October 2024. The search was not restricted by year of publication, and only English papers were included. The reference lists of included studies, relevant reviews and books were also screened to identify other possible inclusions. Please see supplementary materials for a full description of the search strategy that was implemented (Appendix I).

2.2. Eligibility criteria

Following an initial search of relevant literature, eligibility criteria were drafted and amended by the author team. In line with the PICOS approach (Liberati et al., 2009) inclusion criteria were selected as follows:

I. Participants: Studies involving team sports officials (i.e., referees, umpires).
II. Intervention: Studies were included if they investigated the effectiveness of an intervention for the development of decision-making skill and characteristics in sport officials.

III. Comparison*:* Studies observing interventions-alone and differences in outcome variables through different modes of decision-making practice. It was not necessary for studies to include a Control group.
IV. Outcome: Studies need to include an outcome measure related to the effectiveness of decision-making interventions on the learning of motor skills. This needed to be measured through a pre-test and post-test (including optional additional retention and/or transfer test [s]).
V. Study design: This review considered only randomized control trials and non-randomized studies (e.g., quasi-experimental designs).

2.3. Study selection

All articles identified from the search were exported into the screening and data extraction tool Covidence (Covidence online systematic review platform, Veritas Health Innovation Ltd., Melbourne Australia) with all duplicates removed. Titles and abstracts were independently screened by the first and second author to assess the papers' suitability, and all papers beyond the scope of this systematic review were excluded. All eligible articles then proceeded to full-text screening, where the first and second author independently assessed their eligibility according to predetermined inclusion and exclusion criteria. Any disagreements between the reviewers were resolved by consensus or arbitration through the senior author.

2.4. Data extraction

Extraction was undertaken by the first author, while the second author checked the extracted data for accuracy and completeness. Disagreements were resolved by consensus or by the senior author. Reviewers were not blinded to authors, institutions or manuscript journals. The Cochrane Consumers and Communication Review Group's data extraction protocol was used to extract information from eligible studies, including the following variables: (1) publication details (author, year); (2) participant characteristics (number of participants, country, age, sex, skill level); (3) sport investigated, training tool/technology implemented; (4) intervention duration and session frequency, (5) the testing protocol, type of outcome measures used; (6) and effect sizes. Skill level of the officials was defined by Kittel et al. (2019c), in three performance levels: Elite, Sub-elite, and Amateur.

2.5. Assessment of reporting quality and risk of bias

The third and fourth authors independently evaluated included studies using a modified Downs and Black index to assess risk of bias and reporting quality (Downs & Black, 1998). This scale includes 14 original items and rates each item as 0 or 1, with higher scores indicating higher quality studies (Thurlow et al., 2024; Weakley et al., 2023). When there was no clear information to accurately evaluate an item, a score of 0 was assigned. Articles with disagreements regarding their rating were resolved through discussion and arbitration with the first author.

2.6. Data analysis

Cohen's Kappa was used to determine inter-rater reliability between the two reviewers in the Title & Abstract screening and Full-text screening stages of the review process. As per Landis and Koch (1977), Cohen's Kappa (k) was categorised as poor agreement (<0.00), slight agreement (0.00–0.20), fair agreement (0.21–0.40), moderate agreement (0.41–0.60), substantial agreement (0.61–0.80), almost perfect agreement (0.81–1.00). Decision-making performance data were pooled together using Comprehensive Meta-analysis Version 4 (CMA; Biostat Inc., Englewood, NJ, USA). The effectiveness of decision-making training was assessed by calculating pre–post within-group changes. Where this data was not available from the manuscript (i.e., mean and standard deviation for pre- and post-tests), the authors contacted the corresponding author of included studies for the raw data. This data was not sourced for five studies, leading to nine studies being included in subsequent analysis. Specifically, a random-effects meta-analysis of standardised mean differences (SMD), expressed as Hedges' g where possible, was performed. In this analysis, each study was treated as a random effect and was group within studies. Between-group differences have been calculated where possible to indicate changes between different types of decision-making training. According to Cochrane guidelines (Deeks et al., 2022), statistical heterogeneity was assessed using the Q and I^2 statistics to show the dispersion of true effects, expressed as Hedges' g, between predefined subgroups (Higgins et al., 2021). The I^2 statistic was interpreted as follows: 0 %–40 % was considered 'might not be important', 30 %–60 % 'may represent moderate heterogeneity', 50–90 % 'may represent substantial heterogeneity', and 75–100 % 'considerable heterogeneity' (Higgins et al., 2021). Heterogeneity between studies was explored further via moderator analysis using subgroup analysis for the following categorical variables: (1) sport type; (2) skill level of officials; (3) training tool/technology used; (4) training duration; (5) frequency of training sessions; and (6) outcome measure employed (e.g., offside, flag or non-flag). To compare mean effects of different levels of each sub-group, we conducted *Z*-tests following Borenstein et al. (2019). Possible reporting bias was tested by inspection of the funnel plot and Egger's trim-and-fill method (Egger et al., 1997).

3. Results

The initial search yielded 3260 studies. Following the removal of duplicates, 2708 underwent title and abstract screening, 53 articles were subject to full text review and 39 studies were excluded for not meeting the eligibility criteria. There was moderate agreement (k = 0.52) between the two reviewers at the Title & Abstract screening stage, and substantial agreement (k = 0.80) at the Full-text screening stage. This process resulted in a total of 14 studies included in the review (see Fig. 1).

3.1. Study characteristics

Across all included studies, the total number of participants were 512, with an average sample size of 37 participants per study. When considering the studies by sport, the most investigated sport was Soccer ($n = 9$; 64 %), followed by Australian Football ($n = 4$; 29 %), and Rugby Union was represented in one study. Five (36 %) studies involved sub-elite level officials; four (29 %) studies involved elite level officials, and novice level officials were examined in four studies (29 %). One study involved a mixture of elite and novice officials. The most frequently employed training tool was video technology ($n = 10$, 71 %), followed by web-based/computer simulations ($n = 4$; 29 %). 360°VR, combined video and computer simulations, and paper-based logbook testing were utilised in one study each. A summary of study characteristics and participants of included studies are presented in Table 1.

3.2. Outcomes for the assessment of reporting quality and risk of bias

Table 2 summarises the outcomes of the modified Downs and Black scale for the assessment of reporting quality and risk of bias. Results ranged from 5 to 10, with a mean score of 8.4 ± 1.5.

