Cogent Social Sciences

ISSN: 2331-1886 (Online) Journal homepage: www.tandfonline.com/journals/oass20

# National bias in international sports judging: a scoping review

Patrick Alexander Braeunig

**To cite this article:** Patrick Alexander Braeunig (2025) National bias in international sports judging: a scoping review, Cogent Social Sciences, 11:1, 2594841, DOI: 10.1080/23311886.2025.2594841

**To link to this article:** https://doi.org/10.1080/23311886.2025.2594841

COGENT SOCIAL SCIENCES
2025, VOL. 11, NO. 1, 2594841
https://doi.org/10.1080/23311886.2025.2594841

SPORT | REVIEW ARTICLE

OPEN ACCESS Check for updates

# National bias in international sports judging: a scoping review

Patrick Alexander Braeunig

Sport Science Institute, Saarland University, Saarbrücken, Germany

**ABSTRACT**

The subjective nature of performance evaluation in aesthetic sports competitions makes it vulnerable to bias. Among these, national bias—where judges' scores are influenced by their own and the athlete's nationality—poses a significant threat to the integrity of competition outcomes. Existing literature is fragmented across sports, types of bias, theoretical explanations, and methodological approaches, and lacks an overarching synthesis. To address this gap, the present scoping review—conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2000 guidelines—synthesizes existing literature across four dimensions: (1) forms of national bias, (2) underlying mechanisms, (3) degree of intentionality, and (4) proposed mitigation strategies. Contributions addressing at least one of these dimensions were included regardless of type or quality.The review identifies national bias as a multifaceted phenomenon that manifests in several forms, including favoritism toward compatriots, penalization of their competitors, vote trading among judges, and reactive scoring based on perceived or expected national bias from colleagues. Vote trading reflects intentional manipulation, whereas other forms may arise unintentionally through cognitive or social mechanisms, complicating their detection and regulation. Existing countermeasures demonstrate limited effectiveness and often entail trade-offs, underscoring the need for more context-sensitive and robust interventions.

**ARTICLE HISTORY**
Received 8 May 2025
Revised 13 August 2025
Accepted 20 November 2025

**KEYWORDS**
Nationalism; patriotism; favoritism; corruption; scoring; sports competition

**SUBJECTS**
Sport Psychology; Sociology of Sport; Sport and Politics

## 1. Introduction

In 'aesthetic sports' (McFee, 2013, p. 2)—which comprise approximately 30% of Olympic events (Osório, 2020)—judges play a central role in determining outcomes (Klein et al., 2014; McFee, 2013). Despite their extensive training and expertise (Landers, 1970; Russell, 2001; Sala et al., 2007; Soler, 2021; Ste-Marie, 1999; Weekley & Gier, 1989; Whissell et al., 1993; Wolframm, 2023; Yang, 2006), even the elite judges at such events are not immune to cognitive biases. These biases stem from multiple influences on the judges' assessments, including the judges' and athletes' nationalities (Graf, 2010; Plessner & Haar, 2006), which introduce 'national bias' (Heiniger & Mercier, 2019, p. 1). The influence of national and political affiliations on judging in sports competitions has been recognized since the emergence of political blocs during the Cold War era (Bailie, 1965; Criley, 1972, 1976; Wettstone, 1968), when the Olympic Games and international championships served as arenas where nations competed for global prestige, legitimacy, and national solidarity (Kestnbaum, 2003; Sala et al., 2007; Seltzer & Glass, 1991; Tang, 2013). Since then, concerns about national bias have persisted, with high-profile controversies such as former gymnast Leonid Arkayev's allegations of a 'mafia of judges' (Gymmedia International, 2004) and the suspension of the Soviet Union's judging corps for evident national favoritism (Brennan, 1996). Unlike many other forms of bias, national bias is explicitly acknowledged in the official regulations of several aesthetic sports disciplines (Fédération Équestre Internationale (FEI), 2023; Fédération Internationale de Ski (FIS), 2020, 2025a; International Skating Union (ISU), 2024) and is widely recognized as one of the most significant sources of distortion in judges' evaluations. As Whissell et al. (1993, p. 355) aptly state: 'the most important focus for fairness should be national bias. Other biases may well be

**CONTACT** Patrick Alexander Braeunig patrick.braeunig@uni-saarland.de Sport Science Institute, Saarland University, Campus Geb. B8 1, Raum 0.13, Saarbrücken, D-66123, Germany

randomized from one competition to another whereas national bias will have more systematic effects'. Due to its persistent and distorting influence, national bias poses a serious threat to the integrity of the judges' assessments and competition outcomes (Daumann, 2015; Dosseville et al., 2014). Its presence may not only diminish the entertainment value and public appeal of competitions (Chien et al., 2016; Emrich & Pierdzioch, 2015) but may also damage the reputation and perceived integrity of national federations and their presidential leadership (Goodwin, 2004)—especially in cases where these entities are directly responsible for nominating judges for international events (Zitzewitz, 2006). Furthermore, national bias may erode trust in the impartiality of other judges and undermine the integrity of international federations as regulatory authorities, thereby potentially affecting a wide array of stakeholders, including sponsors (Chien et al., 2016).

Given its prevalence and implications, national bias has become the most extensively studied form of bias in international sports judging. However, the literature remains fragmented and lacks a cohesive synthesis. Research has primarily focused on the 'patriotic bias' (Sala et al., 2007, p. 18) form of national bias, where judges benefit their own compatriots with inflated scores (Heiniger & Mercier, 2019). Less studied are other forms, including 'competitor bias' (Braeunig, 2024, p. 254), where judges provide deflated scores to competitors of their own compatriots, and 'political […] bias' (Ball, 1973, p. 65), where judges favor athletes from politically allied nations (positive political bias) or penalize those from adversarial nations (negative political bias) (Ball, 1973; Sala et al., 2007; Zitzewitz, 2006). Unlike patriotic or competitor bias, political bias thus reflects broader geopolitical alignments rather than direct national affiliation or competitive relationships. A fourth documented form of national bias is 'Vote trading' (Sandberg, 2018, p. 2133; Zitzewitz, 2006, p. 69; 2014, p. 4), wherein judges collude by reciprocating favorable scores across events to advantage certain athletes from their own countries (Sandberg, 2018; Zitzewitz, 2006, 2014). Additionally, national bias may also occur in two types of indirect 'national bias' (Sandberg, 2018, p. 2146). The first 'compensating bias' (Zitzewitz, 2006, p. 75), involves judges attempting to counterbalance each other's patriotic bias (Zitzewitz, 2006), while the second—henceforth called *affirming bias*—refers to judges reinforcing each other's patriotic bias (Bouwens et al., 2022; Krumer et al., 2022; Sandberg, 2018; Zitzewitz, 2006).

The different forms of national bias have been studied across various sports, including figure skating (Sala et al., 2007; Zitzewitz, 2006, 2014), ski jumping (Krumer et al., 2022; Lyngstad et al., 2020), artistic, aerobic, and rhythmic gymnastics (Heiniger & Mercier, 2019, 2021), dressage (Sandberg, 2018) and diving (Emerson et al., 2009). Additionally, the suspected mechanisms underlying national bias are manifold and partially related to socialization mechanisms or rules that are characteristic of only certain sports (Lyngstad et al., 2020; Zitzewitz, 2006, 2014). They thus appear scattered through research, ranging from intentional score manipulation (Lyngstad et al., 2020; Zitzewitz, 2006) to unconscious cognitive processes (Lyngstad et al., 2020; Ste-Marie, 1996; Ste-Marie & Lee, 1991). Furthermore, the degree of intentionality remains ambiguous, complicating efforts to implement effective countermeasures. Several strategies have been proposed to mitigate national bias, including expanding judging panels (Soler, 2021; Zitzewitz, 2014), anonymizing scores (Fang & Ho, 2024; Zitzewitz, 2014), and applying mathematical models to adjust scores and rankings for estimated bias (Ansorge & Scheer, 1988; Osório, 2017; Roetzheim & Muzyczko, 1986). However, each of these approaches faces limitations that are discussed across multiple contributions: expanding panels faces logistical challenges (Myers et al., 2006) and raises the likelihood of co-national judging (Heiniger & Mercier, 2019; Soler, 2021), anonymization may obscure bias (Zitzewitz, 2014) or encourage strategic manipulation (Lee, 2008; Osório, 2017; Truchon, 2004), and mathematical corrections struggle with the variability of bias between scoring cases (Campbell & Galbraith, 1996; Heiniger & Mercier, 2019, 2021; Zitzewitz, 2006) as well as potential strategic adaptations among the judges (Scheer & Ansorge, 1987).