3.3. Overall effect of decision-making training

A total of 78 individual effects were extracted from the included studies. To account for statistical dependence a composite approach was implemented, providing an overall conservative effect for each study. Using CMA software, this approach uses the mean of study outcomes to

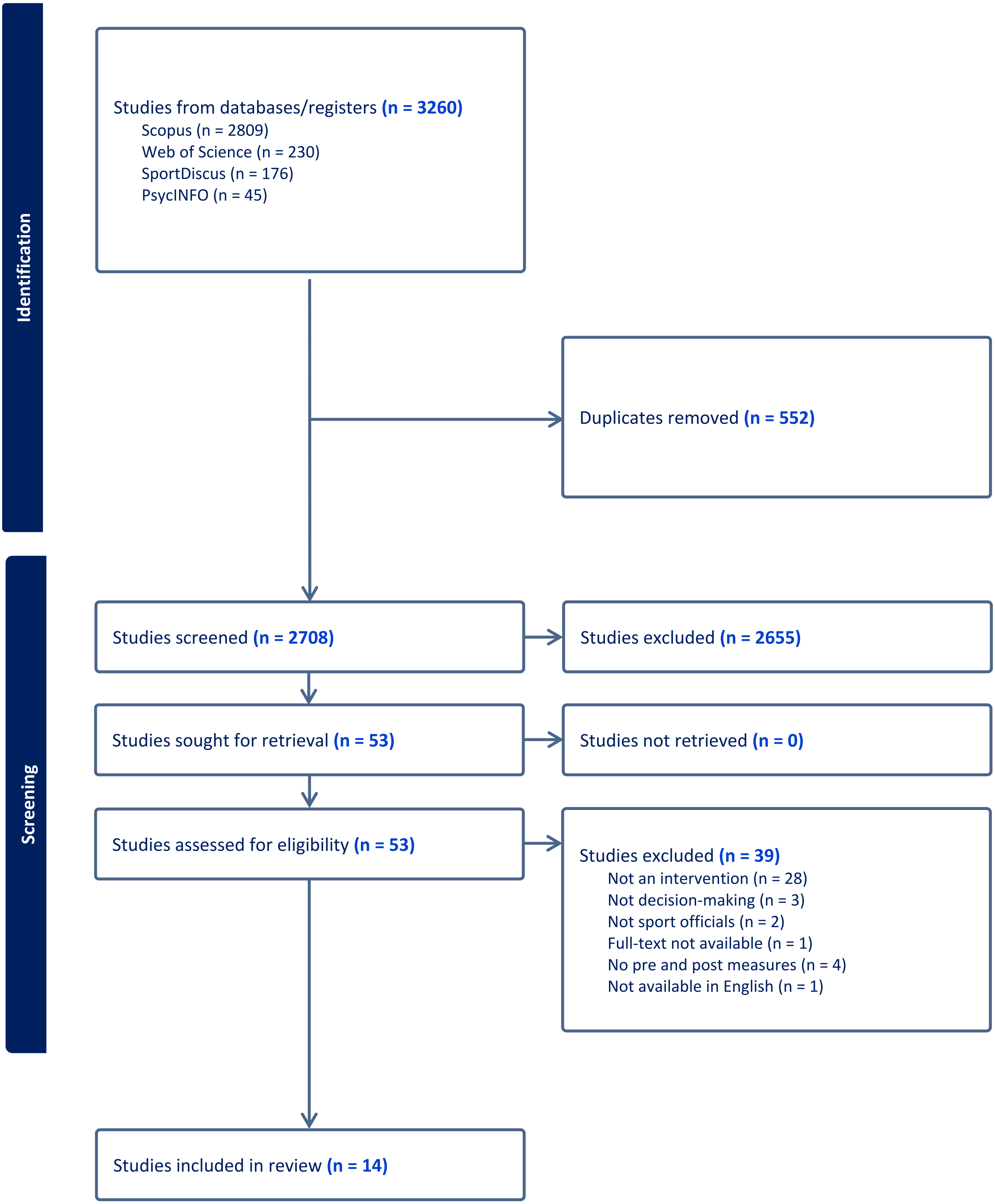

Fig. 1. PRISMA flow chart of search process.

Table 1
Summary of included studies in this review.

Author	Sport	Type of decision	Participants	Intervention overview	Intervention duration & frequency	Training scenarios number (total)	Feedback provided	Testing methods	Key results & conclusions
Armenteros et al. (2018)	Soccer	Offside	8 sub-elite assistant referees	Video training consisting of six random series of 40 ’cases’ per day	3 days	240 videos	N/A	Video-based pre-test, field-based post-test	Experimental group demonstrated continues improvement over time significantly better than the control group
Catteeuw, Gilis, Wagemans, and Helsen (2010)	Soccer	Offside	24 elite assistant referees	Video training and computer simulations	4 weeks, once per week	160 (80 video, 80 computer)	Yes	Video-based pre- and post-test	Video-based training led to an improvement in response accuracy in post-test, whereas control group did not change
Catteeuw, Gilis, Jaspers, et al. (2010)	Soccer	Offside	40 elite assistant referees	Video training, computer simulations	4 weeks, once per week	120 videos or computer animations	Yes	Video-based pre- and post-test	Both video-based training and computer simulation training improved following the training intervention, and no improvement for the control group
Gulec and Yilmaz (2016)	Soccer	Player to play fouls/ penalties	54 referees (some novice, some expert)	Dice-based game where players answered questions (true/ false, multiple choice, or video questions)	Participants able to choose the number of sessions playing the game	N/A	N/A	Multiple choice and video-based pre- and post-test	Game-based learning is better than paper-based.
Kittel et al. (2019a)	Australian Football	Player to play fouls/ penalties	20 novice umpires	Video-based training, video-based training combined with high intensity interval training, control group	8 weeks, once per week	96 videos	Yes	Video-based pre- and post-test	No benefit completing video-based training combined with, or before, high intensity interval training
Kittel et al. (2020)	Australian Football	Player to play fouls/ penalties	32 novice umpires	Video-based training, 360VR video-based training, control group	5 weeks, once per week	125 videos	Yes	Video-based pre- and post-test. Retention test also used	Both intervention groups did not improve over the course of the intervention, but 360VR performed significantly better than control in retention test
Kittel et al. (2023)	Australian Football	Player to play fouls/ penalties	17 novice umpires	Real-time video training, above real-time video training, control group	5 weeks, once per week	60 videos	Yes	Video-based pre- and post-test	Above real-time video-based training is not appropriate for amateur umpires
Larkin, Mesagno, Berry, Spittle, et al. (2018)	Australian Football	Player to play fouls/ penalties	52 novice umpires	Video-based training (no feedback provided), control group	12 weeks, once per week	1040 videos	No	Video-based pre- and post-test. Retention test also used	Decision-making improved for less experienced participants in the retention test
Mascarenhas et al. (2005)	Rugby Union	Player to play fouls/ penalties	56 elite referees (top 20, mid panel, low panel)	Video-based training (from referee perspective)	6 weeks, once per week	25 videos	Yes	Video-based pre- and post-test	Decision-making improved the most for lowest ranked participants in the intervention group
Put et al. (2013)	Soccer	Offside	18 sub-elite referees	Video-based training (from referee perspective) and computer animations as one group, control group	4 sessions, frequency unknown	240 (120 video, 120 computer)	Yes	On-field and off-field offside tests	Training group increased accuracy, and had less flag errors in both the on- and off-field tests