In summary, research on national bias in sports judging encompasses a wide range of disciplines, forms of manifestation, theoretical frameworks, and proposed countermeasures. However, findings remain dispersed across numerous studies, contributing to a fragmented body of literature. Additionally, while early research has been criticized for statistical limitations (Emerson et al., 2009; Heiniger & Mercier, 2019; Sala et al., 2007), more recent studies employ advanced methodologies, such as mixed-effects regression analyses, which account for confounding variables like judge leniency (Emerson et al., 2009; Emerson & Meredith, 2011; Heiniger & Mercier, 2019; Krumer et al., 2022). Despite these methodological advancements, the field lacks a comprehensive synthesis to establish a cohesive theoretical foundation and integrate existing findings systematically.

To address this issue, the present study undertakes a scoping review of national bias in international sports judging. This review is guided by the central research question: *What insights does the existing body of literature provide regarding national bias in international sports judging?* The review is structured around four specific sub-questions, each targeting a key dimension of national bias research to provide a comprehensive synthesis of the field:

1. What forms of national bias have been studied in the context of international sports judging?
2. What mechanisms potentially contributing to national bias are identified and discussed in the extant literature?
3. To what extent does the literature suggest that national bias is exhibited intentionally or unintentionally in international sports judging?
4. What mitigation strategies are proposed in the literature to reduce or eliminate national bias, and what limitations or unintended consequences are associated with these measures?

Section 2 describes the methodological approach of the scoping review. Section 3 presents the findings, structured according to the four sub-questions: forms of national bias (3.1), underlying mechanisms (3.2), intentionality (3.3), and proposed countermeasures (3.4). Section 4 offers a critical discussion of the findings and limitations of the scoping review, and outlines directions for future research.

## 2. Methods

To address the research questions and provide a comprehensive synthesis of the existing literature on national bias in international sports judging, a scoping review was conducted following the PRISMA-2000 guidelines (Page et al., 2021) and the PRISMA extension for scoping reviews (Tricco et al., 2018). A systematic search of electronic databases, including *Web of Science*, *PubMed*, *PsycInfo*, *PsycArticles*, and *ScienceDirect*, was performed. Studies were eligible for inclusion if they were written in English and met at least one of the following criteria:

1. Empirically investigated national bias in international sports judging.
2. Contributed to theoretical or conceptual discussions regarding the mechanisms underlying national bias.
3. Discussed strategies or interventions aimed at mitigating national bias.

No restrictions were imposed concerning publication type, research methodology, statistical approach, or the specific sport analyzed. In line with scoping review guidelines (Tricco et al., 2018), studies were included regardless of their overall quality or risk of bias. However, in recognition of the critical role of stringent statistical testing and judge leniency as a confounding variable in national bias research (Emerson et al., 2009; Emerson & Meredith, 2011), a partial critical appraisal of the included empirical studies was conducted. This appraisal was informed by existing literature emphasizing the potential for biased outcomes when judge leniency is unaccounted for, as well as the frequent absence of stringent statistical testing—such as reliance on simple numerical comparisons of bias estimates rather than inferential statistics. As Munn et al. (2018) have noted, scoping reviews may incorporate such appraisals where appropriate. Peters et al. (2021, p. 2) further outline that 'Scoping reviews typically identify, present and describe relevant characteristics of included sources […]'.

Given that failure to account for judge leniency can significantly distort estimates of national bias (Emerson & Meredith, 2011), its inclusion in the present review is considered both justified and beneficial. Additionally, while analytical approaches have shifted from sign tests, permutation tests, and t-tests toward more robust multivariate regression models capable of simultaneously accounting for confounding variables such as judge leniency, even in more recent studies, statistically rigorous testing is not always employed. Therefore, within the results section addressing various forms of national bias, attention is given to the presence or absence of both judge leniency control and stringent statistical testing. Table 1 summarizes this information for each study included in the review. In interpreting the findings, the results section explicitly identify which studies did not employ judge leniency controls and stringent statistical analysis. The results of these studies are reported in line with

**Table 1.** Contributions included in the review and their respective contributions.

| Author, year | Sport/discipline | Employed statistics | Control of judge leniency | Discusses mechanisms of bias | Discusses measures against national bias | Empirical findings | | | | | | | | Remarks |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Patriotic bias | Competitor bias | Positive political bias | Negative political bias | Vote trading | Indirect compensating bias | Indirect affirming bias | Strategic voting | |
| Ansorge and Scheer (1988) | Men's and Women's Artistic Gymnastics | Sign tests | No | No | Yes | X | X | | | | | | | |
| Ball (1973) | Figure skating | Kendall's W and t-test | Yes | Yes | No | | | | -* | | | | | *Statistical inference based on averaging of Kendall's W coefficients, which may heavily bias results (Wirtz & Caspar, 2002). |
| Bouwens et al. (2022) | Ski jumping | Multiple linear mixed effects Regression | Yes | Yes | Yes | X | | | | | | ≈* | | *Only significant without spectators present. |
| Bring and Carling (1994) | Figure skating | Descriptive statistics | No | No | Yes | ≈* | | | | | | | | *No inferential statistics. |
| Bruine de Bruin (2005) | Figure skating | Multiple linear regression | No | No | No | -* | | | | | | | | *Main focus: serial position effects. |
| Bruine de Bruin (2006) | Figure skating | Multiple linear regression probit regression | No | No | No | -* | | | | | | | | *Main focus: serial position effects. |
| Callahan et al. (2016) | Men's and Women's Artistic Gymnastics | Multiple linear regression | No | Yes | No | | | X | | | | | | |
| Campbell and Galbraith (1996) | Figure skating | Sign test and multiple linear regression | No | Yes | Yes | X | | | | | | | ≈* | *Indication of strategic voting without stringent statistical testing. |
| Deuel (1989) | Dressage | ANOVA | No | No | No | X | | | | | | | | |
| Emerson et al. (2009) | Diving | Multiple linear regression | Yes | Yes | No | X | | | | | | | | |
| Emerson and Meredith (2011) | Diving | Permutation test | Yes | Yes | No | X | | | | | | | | |
| Fang and Ho (2024) | Figure skating | Multiple linear mixed effects regression | No | Yes | Yes | X | | | - | | - | - | | |

(*Continued*)

**Table 1.** Continued.

| Author, year | Sport/discipline | Employed statistics | Control of judge leniency | Discusses mechanisms of bias | Discusses measures against national bias | Empirical findings | | | | | | | | Remarks |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Patriotic bias | Competitor bias | Positive political bias | Negative political bias | Vote trading | Indirect compensating bias | Indirect affirming bias | Strategic voting | |
| Fenwick and Chatterjee (1981) | Figure skating | Inter-judge rank correlation and ANOVA | Yes | Yes | Yes | ≈* | | ≈* | ≈* | | | | | *Statistical inference based on averaging of Spearman Correlation coefficients, which may heavily bias results (Wirtz & Caspar, 2002). |
| Hawson et al. (2010) | Dressage | Restricted maximum likelihood modelling procedure | No | No | No | - | | | | | | | | |
| Heiniger and Mercier (2019) | Acrobatic gymnastics | Multiple linear mixed-effects Regression | Yes | No | Yes | - | - | | | | | | ≈* | *Indication of strategic voting, as national bias estimates are presumably higher for top-placed athletes than overall estimates. However, no statistical inference testing is conducted. |
| | Aerobic gymnastics | | | | | **X** | - | | | | | | ≈* | |
| | Men's artistic gymnastics | | | | | **X** | - | | | | | | ≈* | |
| | Women's artistic gymnastics | | | | | **X** | - | | | | | | ≈* | |
| | Rhythmic gymnastics | | | | | **X** | - | | | | | | ≈* | |
| | Trampoline gymnastics | | | | | - | - | | | | | | ≈* | |
| Heiniger and Mercier (2021) | Acrobatic gymnastics | Multiple linear regression and multi-level linear mixed-effects regression* | Yes | No | Yes | - | - | | | | | | -** | *Multiple linear and multilevel linear mixed effects regression come to comparable results according to the authors. The results presented are from the multiple linear regression model. **No stringent statistical testing for strategic voting. |
| | Aerobic gymnastics | | | | | **X** | - | | | | | | ≈** | |
| | Men's artistic gymnastics | | | | | **X** | - | | | | | | ≈** | |
| | Women's artistic gymnastics | | | | | **X** | **X** | | | | | | -** | |
| | Rhythmic gymnastics | | | | | **X** | **X** | | | | | | -** | |
| | Trampoline gymnastics | | | | | - | - | | | | | | -** | |

(*Continued*)

**Table 1.** Continued.

| Author, year | Sport/discipline | Employed statistics | Control of judge leniency | Discusses mechanisms of bias | Discusses measures against national bias | Empirical findings | | | | | | | | Remarks |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Patriotic bias | Competitor bias | Positive political bias | Negative political bias | Vote trading | Indirect compensating bias | Indirect affirming bias | Strategic voting | |
| Krumer et al. (2022) | Ski jumping | Multiple linear mixed-effects Regression | Yes* | Yes | Yes | X | | | | | - | - | -** | *Judge leniency as leniency over an entire season. **Indication of no strategic voting without stringent statistical testing. |
| Lee (2008) | Figure skating | Autoregressive model | Yes | Yes | Yes | X | | | | | | | X | Primary objective: outlier aversion. |
| Leskošek et al. (2012) | Men's artistic gymnastics | Descriptive statistics | No | No | Yes | ≈* | | | | | | | | *No inferential statistics. |
| Looney (2004) | Figure skating | Many-facet Rasch model | Yes | No | No | ≈* | | | | | | | | *No inferential statistics, national bias just a probable cause of some deviant scores. |
| Lyngstad et al. (2020) | Ski jumping | Multiple linear mixed effects regression | Yes | Yes | No | X | | | | | | | | |
| Morgan and Rotthoff (2014) | Men's artistic gymnastics | Multiple linear mixed-effects regression | No | No | No | X | | | | | | | -* | *Indication of no strategic voting without stringent statistical testing. |
| | Women's artistic gymnastics | | | | | X | | | | | | | -* | |
| Myers et al. (2006) | Muaythai | Multilevel model | Not possible | No | Yes | X* | | | | | | | | *No acknowledgement of the well-known confounding color effect of protective clothing. |
| Popović (2000) | Rhythmic gymnastics | Sign tests and t-tests | No | No | Yes | - | - | | | | | | | |
| Rotthoff (2013) | Artistic gymnastics | Multiple linear regression | No | No | No | - | | | | | | | | Primary objective: sequential effects. |
| Sala et al. (2007) | Figure skating | Multiple linear regression | Yes | Yes | No | X | | - | X | | | | ≈* | *Indication of strategic voting without stringent statistical testing. |
| Sampaio (2012) | Surfing | Multiple linear mixed effects Regression | No | Yes | No | - | X | | | | | | ≈* | *Indication of strategic voting without stringent statistical testing. |
| Sandberg (2018) | Dressage | Multiple linear mixed-effects Regression | Yes | Yes | Yes | X | - | - | - | | | X | -* | *Indication of no strategic voting without stringent statistical testing. |

(*Continued*)

**Table 1.** Continued.

| Author, year | Sport/discipline | Employed statistics | Control of judge leniency | Discusses mechanisms of bias | Discusses measures against national bias | Empirical findings | | | | | | | | Remarks |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Patriotic bias | Competitor bias | Positive political bias | Negative political bias | Vote trading | Indirect compensating bias | Indirect affirming bias | Strategic voting | |
| Scholten et al. (2020) | Ski jumping | Multiple linear mixed effects Regression | Yes* | Yes | Yes | **X** | | | | | | | -** | *Judge fixed effect covers average judge leniency over 41 competitions. **Indication of no strategic voting without stringent statistical testing. |
| Seltzer and Glass (1991) | Figure skating | Descriptive statistics, t-test and ANOVA | Yes | No | No | **X** | | **X*** | **X*** | | | | | *Significant political bias only in certain competitions and between certain blocs. |
| Soler (2021) | Rhythmic gymnastics | Descriptive statistics. inter-judge-correlations, Kendall's W, ANOVA | No | No | Yes | ≈* | | | | | | | | *Indication of patriotic bias without stringent statistical testing. |
| Weekley and Gier (1989) | Figure skating | ANOVA | No | No | Yes | | | - | | | | | | |
| Whissell et al. (1993) | Figure skating | 'One-tailed tests of the significance of probabilities' (S.356) | No | No | Yes | **X*** | | | | | | | | |
| Wolframm (2023) | Dressage | Multiple linear regression | No | Yes | Yes | **X*** | | | | | | **X*** | | *Regression on the mean judge score (TS). |
| Yang (2006) | Figure skating | Multiple linear regression and multiple linear mixed effects regression | No | Yes | Yes | **X** | | ≈ | | | | | | |
| Zitzewitz (2006) | Figure skating | Multiple linear mixed effects regression | Yes | Yes | Yes | **X** | | **X** | **X** | | | **X** | **X** | |
| | Ski jumping | | | | | **X** | | - | - | | **X** | | **X** | |

(*Continued*)

**Table 1.** Continued.

| Author, year | Sport/discipline | Employed statistics | Control of judge leniency | Discusses mechanisms of bias | Discusses measures against national bias | Empirical findings | | | | | | | | Remarks |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Patriotic bias | Competitor bias | Positive political bias | Negative political bias | Vote trading | Indirect compensating bias | Indirect affirming bias | Strategic voting | |
| Zitzewitz (2014) | Figure skating | Multiple linear mixed effects regression | Yes | Yes | Yes | ≈* | | ≈** | ≈** | ≈* | | | **X*** | *Estimated as an aggregate of patriotic bias and vote trading ('Compatriot Judge Effect') **Positive political bias just for east bloc on east bloc countries in the GOE scoring. Negative political bias only for western judges on eastern bloc athletes. |
| Allen et al. (2021) | Artistic gymnastics | – | – | No | Yes | | | | | | | | | |
| Ansorge and Scheer (1984) | Artistic gymnastics | – | – | Yes | No | | | | | | | | | |
| Bailie (1965) | Artistic gymnastics | – | – | Yes | No | | | | | | | | | |
| Balinski and Laraki (2014) | Figure skating | – | – | No | Yes | | | | | | | | | |
| Bar-Eli et al. (2011) | - | – | – | Yes | Yes | | | | | | | | | |
| Bassett and Persky (1994) | Figure skating | – | – | No | Yes | | | | | | | | | |
| Berry (2002) | Figure skating | – | – | No | Yes | | | | | | | | | |
| Brennan (1996) | Figure skating | – | – | Yes | Yes | | | | | | | | | |
| Criley (1972) | Artistic gymnastics | – | – | Yes | No | | | | | | | | | |
| Criley (1976) | Artistic gymnastics | – | – | Yes | No | | | | | | | | | |
| Emerson and Arnold (2011) | Figure skating | – | – | No | Yes | | | | | | | | | |
| Emerson (2007) | Figure skating | – | – | No | Yes | | | | | | | | | |
| FEI (2018) | Dressage | – | – | Yes | Yes | | | | | | | | | |
| Heuschmann (2007) | Dressage | – | – | Yes | Yes | | | | | | | | | |
| Kirkbride (2013) | - | – | – | No | Yes | | | | | | | | | |
| Landers (1970) | Artistic gymnastics | – | – | Yes | Yes | | | | | | | | | |
| McFee (2013) | - | – | – | Yes | Yes | | | | | | | | | |