(continued on next page)

Table 1 (*continued*)

Author	Sport	Type of decision	Participants	Intervention overview	Intervention duration & frequency	Training scenarios number (total)	Feedback provided	Testing methods	Key results & conclusions
Put, Wagemans, Pizzera, et al. (2016)	Soccer	Offside	96 elite referees	Increasing speed group (75 %, 100 %, 125 %), Decreasing speed group (125 %, 100 %, 75 %), variable speed group.	2 days, 3 sessions	60 videos	Yes	Video-based pre- and post-test	Decreasing speed group had best training adaptations following intervention
Put, Wagemans, Spitz, et al. (2016)	Soccer	Offside	20 elite referees	Video-based training (from referee perspective) and computer animations as one group, control group	8 weeks, 12 sessions	720 (360 video, 360 computer)	Yes	Video-based pre- and post-test	Web-based training using videos led to a performance increase following the intervention
Schweizer et al. (2011)	Soccer	Player to play fouls/ penalties	53 sub-elite referees	Video-based training (with repetition following feedback), video-based training (no repetition), control group	7 sessions	144 videos	Yes	Video-based pre- and post-test	Both training groups (with and without repetition) improved following the intervention
Van Biemen et al. (2018)	Soccer	Player to play fouls/ penalties	22 elite referees	Blurred footage group, normal footage group	1 session	70 videos	Yes	Video-based pre- and post-test	Blurred training led to significantly greater training effects in decision-making than regular non-blurred footage

Table 2
Quality of reporting of included studies.

Study	Quality Assessment Domains													
	1	2	3	6	7	10	12	15	18	20	22	23	25	27	Total/ 14
Armenteros et al. (2018)	✓	✓	✓	✓	✗	✗	✗	✗	✓	✗	✓	✗	✗	✗	6
Catteeuw et al. (2010b)	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗	✓	✗	✗	✗	8
Catteeuw et al. (2010a)	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗	✓	✗	✗	✗	8
Gulec and Yilmaz (2016)	✓	✓	✗	✓	✗	✗	✗	✗	✗	✗	✓	✓	✗	✗	5
Kittel et al. (2019a)	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✗	✗	10
Kittel et al. (2020)	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✗	✗	10
Kittel et al. (2023)	✓	✓	✓	✓	✗	✓	✗	✗	✓	✓	✓	✓	✗	✗	9
Larkin et al. (2018b)	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✗	✗	10
Mascarenhas et al. (2005)	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✗	✗	✗	9
Put et al. (2013)	✓	✓	✓	✓	✓	✗	✗	✗	✓	✓	✓	✗	✗	✗	8
Put et al. (2016a)	✓	✓	✓	✓	✗	✓	✗	✗	✓	✗	✓	✗	✗	✗	7
Put et al. (2016b)	✓	✓	✓	✓	✗	✓	✗	✗	✓	✓	✓	✓	✗	✗	9
Schweizer et al. (2011)	✓	✓	✓	✓	✓	✗	✗	✗	✓	✗	✓	✓	✗	✗	8
Van Biemen et al. (2018)	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✗	✗	10

compute a composite score for each of the included studies. This approach provides a more conservative estimate of the overall effect by calculating a variance that accounts for the correlation between different study outcomes (Borenstein et al., 2022). Subsequently, a total of 27 individual effects were used for analysis from the following conditions: decision-making training and control. Table 3 displays the overall results for subgroups and comparisons conducted between subgroups. From a cumulative sample of N = 251 participants, an overall moderate significant effect was observed ($g = 0.680$; 95 % CI = 0.381, 0.980; $p < .001$) for decision-making training on performance outcomes relative to control conditions. There was no evidence of publication bias (Kendall's tau, Z = 1.25, p = 0.21; Egger's test, t (7) = 1.01, $p = .34$). Duval and Tweedie's trim-and-fill analysis imputed one effect size to the right of the mean and marginally changed the observed hedge's g value ($g = 0.728$, 95 % CI = 0.439, 1.017). Fig. 2 presents a forest plot of individual study effects and the overall effects for decision-making training on performance outcomes (see Table 4).

Analysis revealed statistical heterogeneity of decision-making training effects on performance was significant ($Q = 25.86$, $p < .001$) and was indicative of substantial heterogeneity ($I^2 = 69.1$ %) (Higgins & Green, 2008), indicating the variability within the included studies could be due to other moderating factors rather than sampling error. Subsequently, additional subgroup analyses were conducted on pre-defined DM intervention variables to locate potential sources of

Table 3
Effects of included studies with 95 % confidence intervals and sub-groups comprising each Decision-making training variable.