(*Continued*)

**Table 1.** Continued.

| Author, year | Sport/discipline | Employed statistics | Control of judge leniency | Discusses mechanisms of bias | Discusses measures against national bias | Empirical findings | | | | | | | | Remarks |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Patriotic bias | Competitor bias | Positive political bias | Negative political bias | Vote trading | Indirect compensating bias | Indirect affirming bias | Strategic voting | |
| Osório (2017) | - | – | – | No | Yes | | | | | | | | | |
| Osório (2020) | - | – | – | Yes | Yes | | | | | | | | | |
| Plessner and Haar (2006) | - | – | – | Yes | Yes | | | | | | | | | |
| Roetzheim and Muzyczko (1986) | Artistic gymnastics | – | – | No | Yes | | | | | | | | | |
| Russell (2001) | - | – | – | No | Yes | | | | | | | | | |
| Scheer and Ansorge (1987) | Artistic gymnastics | – | – | No | Yes | | | | | | | | | |
| Ste-Marie and Lee (1991) | Artistic gymnastics | – | – | Yes | Yes | | | | | | | | | |
| Ste-Marie (1996) | Artistic gymnastics | – | – | Yes | No | | | | | | | | | |
| Ste-Marie (2003) | Artistic gymnastics | – | – | Yes | Yes | | | | | | | | | |
| Ste-Marie et al. (2001) | Artistic gymnastics | – | – | Yes | Yes | | | | | | | | | |
| Truchon (2004) | Figure skating | – | – | No | Yes | | | | | | | | | |
| Wettstone (1968) | Artistic gymnastics | – | – | Yes | Yes | | | | | | | | | |
| Wu and Yang (2004) | Figure skating | – | – | No | Yes | | | | | | | | | |
| Goodwin (2004) | Figure skating | – | – | Yes | No | | | | | | | | | |

*Note.* **X** = statistically significant finding ($p < 0.05$), ≈ = Indication without statistical inference, **-** = no statistically significant finding ($p > 0.05$).

scoping review standards (Tricco et al., 2018), though caveats are clearly noted regarding limitations in analytical rigor.

The full review checklists, protocol, and additional documentation such as compiled funding statements can be accessed through the Open Science Framework: https://osf.io/8u6cm/overview?view_only=c8a4298433c14190865c62cce6cbf7a8.

### 2.1. Search strategy

Preliminary exploratory searches identified key search terms: *sport* (Heiniger & Mercier, 2019, 2021), *judge* or *judging* (Ansorge & Scheer, 1988; Ball, 1973; Fang & Ho, 2024; Soler, 2021), *rating* (Bassett & Persky, 1994; Emerson et al., 2009; Soler, 2021), *scoring* (Ball, 1973; Deuel, 1989; Soler, 2021), *bias* (Ansorge & Scheer, 1988; Campbell & Galbraith, 1996), *national* or *nationalistic* or *nationalism* (Deuel, 1989; Fang & Ho, 2024; Heiniger & Mercier, 2019; Soler, 2021; Whissell et al., 1993; Zitzewitz, 2006), *error* (Heiniger & Mercier, 2019, 2021), *corruption* (Fang & Ho, 2024; Soler, 2021; Zitzewitz, 2014), *favoritism* (Ball, 1973; Bouwens et al., 2022; Fang & Ho, 2024; Krumer et al., 2022; Soler, 2021), and '*patriotic bias*' (Sala et al., 2007, p. 18). These terms were compiled into the following initial search string:

$$\textbf{Sport AND}(\textbf{judg}*\textbf{OR rat}*\textbf{OR scor}*)\textbf{AND}((\textbf{Bias AND}(\textbf{National}*\textbf{OR Patriot}*))\textbf{OR Error OR Corrupt}*\textbf{OR Fav}*\textbf{OR Patriot}*)$$

Database-specific modifications to this search string were made to transparently exclude false positive findings. The full search strategies for all databases, including any filters and limits used, are presented in Tables A1–A5 of the Appendix. This process yielded 317 findings, of which 23 were found eligible following full-text screening.

### 2.2. Reference tracking

To address common database limitations (Richards, 2006; von Elm et al., 2019), an iterative reference-tracking approach was employed. Initially, 761 references were extracted from the 23 eligible contributions. After removing duplicates and already included articles, 608 references were screened, yielding 29 additional eligible contributions. A second iteration extracted 598 references from these new contributions (including 117 duplicates and 78 previously screened contributions), resulting in 12 further eligible contributions. A third and final iteration identified one additional eligible study, concluding the reference-tracking process.

### 2.3. Data charting and synthesis

Data charting was conducted in alignment with the four central areas of inquiry: (1) the forms of national bias examined in the context of international sports judging, including supporting empirical evidence and context-specific patterns; (2) the underlying mechanisms of national bias proposed in the existing literature, along with any available empirical support; (3) indications and evidence regarding the intentional or unintentional exhibition of national bias as discussed in prior research; and (4) proposed measures to mitigate or eliminate national bias in international sports judging, including potential limitations and unintended consequences.

Table 1 in the Results section provides an overview of the specific categories of information extracted from each included source. The synthesis of findings was structured to deliver:

1. A comprehensive overview of the various forms of national bias studied in the context of international sports judging, highlighting both empirical evidence and any contextual nuances;
2. A summary of mechanisms identified in the literature as (potentially) contributing to national bias;
3. An aggregation and interpretation of findings from the extant literature concerning the potential intentionality of national bias; and

4. A synthesis of countermeasures to national bias proposed in the extant literature, including a critical assessment of their limitations and possible unintended effects.

Given the broad scope of the review and the heterogeneity of methodological approaches and reported data, the data synthesis primarily follows a qualitative, narrative approach structured around the four guiding research questions. Where appropriate, descriptive quantitative summaries are also provided.

## 3. Results

Figure 1 illustrates the literature search process, which yielded 65 eligible contributions. Additionally, three eligible contributions identified through exploratory searches (Ball, 1973; Fang & Ho, 2024; Soler, 2021) were not captured by the systematic search but were included due to their relevance. Consequently, this review encompasses a total of 68 studies, of which 37 employ statistical analyses of scoring data to empirically estimate national bias. The remaining contributions consist of theoretical discussions published in academic journals, sports magazines, and books.

Table 1 summarizes these contributions, listing empirical (upper section) and non-empirical (lower section) contributions alphabetically. It provides details on the year of publication, the sports examined, the statistical methods employed, and the control of judge leniency, as well as the empirical findings of national bias in its several forms. Additionally, the extraction of discussed bias mechanisms and proposed countermeasures is indicated. Lastly, the table also includes remarks on notable limitations to aid in the interpretation of results.

### 3.1. Forms of national bias

Of the 68 contributions reviewed, 37 empirically investigate different manifestations of national bias. Table 2 summarizes the number of studies conducted for each sport and each form of bias.