Sub-group	*j*	*k*	*g*	95 % CI		*df*	*z*	*p*
				LL	*UL*			
Decision-making training overall	8	27	0.687	0.415	0.959	26	4.950	**<.001**
Expertise level								
Amateur	3	6	0.271	−0.145	0.687	5	1.276	*p* = .202
Semi-professional	5	10	0.890	0.518	1.262	9	4.686	***p* < .001**
Sport								
Australian Football	4	8	0.237	−0.101	0.575	7	1.373	*p* = .170
Rugby Union	1	3	0.901			2	3.105	***p* < .001**
Soccer	4	6	1.051			5	4.365	***p* < .001**
Training tool								
2-D Video	7	13	0.717	0.382	1.052	12	4.200	***p* < .001**
Training Duration (weeks)								
4–6	4	8	0.629	0.321	0.936	7	4.003	***p* < .001**
6+	2	4	0.162	−0.286	0.611	3	0.711	p = .477
Training Frequency (per week)								
1	7	13	0.451	0.207	0.695	12	3.626	***p* < .001**
Decision-making performance outcome								
Accuracy		13	0.652	0.329	0.975	12	3.959	***p* < .001**
Correct flag infringement		4	1.229	0.421	2.037	3	2.980	***p* < .001**
Correct non-flag identification		4	0.370	−0.126	0.866	3	1.463	*p* = .143
Offside identification		4	1.480	1.010	1.950	3	6.170	***p* < .001**

Note: *j* = number of studies; *k* = number of effect sizes; *g* = hedge's g; CI = confidence interval; *LL* = lower limit; *UL* = upper limit; d*f* = degrees of freedom. *Z*-scores and associated *p*-values indicate whether the effects were significantly different from 0.

Table 4
Difference in Hedge's g for included studies with 95 % confidence intervals and sub-groups comprising each moderator variable.

Sub-group	Difference	95 % CI		*Q*	*z*	*p*
		LL	*UL*			
Expertise level						
Amateur - Semi-professional	0.619	0.060	1.177	4.81	2.172	***p* < .05**
Sport						
Rugby Union - AFL	0.664	0.002	1.325	9.00	1.967	***p* < .05**
Soccer -AFL	0.814	0.233	1.394	9.00	2.748	***p* < .001**
Rugby Union - Soccer	−0.150	−0.889	0.589	9.00	−0.398	*p* = .691
Training Duration (weeks)						
6+ - 4-6	−0.466	−1.010	0.077	9.47	−1.681	*p* = 0.093
Decision-making performance outcome						
Flag infringement – Accuracy	0.576	−0.294	1.447	20.63	1.298	*p* = 0.194
Offside – Accuracy	0.827	0.257	1.398	20.63	2.843	***p* < .001**
Non-Flag infringement – Accuracy	−0.282	−0.874	0.309	20.63	−0.935	*p* = 0.350
Flag infringement - Offside	−0.251	−1.186	0.684	20.63	−0.526	*p* = 0.599
Non-flag infringement – Offside	−1.110	−1.793	−0.426	20.63	−3.184	***p* < .001**
Non-flag infringement – Flag Infringement	−0.859	−1.807	0.089	20.63	−1.775	*p* = 0.076

Note: CI = confidence interval; *LL* = lower limit; *UL* = upper limit. *Q* statistic and degrees of freedom were used to test for heterogeneity of effect size variance. *Z*-scores and associated p-values indicate whether effects were significantly different from 0.

variability.

3.4. Decision-making performance outcome

Studies were categorised into seven decision performance outcomes (or markers); Accuracy, Confidence in decision, Decision consistency, Correct Flag infringement identification, Correct Non-flag infringement identification, Offside identification, and Recognition. The Decision consistency and Recognition subgroups contained only 1 effect size, respectively, resulting in exclusion of both subgroups from comparative analyses. Results indicated that decision-making training significantly improved decision making accuracy ($g = 0.65$; 95 % CI = 0.33, 0.98; $p < .001$), Correct Flag infringement identification ($g = 1.23$; 95 % CI = 0.42, 0.2.04; $p < .001$), and Offside identification ($g = 1.48$; 95 % CI = 1.01, 1.95; $p < .001$), but not Correct Non-flag infringement identification ($g = 0.37$; 95 % CI = −0.13, 0.87; $p = 0.08$). The variability between sub-groups was heterogenous ($Q = 20.63$, $p < .001$), indicating that effect sizes significantly varied between sub-groups. Z –tests showed that decision-making training had a significantly larger effect on offside performance outcomes than accuracy measures ($z = 2.843$; 95 % CI = 0.257, 1.398; $p < .001$). Decision-making training was significantly more effective at developing correct identification of offside penalties than non-flag infringements ($z = -3.184$; 95 % CI = −1.793, −0.426; $p < .001$). No further significant differences were found between other decision-making performance measures (Table 2).

3.5. Sport

Three sports were included for sub-group analysis: Australian Football, Rugby Union, and Soccer. Analysis showed that decision-making training significantly improved performance in Soccer ($g = 1.05$; 95 % CI = 0.58, 1.52; $p < .001$) and Rugby Union ($g = 0.90$; 95 % CI = 0.33, 1.47; $p < .001$), however this was not the case in Australian Football ($g = 0.24$; 95 % CI = −0.10, 0.58; $p = .17$). The variability between sub-groups was heterogenous ($Q = 9.00$, $p = .01$), indicating that effect sizes significantly differed (Table 2). Z-tests revealed that decision-making training in Rugby Union ($z = 1.967$; 95 % CI = 0.002, 1.325; $p = .049$) and Soccer ($z = 2.748$; 95 % CI = 0.233, 1.394; $p < .001$) improved performance significantly more than in Australian Football. There were no significant differences between Rugby Union and Soccer ($p = .691$).

Study Name	Statistics for each study				
	Hedges's g	Lower limit	Upper limit	Z-Value	p-Value
Catteeuw et al., 2010a	1.062	0.716	1.409	6.008	0.000
Catteeuw et al., 2010b	0.790	0.210	1.369	2.670	0.008
Kittel et al., 2023	0.132	-0.641	0.905	0.336	0.737
Kittel et al., 2019	0.150	-0.556	0.857	0.418	0.676
Kittel et al., 2020	0.253	-0.433	0.939	0.722	0.470
Larkin et al., 2018	0.171	-0.409	0.750	0.577	0.564
Maascarenhas et al., 2005	0.542	0.241	0.843	3.527	0.000
Put et al., 2013	1.011	0.613	1.409	4.981	0.000
Schweizer et al., 2011	1.528	0.994	2.062	5.613	0.000
Pooled	0.680	0.381	0.980	4.447	0.000
Prediction Interval	0.680	-0.263	1.623		

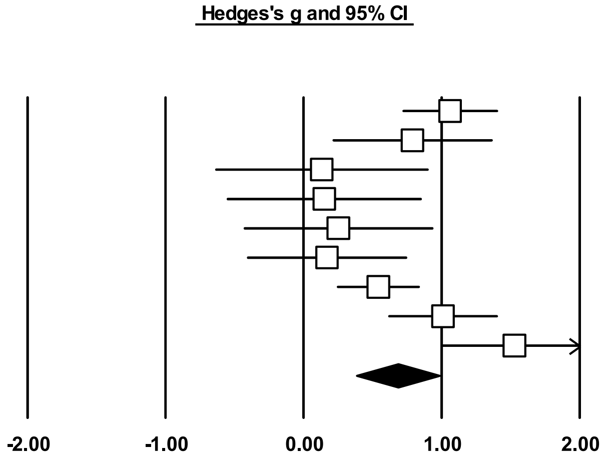

Fig. 2. Forest plot showing the effect of decision-making training in k = 9 studies, $g = 0.680$ (95 % CI = 0.381, 0.980). The pooled hedge's g is depicted by the diamond. Squares denote the effect of decision-making training for each individual study and the horizontal lines denote the 95 % CI for each individual study.

3.6. Expertise level

Studies were categorised into three expertise levels: Amateur, Sub-elite, and Elite. The Elite sub-group contained only 1 effect resulting in exclusion from comparative analysis. Results indicated that decision-making training positively improved performance across all levels of expertise (Amateur; $g = 0.27$; 95 % CI = −0.15, 0.69; $p < .20$) (Sub-elite; $g = 0.89$; 95 % CI = 0.52, 1.26; $p < .001$) and significantly so for Sub-elite performers. Between-group variability was not significantly heterogeneous ($Q = 4.81$, $p = .09$), suggesting that effect sizes did not vary between groups (Table 2).

3.7. Training tool

Decision-making training tools were categorised into four sub-groups; 360°VR, Combined computer and video, Computer only, and Video only. The 360°VR, Combined computer and video, Computer only sub-groups contained only 1–2 effects and, therefore, were excluded from comparative analysis. Decision-making training using video only tools were found to have a significant positive effect on performance ($g = 0.72$; 95 % CI = 0.38, 1.05; $p < .001$).

3.8. Training duration

Training interventions less than one week contributed 2 effect sizes, and therefore, were excluded from further analyses. Decision-making training was found to have a significant positive effect for intervention durations of 4–6 weeks ($g = 0.63$; 95 % CI = 0.32, 0.94; $p < .001$), which was not the case for interventions of 6 weeks or more ($g = 0.16$; 95 % CI = −0.29, 0.61; $p = .48$). The variability between sub-groups was heterogenous ($Q = 9.47$, $p = .02$), indicating effect sizes significantly varied between sub-groups. However, z-tests showed no significant differences between interventions of 4–6 weeks and 6 weeks or more ($p = .093$).

3.9. Training frequency

Weekly training frequencies of 2–3 and 4–5 sessions per week contributed only 2 effect sizes, respectively and, therefore, were removed from further analysis. Subsequently only one subgroup remained for analysis. Results showed decision-making training has a significant impact on performance when implemented for one session per week ($g = 0.45$; 95 % CI = 0.21, 0.69; $p < .001$).

4. Discussion

This review synthesised current understanding around the effectiveness of decision-making training with team sport officials, with 14 studies meeting the inclusion criteria, and data able to be analysed for nine of these studies. Decision-making training had an overall significant positive effect on decision-making performance. Sub-group analyses indicated that decision-making training significantly improved the accuracy of objective decision-making, such as flag infringements, and offside identification. Moderate improvements were observed in subjective decision-making outcomes, such as foul/penalty identification. In addition, analysis revealed several moderator variables that impact the effect of decision-making training on decision-making performance, such as type of sport, expertise level, and training frequency/duration. However, there are some considerations for the reader when reviewing the results. For instance, this is an emerging research area with 14 studies included (and nine able to be analysed), and there was high heterogeneity (i.e., there was a significant amount of variability between the methodology of included studies).

This meta-analysis identified that decision-making performance can be significantly improved through off-field training methods (e.g., 2D video only) in sport officials. Indeed, recent discussions in the literature critique existing decision-making training for sports official for being decontextualised, suggesting it may not appropriately reflect the demands of competition (O'Brien & Rynne, 2021). Our findings indicate that although more immersive forms of training may be preferred, decision-making training that are considered less representative (i.e., off-field video only) is not completely decontextualised, and may provide learners with an adequate level of relevant information to successfully facilitate the development of decision-making skills. It is important to note that sub-group analyses indicated that such off-field methods of training were more effective for types of decisions that are more objective in nature (e.g., offside penalties), and do not require interpretation of the laws for game management (Raab et al., 2021), compared with more subjective decision-making (e.g., identification of fouls), which involve balancing how a decision contributes to the fairness, safety and entertainment of the competition (Russell et al., 2019). However, it is important to note that when analysing by sub-group, there were a smaller number of effect sizes to make inferences on the data, and more research should be done to build on the current review. Taken together, these findings suggest more decontextualised forms of decision-making training could enhance performance in scenarios where decisions rely on fixed criteria (e.g., offside). Whereas, for subjective decision-making scenarios, training tools need to be progressed beyond off field methods to successfully replicate the complexity of in-game scenarios. A limited number of studies supported this idea in the present review, with emerging evidence utilising more immersive decision-making training tools like 360°VR, revealing significant improvements in subjective decision-making outcomes (e.g., identification of foul play) (Kittel et al., 2019; Kittel et al., 2020; Kittel et al., 2023). Recently, studies have emerged highlighting the use of technologies such as 360°VR for decision-making in sport officials, that although they did not meet the inclusion criteria for this review and subsequent meta-analysis, yield important insights. For example, Boyer et al. (2023) interviewed officials from different sports (Soccer, Rugby, Handball) on their use of 360°VR, highlighting it as a useful tool for reflective practice. Other research has explored virtual environments (i.e., VR) in Soccer referees, reporting that decision-making and visual search behaviours are similar in VR to the real world (van Biemen et al., 2023). While there is emerging evidence for the use of technologies such as 360°VR and VR for sports officials (Kittel et al., 2024), more research is required to explore the effectiveness of these technologies in intervention-based research.