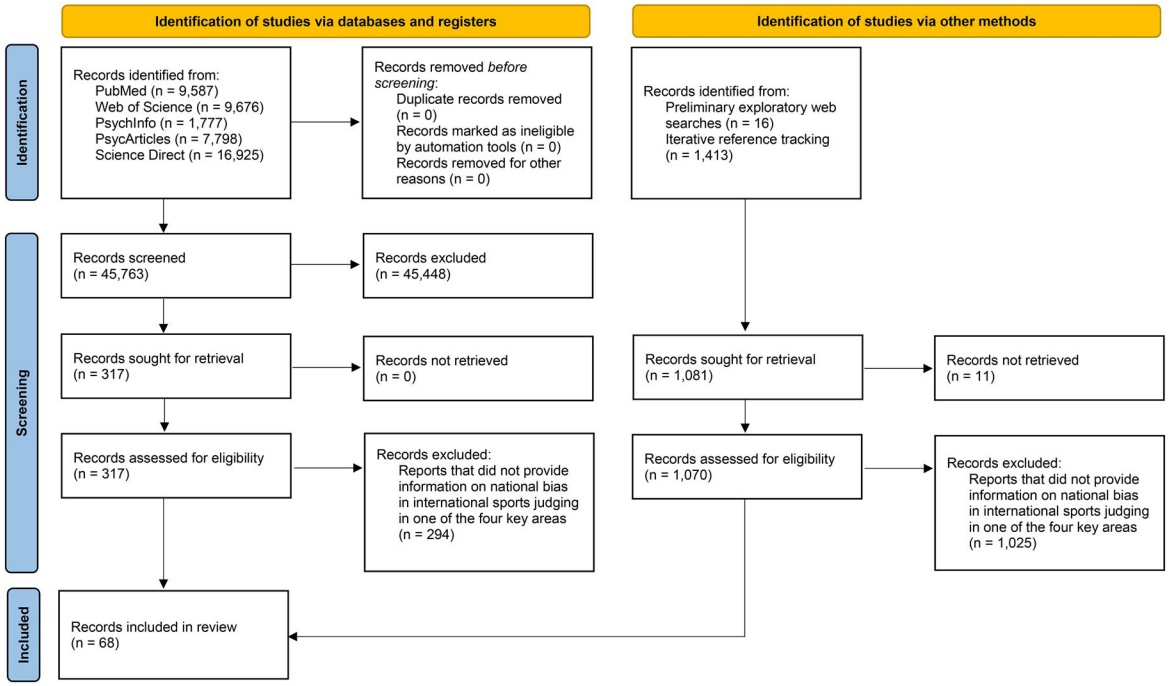


**Figure 1.** Schematic representation of the literature search process following PRISMA 2020 guidelines, adapted with minor modifications from Page et al. (2021) in accordance with the recommended standards for figure presentation.

**Table 2.** Number of empirical studies on the various forms of national bias in international sports judging.

| Sport | Patriotic bias | Competitor bias | Political bias | Vote trading | Indirect bias | Overall* |
|---|---|---|---|---|---|---|
| Figure skating | 15 | 0 | 9 | 1 | 2 | **16** |
| Artistic gymnastics | 6 | 3 | 1 | 0 | 0 | **7** |
| Ski jumping | 5 | 0 | 1 | 0 | 3 | **5** |
| Rhythmic gymnastics | 4 | 3 | 0 | 0 | 0 | **4** |
| Dressage | 4 | 1 | 1 | 0 | 2 | **4** |
| Aerobic gymnastics | 2 | 2 | 0 | 0 | 0 | **2** |
| Acrobatic gymnastics | 2 | 2 | 0 | 0 | 0 | **2** |
| Trampoline gymnastics | 2 | 2 | 0 | 0 | 0 | **2** |
| Diving | 2 | 0 | 0 | 0 | 0 | **2** |
| Surfing | 1 | 1 | 0 | 0 | 0 | **1** |
| Muaythai | 1 | 0 | 0 | 0 | 0 | **1** |
| Overall* | **34** | **6** | **11** | **1** | **6** | |

*As several studies examine multiple forms of national bias and include analyses of more than one sport, the total number of contributions does not correspond to the sum of the rows or columns.

The most extensively studied form of national bias is patriotic bias, characterized by judges awarding inflated scores to athletes from their own country. This form of bias has been examined in 34 studies to date. In contrast, competitor bias—where judges assign deflated scores to athletes who compete directly against their compatriots—has received considerably less scholarly attention, with only six studies investigating this second fundamental form of national bias besides patriotic bias (Heiniger & Mercier, 2019). Political bias, in which judges favor athletes from politically allied nations or disadvantage those from politically adversarial countries, has been the focus of eleven studies. Vote trading, involving collusion among judges who exchange favorable scores across events or competitions to reciprocally benefit athletes from their own countries, has been empirically examined in only one study. Finally, six studies have investigated scoring data from international aesthetic sports competitions to explore two indirect forms of national bias: compensating bias, where judges offset each other's patriotic scoring tendencies, and affirming bias, where judges reinforce such tendencies.

The following sections (3.1.1–3.1.5) provide a critical synthesis of empirical findings related to each of these identified forms of national bias.

### *3.1.1. Patriotic bias*

Overall, 34 empirical studies investigated national bias in its patriotic bias form, with 28 reporting (significant) findings or indications, some based on purely descriptive analyses. Only a few (8) studies (partially) found no evidence of this bias in figure skating (Bruine de Bruin, 2005, 2006), artistic gymnastics (Morgan & Rotthoff, 2014; Rotthoff, 2015), dressage (Hawson et al., 2010), rhythmic gymnastics (Popović, 2000), and acrobatic and trampoline gymnastics (Heiniger & Mercier, 2019, 2021). However, most of these studies suffer from considerable methodological limitations, particularly the failure to control for judge leniency, the key confounding factor in national bias research (Emerson et al., 2009; Emerson & Meredith, 2011).

Of the 34 empirical studies, 17 did not account for judge leniency. Among the remaining 17 studies, four relied on basic statistical tests such as t-tests (Seltzer & Glass, 1991), ANOVA (Deuel, 1989), permutation tests (Emerson & Meredith, 2011), and many-facet Rasch modelng (Looney, 2004), which limits their analytical power. Additionally, one study on Muay Thai (Myers et al., 2006) failed to account for the well-documented influence of clothing color in combat sports (Hill & Barton, 2005; Sorokowski et al., 2014).

The remaining 12 studies employed linear regression modeling, allowing for the control of multiple confounding influences. These studies consistently identified significant patriotic bias in figure skating (Lee, 2008; Sala et al., 2007; Zitzewitz, 2006, 2014), ski jumping (Bouwens et al., 2022; Krumer et al., 2022; Lyngstad et al., 2020; Scholten et al., 2020; Zitzewitz, 2006), artistic, aerobic, and rhythmic gymnastics (Heiniger & Mercier, 2019, 2021), diving (Emerson et al., 2009), and dressage (Sandberg, 2018). Notably, Sandberg (2018) observed that dressage judges also tend to favor horses from their own countries; however, this preference did not amplify their bias towards compatriot riders.

The magnitude of patriotic bias, when expressed in comparable statistical units, ranges from 13% to 50% of the within-performance standard deviation of scores[1] (Bouwens et al., 2022; Krumer et al., 2022; Lyngstad et al., 2020; Sala et al., 2007; Sandberg, 2018; Zitzewitz, 2006, 2014). However, it varies significantly across judges and nationalities (Heiniger & Mercier, 2019; Krumer et al., 2022; Lyngstad et al., 2020;

Sala et al., 2007; Zitzewitz, 2006) and correlates with the country's Transparency International Corruption Perceptions Index (Krumer et al., 2022; Zitzewitz, 2006) and geographically, with increased bias observed in judges from more eastern nations (Zitzewitz, 2006). However, no such correlation was found in gymnastics disciplines (Heiniger & Mercier, 2021), and Sandberg (2018) found no significant relationship between patriotic bias and nationalism indicators in dressage. Furthermore, while no gender differences were observed in the extent of patriotic bias (Sandberg, 2018), younger ski jumping judges appear less biased than their older colleagues (Scholten et al., 2020).

Some studies additionally suggest that patriotic bias is more pronounced in final competition rounds and among higher-ranked athletes, indicating a potential element of intentionality (Heiniger & Mercier, 2019, 2021; Zitzewitz, 2006). However, other studies report no evidence that higher stakes, such as medal contention, exacerbate patriotic bias (Campbell & Galbraith, 1996; Krumer et al., 2022; Lee, 2008; Morgan & Rotthoff, 2014; Sala et al., 2007; Sandberg, 2018; Scholten et al., 2020).

Finally, patriotic bias is believed to be more prominent in subjective scoring aspects, where greater flexibility in scoring allows for more substantial unsanctioned distortions (Fang & Ho, 2024; Fenwick & Chatterjee, 1981; Lee, 2008; Sandberg, 2018; Yang, 2006; Zitzewitz, 2006, 2014). However, these beliefs are mostly drawn from nominal differences in bias estimates and lack robust statistical inference testing (Fang & Ho, 2024; Fenwick & Chatterjee, 1981; Lee, 2008; Zitzewitz, 2006), and two studies found no significant differences between subjective and objective scoring aspects (Sandberg, 2018; Yang, 2006).

### *3.1.2. Competitor bias*

Competitor bias has received considerably less scholarly attention than patriotic bias, with only six studies investigating its presence. Of these, three studies reported statistically significant findings in artistic gymnastics (Ansorge & Scheer, 1988; Heiniger & Mercier, 2021) and surfing (Sampaio, 2012), while only Heiniger and Mercier (2021) study on competitor bias accounted for judge leniency. The magnitude of competitor bias was minimal in this study, accounting for only 5–7% of the within-performance standard deviation of scores. Additionally, no evidence of competitor bias was found in acrobatic, aerobic, men's artistic gymnastics, or trampoline gymnastics (Heiniger & Mercier, 2021). Similarly, Sandberg (2018), also accounting for judge leniency, reported no significant competitor bias in dressage.

The common assumption among these studies was that close competitive relationships, the basis of competitor bias, persist between athletes ranked adjacently in the respective competition (Ansorge & Scheer, 1988; Heiniger & Mercier, 2019, 2021; Popović, 2000; Sandberg, 2018). However, competitive tension may be more pronounced among medal contenders (Campbell & Galbraith, 1996; Heiniger & Mercier, 2019, 2021) or among athletes positioned near critical qualification thresholds (Sandberg, 2018), which may explain the limited empirical evidence for competitor bias (Braeunig, 2024).

Thus, Sampaio's (2012) study on surfing, where competitive relationships form more explicitly in head-to-head heats, marks a notable exception. Sampaio found substantial competitor bias, with indications of strategic scoring behavior: competitor bias was evident when a compatriot was at risk of losing but diminished when the compatriot was already winning. However, Sampaio's findings are again limited by the absence of controls for judge leniency (Emerson et al., 2009; Emerson & Meredith, 2011).

### *3.1.3. Political bias*

Since Criley (1972, 1976) first acknowledged the influence of national politics on scoring in aesthetic sports, ten studies have investigated political bias in sports judging, with six studies reporting significant findings or indication of political bias in artistic gymnastics (Callahan et al., 2016) and figure skating (Fenwick & Chatterjee, 1981; Sala et al., 2007; Seltzer & Glass, 1991; Zitzewitz, 2006, 2014). Among these, only three studies accounted for judge leniency and provided significant evidence of bloc-based biases during the Cold War (Sala et al., 2007; Seltzer & Glass, 1991; Zitzewitz, 2006).

Seltzer and Glass (1991) analyzed figure skating scores from 1968 to 1988, identifying a pattern of 'bloc judging' (Yang, 2006, p. 10; Zitzewitz, 2006, p. 69; 2014, p. 19). Their findings indicate that judges from the Western bloc tended to assign lower scores to Eastern bloc athletes, whereas Eastern bloc judges displayed mild favoritism toward compatriots while demonstrating bias against Western

athletes. However, neutral judges occasionally exhibited greater bias than their bloc-affiliated counterparts. Furthermore, methodological concerns such as the misclassification of neutral nations and the overrepresentation of U.S. judges, impose considerable limitations on these findings (Seltzer & Glass, 1991). Similarly, Sala et al. (2007) identified bloc bias between NATO and Warsaw Pact nations in figure skating competitions spanning 1948–2002. Their findings suggest that bilateral political relations did not necessarily result in reciprocal scoring advantages, and bloc-based bias diminished after the Cold War. Zitzewitz (2006) used maximum likelihood estimation to infer political blocs based on scoring patterns, revealing results that closely mirrored the alignments observed in the 2002 Olympic figure skating vote trading controversy. However, this study found no considerable intra-bloc bias and no evidence of political bias in ski jumping and Emerson et al. (2009) noted an absence of rigorous statistical testing.

Additionally, some research suggests that cultural factors may also shape judges' scoring behavior. Fenwick and Chatterjee (1981) argue that low score correlations between Warsaw Pact and NATO judges may stem from cultural differences rather than political affiliations. Their claim is supported by positive correlations between West and East German judges, which Seltzer and Glass (1991) later cited as further evidence of cultural influence. However, Fenwick and Chatterjee (1981) analyses did not account for judge leniency and their methodology, involving averaging correlation coefficients, is substantially flawed (Wirtz & Caspar, 2002).

Further evidence of non-political influences comes from Yang (2006), who found that judges exhibited favoritism toward athletes from neighboring countries, nations with similar legal systems, and those frequently covered in national media. Similarly, Callahan et al. (2016) observed that gymnastics judges awarded higher scores for difficult routines when the athletes' home countries had trade agreements with the judges' nations. However, Callahan et al. (2016) did not rule out political influences, and both Callahan et al. (2016) and Yang (2006) did not account for judge leniency, which limits the validity of their findings. Lastly, Sandberg (2018) demonstrated that dressage judges exhibited national bias independent of bilateral trust or historical conflicts, further suggesting that political bias alone does not fully account for observed scoring patterns.

#### *3.1.4. Vote trading and compatriot judge effect*

The phenomenon of vote trading has been widely acknowledged in academic literature, with one of the most notorious examples being the alleged collusion between French judge Marie-Reine Le Gougne and a Russian judge during the 2002 Olympic pairs skating event in Salt Lake City (Emerson et al., 2009; Kirkbride, 2013; Lyngstad et al., 2020; McFee, 2013; Yang, 2006; Zitzewitz, 2014). Vote trading may occur within a single competition or across multiple events, wherein judges exchange favorable scores to mutually benefit their compatriots (Sandberg, 2018).

Despite some anecdotal evidence, including Le Gougne's initial admission of collusion, which she later retracted, empirical analyses of scoring data have found no direct evidence of misconduct by the French judge (Lyngstad et al., 2020). To date, only one empirical study has touched on vote trading, reporting a significant 'compatriot judge effect' (Zitzewitz, 2014, p. 13), which encapsulates both patriotic bias and vote trading. However, anonymized scoring data prevented the differentiation of these two components, which amounted to 13.7% of the within-performance standard deviation under the Code of Points system and 17.8% under the older 6.0 scoring system (Zitzewitz, 2014). This places the compatriot judge effect on the lower end of previously estimated patriotic bias magnitudes, suggesting that vote trading does not significantly contribute to the compatriot judge effect beyond patriotic bias.

#### *3.1.5. Indirect national bias – compensating bias and affirming bias*

A more nuanced form of national bias arises when judges adjust their scoring in response to perceived or suspected patriotic bias among their colleagues. This phenomenon has been investigated by six studies, three of which reported significant findings in figure skating and ski jumping (Zitzewitz, 2006) and dressage (Sandberg, 2018; Wolframm, 2023). Among these, two studies accounted for judge leniency (Sandberg, 2018; Zitzewitz, 2006).