When considering type of sport, decision-making training was found to be more effective for improving decision-making for player-to-player fouls in Rugby Union (Mascarenhas et al., 2005) and Soccer (Schweizer et al., 2011), relative to Australian Football (Kittel et al., 2019a, 2020, 2023; Larkin, Mesagno, Berry, Spittle, et al., 2018). An explanation for this may the complex decision-making processes to make certain decisions in Australian Football, such as the 'holding the ball' interpretation unique to this sport (Larkin, Mesagno, Berry, & Spittle, 2018). Further, this is supported by the top two attributes for effective Australian Football umpiring performance being knowledge of the laws, followed by decision-making (Kittel et al., 2019b). This highlights that in sports like Australian Football, where decisions often hinge on nuanced rules (e.g., "holding the ball"), integrating knowledge of the laws within decision-making training becomes essential. Training programs should aim to combine law-based education with decision-making practice to ensure that officials not only understand the rules but develop the skills to apply them accurately in pressure situations (Cunningham et al., 2022). Further research is recommended to understand the impact decision-making training may have on other sports that may have different officiating demands to the three sports included here.

It is also interesting to note that the Video Assistant Referee (VAR) was introduced to Soccer in 2017–2018 (Lago-Peñas et al., 2019), and each of the studies in Soccer in this review were published around this time, or earlier. It is surprising that no studies exploring the effectiveness of decision-making have been published since then, given a key premise of this training is to hasten expertise (Helsen et al., 2019, pp. 250–266). It is, therefore, surprising there are no studies since determining the effectiveness of decision-making training in Soccer referees for two key reasons. Firstly, even in competitions where VAR is available, not every decision is able to be reviewed, and an accurate initial decision should be made regardless (Spitz et al., 2021). Secondly, most competitions in Soccer (or any sport for that matter) do not have access to technology such as the VAR at levels below the elite. Researchers in this domain need to ensure they are exploring ways to develop decision-making skill in emerging officials, where they do not have the ability to use/rely on video technology in competitions.

The most common training tool was using video-based methods, which was found to have a positive effect on decision-making skill development. This was the only training tool that had enough effect sizes to be able to be used in the meta-analysis component of this review. This finding for sports officials aligns with the literature on video-based training in athletes that has reported that it is effective tool for the development of decision-making skill (Larkin et al., 2015). There was a mixture of different video-based approaches used in the studies identified, including match broadcast (Larkin, Mesagno, Berry, Spittle, et al., 2018), first-person training drills (Put et al., 2013), computer animations (Put, Wagemans, Spitz, et al., 2016), Put, Wagemans, Spitz, et al., 2016nd 360°VR of training-based situations (Kittel et al., 2020). Also, other studies use certain video manipulations to demonstrate varied training effects to that of normal video presentation, including blurring images (van Biemen et al., 2018) and speed of video playback (Put, Wagemans, Pizzera, et al., 2016). There were few studies that utilised other technologies such as 360°VR and computer animations, therefore limiting comparisons with these technologies. Although some of these video-based methods, such as match broadcast video, computer animations, and even first-person training drills may appear to have lower representativeness, they still appear to be effective as decision-making training modalities with sports officials. As highlighted earlier, there was significant heterogeneity in the included studies, as evident when considering the different training tools used across the studies included, with some training tools being only used in one particular study. Therefore, more research needs to be done with specific training tools across multiple skill levels and support to strengthen our knowledge of decision-making training of sport officials. Further research exploring technologies that appear to have higher levels of representativeness may help to determine whether and how important this representativeness may be in decision-making training of sports officials.

Interestingly, the most effective duration of decision-making training was for 4–6 weeks, yet not as effective for 6+ weeks, although there was no difference between the durations. In other psychological training methodologies, such as mental practice and imagery training in sport, systematic reviews and meta-analyses have indicated that programs of 1–6 weeks are effective, but with the magnitude of effect decreasing with increases in program length (Lindsay et al., 2023; Toth et al., 2020). This led Lindsay et al. (2023) to conclude that programs between 1 and 6 weeks were most effective. For decision-making training, this could be because there is a diminishing return of practice (Spittle, 2021) so that additional training of a similar type begins to produce smaller benefits, indicating that the ideal initial program length for an intervention of this type may be between 4 and 6 weeks.

There are some limitations to the current meta-analysis that should be considered when interpreting the results. Firstly, although 14 studies were included in this review, only nine were included in the meta-analysis as having data able to be used for analysis. Therefore, the reader should exercise caution when interpreting the results as although this is a growing research area, the limited number of studies included may contribute to the overall effect being sensitive to results of individual studies. Another limitation is the over-representation of certain sports, specifically Soccer, Australian Football and Rugby Union. More research in other interactor sports such as basketball, hockey, handball and American football would benefit the testing of modes of decision-making training. It is surprising there were no studies from these sports meeting the inclusion criteria, given these are commonly researched in the sport officiating domain (Hancock et al., 2021). Research in non-interactor sports (MacMahon et al., 2014) should be a future focus in this area as this review found no studies aimed at improving judges' evaluations of human performance, such as in aesthetic sports (i.e., gymnastics, figure skating) or racquet sports (e.g., tennis chair umpire). Further understanding of the influence of expertise level on the effectiveness of training and other individual differences in officials are needed to decipher at what expertise levels training might be more effective. How training duration influences training effectiveness needs to be further discerned, particularly considering the inconsistency in duration times administered in the available training studies. With 4–6-week interventions showing more significant improvement, understanding if there are learning ceilings of decision-making training modes and how more complex layering of training modes and focus possibly alter learning outcomes. With video training (in its different forms) the primary decision-making mode in empirical-studies, future research might test accumulated or scaffolded effects of blending different formats of video viewing together in training programs (and retention of knowledge acquired by video training to actual performance). One such example could be exposing sports officials to video with increasing perceptual difficulty, or to use a decision priming strategy (e.g., knowledge vs. infraction) as potentially other variations of training format (MacMahon et al., 2007).

A primary future research interest in the applied area of decision-making training should be to investigate the transfer of skills to the actual officiating environment (almost non-existent across most reviewed studies). Despite the efficacy of video-based training observed in the present review, this approach may not fully capture the dynamic nature of in-game decision-making, highlighted by the reduced effectiveness of such approaches for developing more subjective outcomes. While videos provide opportunities for reviewing and reflecting on specific scenarios, they tend to lack the dynamic, unpredictable elements of live gameplay. Subsequently, officials may struggle to transfer the skills learning in training to on-field performance environments. Immersive and representative video-based learning, through first person (mobile) and fixed video capture remains to have significant potential to address these issues as a decision-making training learning mechanism where transfer of learning might be more easily studied. Often decision-making training studies can become a 'one-off' investigation between researchers and sport organisations that are not extended into meaningful implementation over longer periods to monitor and assess performance changes. The uptake of applied sport science knowledge to the target training sport field has become a growing question more recently (Fullagar et al., 2019), giving credence for more implementation research approaches in the future.