In ski jumping, judges appear to counteract their colleagues' patriotic bias by assigning lower scores to athletes from the same country as the presumably biased judge. This compensating bias is particularly pronounced among judges from countries with lower rankings on the 2001 Transparency International Corruption Perceptions Index and when the compensated judge originates from a nation historically associated with strong patriotic bias (Zitzewitz, 2006). The opposing affirming bias, whereby judges reinforce rather than mitigate national bias, has been observed in figure skating and dressage. Zitzewitz (2006) found that when an athlete's nationality is represented on the judging panel, non-compatriot figure skating judges reportedly align their scores with those of the compatriot judge, thereby amplifying the bias. A similar pattern was observed in international dressage competitions, where judges tended to favor athletes whose nationality was also represented on the panel, potentially due to reluctance to provide outlier scores or temporary in-group dynamics within judging panels (Sandberg, 2018).

As a result, athletes in ski jumping may be disadvantaged by the presence of a compatriot judge, whereas those in figure skating and dressage are more likely to benefit from such representation. However, contrary to Zitzewitz (2006), Bouwens et al. (2022) reported evidence of affirming bias in ski jumping, though their findings were statistically significant only in the absence of an audience. Accordingly, Krumer et al. (2022) argue that controlling for home advantage substantially reduces the significance of indirect bias in ski jumping, thereby raising concerns that prior findings may have been influenced by methodological differences in accounting for this factor.

### *3.2. The mechanisms behind national bias*

The concept of bias often carries implications of deliberate manipulation (Emerson et al., 2009) and national bias is frequently attributed to overt nationalistic attitudes. This perspective is reflected in the FEI's explicit stance: 'Patriotism bias: Is not acceptable, it has nothing to do with psychological phenomena that are common to all of us, but is a simple nationalistic attitude' (FEI, 2018, p. 6). However, empirical research suggests that national bias may arise from non-nationalistic considerations, including unconscious cognitive processes (Emerson et al., 2009; Emerson & Meredith, 2011; Landers, 1970; Lyngstad et al., 2020; Ste-Marie, 1996; Ste-Marie & Lee, 1991).

From the perspective of social cognition research, national bias is believed to emerge in the final stage of a four-step information-processing framework, consisting of (1) *perception*, (2) *categorization*, (3) *memory processes*, and (4) *information integration* (Graf, 2010; Plessner & Haar, 2006). At this final stage, information about an athlete's performance—already encoded and categorized—is integrated with stored memory information to form a judgment (Plessner & Haar, 2006). Several mechanisms have been proposed to explain how bias-inducing information is acquired and subsequently influences decision-making in international sports judging. Lyngstad et al. (2020) provided a foundational overview of these mechanisms, which was later expanded by Braeunig (2024) into the following five key mechanism domains:

1. Intentional national bias mechanism
2. 'Social psychological mechanism' (Lyngstad et al., 2020, p. 252)
3. 'Cultural legitimacy mechanism' (Lyngstad et al., 2020, p. 252)
4. 'Differential professionalization mechanism' (Lyngstad et al., 2020, p. 252)
5. Mere exposure mechanism

These mechanisms, though not mutually exclusive, underscore the complex and multifaceted nature of national bias, potentially arising from a combination of personal interests, career considerations, and unconscious cognitive or social-psychological processes. They also align with a process-sociological framework (Stern, 2010), emphasizing the interplay between individual agency and systemic influences in shaping national bias.

#### *3.2.1. Intentional national bias mechanism*

Intentional national bias occurs when judges consciously adjust scores based on their own national affiliations or external pressures (Lyngstad et al., 2020). Sala et al. (2007) highlight a role conflict among

international judges, who must navigate dual allegiances as they represent both an international sports organization—mandated to ensure impartiality[2]—and their respective national sports association with vested interests (Sala et al., 2007).

This dual allegiance creates tension, as captured by Scheer and Ansorge (1987, p. 6): 'The current system prevents a judge from being fair. To be fair may cost a judge's team dearly'. The likelihood of intentional bias is particularly heightened in contexts emphasizing national identity, such as international championships or team competitions (Osório, 2020; Sala et al., 2007; Sandberg, 2018). Empirical evidence confirms that national bias is more pronounced in team events, where national affiliations are prominently displayed (Zitzewitz, 2006). In state-controlled sports systems, such as those in the former USSR or China, judges may even be politically appointed and expected to align with national interests (Sala et al., 2007) as former figure skating champion Katarina Witt acknowledged: 'They were told to. They had no choice' (Brennan, 1996, p. 74). Similarly, Goodwin (2004)—cited in Sala et al. (2007)—noted: 'Some judges are tightly controlled by the presidents of their national federations…. They know that if they don't follow orders their federation president will never throw them a good judging assignment'. (p. 21). Lastly, Scheer and Ansorge (1987) argue that 'In selecting judges for international competition, federations of competing countries are more interested in sending judges who are skilled at bias than skilled in judging' (p. 6).

However, loyalty and pressure are not the sole drivers of intentional national bias. Career incentives may also contribute to intentional national bias (Fang & Ho, 2024; Goodwin, 2004; Zitzewitz, 2006). Judges may exhibit bias to secure future assignments, maintain favorable relationships with national federations, or meet the expectations of event organizers and stakeholders (Fang & Ho, 2024; Heuschmann, 2007; Zitzewitz, 2006). For example, in figure skating, where national federations select judges, bias levels are higher compared to ski jumping, where judges are appointed by the international federation (Zitzewitz, 2006).

On the contrary, career concerns may also reduce bias, as judges might avoid outlier scores to preserve credibility when under scrutiny (Brennan, 1996; Lee, 2008). Judges may also strategically adjust scores in response to their colleagues' biases. In figure skating and dressage, judges have been observed aligning with biased colleagues, possibly to ensure panel consistency, signal competence, or correspond to a temporary unity on judging panels (Fang & Ho, 2024; Sandberg, 2018; Zitzewitz, 2006).

Conversely, in ski jumping, judges reportedly counteract perceived biases from colleagues (Zitzewitz, 2006), suggesting that bias dynamics may be influenced by sport-specific norms (Bar-Eli et al., 2011).

Lastly, Campbell and Galbraith (1996) describe an informal scoring practice in which, during moments of uncertainty, judges assign the higher of two possible scores to their compatriots, while athletes from other nations receive the higher or lower alternative at random. This leads to an incremental scoring advantage of approximately 0.5 points, as evidenced in their analysis of Olympic figure skating data.

Overall, intentional national bias seems to be shaped by a complex interplay of personal motivations, institutional pressures, and contextual factors, with variations across sports and regulatory frameworks.

#### *3.2.2. Social psychological mechanism*

The social psychological mechanism reflects judges' conscious or unconscious preferences for their compatriots, which manifest as patriotic bias (Landers, 1970; Lyngstad et al., 2020). These 'inherent preferences' (Bouwens et al., 2022, p. 3; Krumer et al., 2022, p. 278) arise from personal relationships between judges and athletes from the same national sports system (Osório, 2020; Sala et al., 2007). Additionally, Sandberg (2018) highlights that judges may also exhibit affirming bias due to temporary unity within judging panels (Sandberg, 2018; Wolframm, 2023).

Present stakeholders, such as spectators, may aggravate (Lyngstad et al., 2020), or reduce national bias (Bouwens et al., 2022) depending on their behavior and context.

#### *3.2.3. Cultural legitimacy mechanism*

The cultural legitimacy mechanism is based on the premise that judges and athletes internalize shared stylistic and evaluative norms within their national sporting systems (Lyngstad et al., 2020). These norms develop through interactions with senior judges, who play a pivotal role in shaping the perceptions of

performance among both athletes and junior judges (Emerson & Meredith, 2011; Lyngstad et al., 2020; Zitzewitz, 2014). At the national level, judges actively engage in training sessions, competitions, and take on advisory roles for athletes and coaches. This mutual exchange of expertise fosters common stylistic preferences and standardized evaluative frameworks (Berry, 2002; Campbell & Galbraith, 1996; Emerson & Meredith, 2011; Lyngstad et al., 2020; Sala et al., 2007; Sampaio, 2012; Zitzewitz, 2014). These frameworks differ across nations, contributing to observed patterns of patriotic bias (Emerson et al., 2009; Goodwin, 2004; Lyngstad et al., 2020; Zitzewitz, 2014).