The present systematic review and meta-analysis provides important insights into the current state of knowledge regarding the effectiveness of decision-making training for sports officials. Overall, results indicated that decision-making training significantly improves performance, demonstrating significant improvements in objective decision-making outcomes. However, more moderate improvements were observed for subjective outcomes, highlighting the challenges of replicating complex in-game scenarios through off-field, video-based training. Video-based training emerged as the most common training tool, however, further investigation is needed to explore the efficacy of more immersive approaches, such as on-field simulations or virtual reality. However, when interpreting these findings, it is important to consider that only officials from three sports were identified within this review (Soccer, Australian Football, Rugby Union), limiting the generalisability of these results to other sports, highlighting a future research agenda for decision-making training in other sports. In addition, there was high heterogeneity, and a small number of studies included in the final meta-analysis, which may increase the sensitivity to the overall effect being influenced by individual studies. A gap in the literature was identified regarding the lack of examination into training that integrates both rule knowledge and decision-making skills, which are proposed to be of particular importance in sports with nuanced rules, like Australian Football. It is recommended that future studies examine the transfer of decision-making training to competitive environments.

CRediT authorship contribution statement

Aden Kittel: Writing – review & editing, Writing – original draft, Visualization, Resources, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Riki Lindsay:** Writing – review & editing, Writing – original draft, Software, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Paul Larkin:** Writing – review & editing, Writing – original draft, Formal analysis. **Michael Spittle:** Writing – review & editing, Writing – original draft, Formal analysis. **Ian Cunningham:** Writing – review & editing, Writing – original draft, Investigation, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Armenteros, M., Benítez, A. J., Flores, R., Sillero-Quintana, M., Cid, M. S., & Simón, J. (2018). The training of soccer assistant referees beyond on-field experience: The use of the interactive video test. *International Journal of Computer Science, 17*(2), 163–174. *in Sport*.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2022). Comprehensive meta-analysis version 4 (Englewood, NJ: Biostat, Inc). Available at: www.Meta-Analysis.com.

Boyer, S., Rochat, N., & Rix-Lièvre, G. (2023). Uses of 360° video in referees' reflectivity training: Possibilities and limitations. *Frontiers in Psychology, 14*, Article 1068396.

Brand, R., Schweizer, G., & Plessner, H. (2009). Conceptual considerations about the development of a decision-making training method for expert soccer referees. *Perspectives on cognition and action in sport*, 181–190.

Catteeuw, P., Gilis, B., Jaspers, A., Wagemans, J., & Helsen, W. (2010a). Training of perceptual-cognitive skills in offside decision making. *Journal of Sport & Exercise Psychology, 32*(6), 845–861.

Catteeuw, P., Gilis, B., Wagemans, J., & Helsen, W. (2010b). Perceptual-cognitive skills in offside decision making: Expertise and training effects. *Journal of Sport & Exercise Psychology, 32*(6), 828–844.

Cunningham, I., Mergler, J., & Wattie, N. (2022). Training and development in sport officials: A systematic review. *Scandinavian Journal of Medicine & Science in Sports, 32*(4), 654–671.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health, 52*(6), 377–384.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj, 315*(7109), 629–634.

Fullagar, H. H., McCall, A., Impellizzeri, F. M., Favero, T., & Coutts, A. J. (2019). The translation of sport science research to the field: A current opinion and overview on the perceptions of practitioners, researchers and coaches. *Sports Medicine, 49*, 1817–1824.

Gulec, U., & Yilmaz, M. (2016). A serious game for improving the decision making skills and knowledge levels of Turkish football referees according to the laws of the game. *SpringerPlus, 5*, 1–10.

Hancock, D. J., Bennett, S., Roaten, H., Chapman, K., & Stanley, C. (2021). An analysis of literature on sport officiating research. *Research Quarterly for Exercise & Sport, 92*(4), 607–617.

Helsen, W. F., MacMahon, C., & Spitz, J. (2019). Decision making in match officials and judges. *Anticipation and decision making in sport* (pp. 250–266).

Higgins, J. P., & Green, S. (Eds.). (2008). *Cochrane handbook for systematic reviews of interventions*. http://www.cochrane-handbook.org.

Kittel, A., Cunningham, I., Larkin, P., Hawkey, M., & Rix-Lievre, G. (2021). Decision-making training in sporting officials: Past, present and future. *Psychology of Sport and Exercise, 56*, Article 102003.

Kittel, A., Elsworthy, N., & Spittle, M. (2019a). Incorporating perceptual decision-making training into high-intensity interval training for Australian football umpires. *Journal of Sports Sciences, 37*(1), 29–35.

Kittel, A., Elsworthy, N., & Spittle, M. (2023). The effectiveness of above real time training for developing decision-making accuracy in australian football umpires. *Research Quarterly for Exercise & Sport, 94*(1), 64–72.

Kittel, A., Larkin, P., Elsworthy, N., Lindsay, R., & Spittle, M. (2020). Effectiveness of 360 virtual reality and match broadcast video to improve decision-making skill. *Science and Medicine in Football, 4*(4), 255–262.

Kittel, A., Larkin, P., Elsworthy, N., & Spittle, M. (2019b). Identification of key performance characteristics of elite Australian football umpires. *International Journal of Sports Science & Coaching, 14*(4), 490–497.

Kittel, A., Larkin, P., Elsworthy, N., & Spittle, M. (2019c). Video-based testing in sporting officials: A systematic review. *Psychology of Sport and Exercise, 43*, 261–270.

Kittel, A., Lindsay, R., Le Noury, P., & Wilkins, L. (2024). The use of extended reality technologies in sport perceptual-cognitive skill research: A systematic scoping review. *Sports Medicine-Open, 10*(1), 128.

Lago-Peñas, C., Rey, E., & Kalén, A. (2019). How does Video Assistant Referee (VAR) modify the game in elite soccer? *International Journal of Performance Analysis in Sport, 19*(4), 646–653.

Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, 363–374.

Larkin, P., Berry, J., Dawson, B., & Lay, B. (2011). Perceptual and decision-making skills of Australian football umpires. *International Journal of Performance Analysis in Sport, 11*(3), 427–437.

Larkin, P., Mesagno, C., Berry, J., & Spittle, M. (2018). Exploration of the perceptual-cognitive processes that contribute to in-game decision-making of Australian football umpires. *International Journal of Sport and Exercise Psychology, 16*(2), 112–124.

Larkin, P., Mesagno, C., Berry, J., Spittle, M., & Harvey, J. (2018). Video-based training to improve perceptual-cognitive decision-making performance of Australian football umpires. *Journal of Sports Sciences, 36*(3), 239–246.

Larkin, P., Mesagno, C., Spittle, M., & Berry, J. (2015). An evaluation of video-based training programs for perceptual-cognitive skill development. A systematic review of current sport-based knowledge. *International Journal of Sport Psychology, 46*(6), 555–586.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Annals of Internal Medicine, 151*(4). W-65-W-94.

Lindsay, R. S., Larkin, P., Kittel, A., & Spittle, M. (2023). Mental imagery training programs for developing sport-specific motor skills: A systematic review and meta-analysis. *Physical Education and Sport Pedagogy, 28*(4), 444–465.

MacMahon, C., Mascarenhas, D., Plessner, H., Pizzera, A., Oudejans, R., & Raab, M. (2014). *Sports officials and officiating: Science and practice*. Routledge.

MacMahon, C., Starkes, J., & Deakin, J. (2007). Referee decision making in a video-based infraction detection task: Application and training considerations. *International Journal of Sports Science & Coaching, 2*(3), 257–265.

Mascarenhas, D. R., Collins, D., Mortimer, P. W., & Morris, B. (2005). Training accurate and coherent decision making in rugby union referees. *The Sport Psychologist, 19*(2), 131–147.

McEwan, G. P., Unnithan, V. B., Easton, C., & Arthur, R. (2023). Training practices and perceptions of soccer officials: Insights from the referee training activity questionnaire. *International Journal of Sports Science & Coaching, 18*(4), 1173–1189.

McEwan, G. P., Unnithan, V. B., Easton, C., Glover, A. J., & Arthur, R. (2024). Decision-making accuracy of soccer referees in relation to markers of internal and external load. *European Journal of Sport Science, 24*(6), 659–669.

Morris, G., & O'Connor, D. (2017). Key attributes of expert NRL referees. *Journal of Sports Sciences, 35*(9), 852–857.

O'Brien, K. A., & Rynne, S. B. (2021). Seen but not heard: Missing the mark in conceptualizing high performance officiating. *Sport in Society, 24*(7), 1110–1121.

O'Connor, D., Larkin, P., & Williams, A. M. (2017). What learning environments help improve decision-making? *Physical Education and Sport Pedagogy, 22*(6), 647–660.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery, 88*, Article 105906.

Pinder, R. A., Davids, K., Renshaw, I., & Araújo, D. (2011). Representative learning design and functionality of research and practice in sport. *Journal of Sport & Exercise Psychology, 33*(1), 146–155.

Put, K., Wagemans, J., Jaspers, A., & Helsen, W. F. (2013). Web-based training improves on-field offside decision-making performance. *Psychology of Sport and Exercise, 14*(4), 577–585.

Put, K., Wagemans, J., Pizzera, A., Williams, A. M., Spitz, J., Savelsbergh, G. J., & Helsen, W. F. (2016). Faster, slower or real time? Perceptual-Cognitive skills training with variable video speeds. *Psychology of Sport and Exercise, 25*, 27–35.

Put, K., Wagemans, J., Spitz, J., Williams, A. M., & Helsen, W. F. (2016). Using web-based training to enhance perceptual-cognitive skills in complex dynamic offside events. *Journal of Sports Sciences, 34*(2), 181–189.

Raab, M., Avugos, S., Bar-Eli, M., & MacMahon, C. (2021). The referee's challenge: A threshold process model for decision making in sport games. *International Review of Sport and Exercise Psychology, 14*(1), 208–228.

Russell, S., Renshaw, I., & Davids, K. (2019). How interacting constraints shape emergent decision-making of national-level football referees. *Qualitative Research in Sport, Exercise and Health, 11*(4), 573–588.

Schweizer, G., Plessner, H., Kahlert, D., & Brand, R. (2011). A video-based training method for improving soccer referees' intuitive decision-making skills. *Journal of Applied Sport Psychology, 23*(4), 429–442.

Spittle, M. (2021). *Motor learning and skill acquisition: Applications for physical education and sport*. Bloomsbury Publishing.

Spitz, J., Wagemans, J., Memmert, D., Williams, A. M., & Helsen, W. F. (2021). Video assistant referees (VAR): The impact of technology on decision making in association football referees. *Journal of Sports Sciences, 39*(2), 147–153.

Thurlow, F., Huynh, M., Townshend, A., McLaren, S. J., James, L. P., Taylor, J. M., Weston, M., & Weakley, J. (2024). The effects of repeated-sprint training on physical fitness and physiological adaptation in athletes: A systematic review and meta-analysis. *Sports Medicine, 54*(4), 953–974.

Toth, A. J., McNeill, E., Hayes, K., Moran, A. P., & Campbell, M. (2020). Does mental practice still enhance performance? A 24 Year follow-up and meta-analytic replication and extension. *Psychology of Sport and Exercise, 48*, Article 101672.

Van Biemen, T., Koedijker, J., Renden, P. G., & Mann, D. L. (2018). The effect of blurred perceptual training on the decision making of skilled football referees. *Frontiers in Psychology, 9*, 1803.

van Biemen, T., Müller, D., & Mann, D. L. (2023). Virtual reality as a representative training environment for football referees. *Human Movement Science, 89*, Article 103091.

Weakley, J., Cowley, N., Schoenfeld, B. J., Read, D. B., Timmins, R. G., García-Ramos, A., & McGuckian, T. B. (2023). The effect of feedback on resistance training performance and adaptations: A systematic review and meta-analysis. *Sports Medicine, 53*(9), 1789–1803.

Williams, A. M., & Ford, P. R. (2013). 'Game intelligence': Anticipation and decision making. In *Science and soccer* (pp. 117–133). Routledge.