Furthermore, McFee (2013) illustrates how scoring patterns—such as USSR judges favoring USSR athletes—can be attributed to common training processes, which condition judges to prioritize specific qualities exhibited by culturally related athletes. Similarly, Fenwick and Chatterjee (1981) argue that cultural factors, rather than political considerations, primarily drive judging behaviors in figure skating. This argument is supported by Seltzer and Glass (1991), who observed similar scoring patterns between East and West German judges during the Cold War. In addition, Landers (1970) suggests that shared political ideologies also play a role in shaping biases between nations. These cultural affiliations influence scoring decisions, often leading to higher evaluations for compatriots or athletes from culturally aligned nations, particularly in subjective performance assessments (Callahan et al., 2016; Fenwick & Chatterjee, 1981; Zitzewitz, 2006). Several factors contribute to this phenomenon, including shared religious or linguistic backgrounds, common national borders, similar legal systems, and reciprocal media representation. Such elements help foster mutual sentiment and trust, potentially explaining instances of positive mutual bias (Callahan et al., 2016; Yang, 2006).

However, globalization may gradually diminish the influence of this mechanism. Empirical evidence suggests that younger judges, having been socialized in a more globalized environment, exhibit lower levels of national bias compared to their older counterparts (Scholten et al., 2020). This trend indicates a potential erosion of culturally ingrained biases over time. Nonetheless, alternative explanations, such as career-related strategic behavior and greater leniency among younger judges, must also be considered when interpreting these findings (Scholten et al., 2020).

#### *3.2.4. Differential professionalization mechanism*

Lyngstad et al. (2020) propose that national professionalization levels within a sport impact judging bias, suggesting that greater professionalization may lead to more impartial judging. Their study found that ski jumping judges from Finland and Norway exhibit lower levels of national bias, which they attribute to the high degree of professionalization in ski jumping within these countries. However, an analysis of temporal trends in ski jumping does not indicate a significant decline in bias over time, suggesting that professionalization may have a limited impact on reducing national bias (Lyngstad et al., 2020). Moreover, while initial results from figure skating seem to provide support for this mechanism (Ball, 1973), flawed statistics and a variety of alternative explanations raise doubts about the results. As a result, the assumption that higher levels of professionalization inherently reduce patriotic bias requires further scrutiny, particularly given that existing evidence is drawn from a limited number of sports and countries.

Moreover, an alternative perspective suggests that increased professionalization may exacerbate national bias rather than mitigate it. Greater formalization and institutionalization within a sport may reinforce cultural legitimacy mechanisms, further embedding nation-specific socialization processes among judges. At the same time, heightened national aspirations for international success could intensify judges' career-related motivations, increasing their incentive to favor compatriot athletes in order to align with national federations or secure future judging opportunities. This complexity underscores the need for a more nuanced examination of the relationship between professionalization and judging bias, rather than assuming a straightforward reduction in bias as professionalization increases.

#### *3.2.5. Mere exposure mechanism*

The mere exposure mechanism posits that repeated exposure to a stimulus, such as an athlete's performance, enhances cognitive fluency, which in turn increases preference for that stimulus. This *fluency heuristic* is particularly influential in non-analytical judgments, such as forming overall impressions (Ste-Marie, 1996). In the context of sports judging, judges who are more familiar with their compatriots'

performances may unintentionally assign higher scores (Ste-Marie, 1996; Ste-Marie & Lee, 1991). Notably, conscious recollection of prior exposure is not necessary for this bias to emerge. Experimental research has demonstrated that memory effects persist even when athletes wear different attire during subsequent performances, indicating that the bias is driven by underlying cognitive mechanisms rather than explicit memory (Moreland & Zajonc, 1977; Ste-Marie, 2003; Ste-Marie et al., 2001). Ste-Marie (1996) experimental studies in artistic gymnastics further support this mechanism, showing that repeated exposure facilitates cognitive fluency, leading to higher scores for more familiar athletes.

However, for exposure to influence judgment, it must exceed a certain threshold, and corrective cognitive processes may mitigate its impact in real competitive settings. This is particularly true when judges actively recall their prior exposure to specific athletes, which may prompt deliberate efforts to counteract potential bias (Ste-Marie, 1996). Consequently, while the mere exposure mechanism may contribute to unintentional bias, its overall impact in real-world competitions is likely limited (Ste-Marie, 1996).

### 3.3. National bias – evidence suggesting an intentional component

National bias in sports judging manifests in multiple forms, ranging from overt collusion to potentially unintentional score distortion. One explicit form, vote trading, involves coordinated score manipulation among judges. Other manifestations, such as competitor bias or indirect bias through score compensation or affirmation, suggest intent, although subconscious mechanisms cannot be entirely dismissed for indirect national bias (Lyngstad et al., 2020; Sandberg, 2018). Correspondingly, patriotic bias is often linked to corruption and favoritism (Emerson et al., 2009), while theoretical frameworks acknowledge that it may also arise unintentionally or through subconscious mechanisms (Emerson et al., 2009; Emerson & Meredith, 2011; Lyngstad et al., 2020; Seltzer & Glass, 1991).

Empirical evidence suggests that some judges engage in 'strategic voting' (Sandberg, 2018, p. 2144), which involves selectively favoring compatriots with strong medal prospects while counterbalancing one's own bias by assigning lower scores to compatriots with minimal chances of success (Morgan & Rotthoff, 2014; Sandberg, 2018). This calculated approach enables judges to avoid detection and sanctions[3] (Bouwens et al., 2022; Sandberg, 2018), reflecting a rational cost-benefit analysis (Sandberg, 2018) and thus indicates intent.

Of the 12 studies that reflected on aspects of strategic voting, eight provide evidence or indications of such judging behavior, with six employing judge leniency control. Heiniger and Mercier (2019, 2021) reported increased national bias during gymnastics finals and among top-eight competitors, indicating that judges 'bend the rules further when it counts' (Heiniger & Mercier, 2019, p. 4; 2021, p. 11). Similarly, ski jumping judges reportedly displayed greater bias in prestigious competitions, decisive competition moments, and disciplines involving more subjective scoring, such as the 90-meter hill (Zitzewitz, 2006). Furthermore, patriotic bias seems more pronounced in sports where scoring is highly subjective, as judges have greater discretion in evaluating performances due to increased variability in scoring (Lee, 2008; Sandberg, 2018; Yang, 2006; Zitzewitz, 2006). Interestingly, some findings additionally suggest that judges adjust their behavior under scrutiny. In high-stakes figure skating events, Zitzewitz (2006) observed reduced patriotic bias, suggesting that judges may avoid scrutiny by (over)compensating for their biases. Similarly, Lee (2008) identified 'outlier aversion', where judges align their scores more closely with those of their peers after assigning an extreme score. This behavior, attributed to career concerns, underscores the social and professional pressures that influence judging decisions.

However, findings of strategic voting often rely on descriptive differences rather than rigorous statistical testing. While Zitzewitz (2006) identified statistically significant differences in patriotic bias between top-10 and non-top-10 finishers, many studies drew their conclusions based on nominal differences without stringent statistical testing (Campbell & Galbraith, 1996; Fang & Ho, 2024; Fenwick & Chatterjee, 1981; Krumer et al., 2022; Lee, 2008; Morgan & Rotthoff, 2014; Sala et al., 2007; Sandberg, 2018; Scholten et al., 2020; Yang, 2006).

Finally, despite some evidence of strategic bias, some research suggests that subconscious mechanisms also contribute to national bias. Scholten et al. (2020) found no increase in patriotic bias during

critical ski jumps, leading them to conclude that bias in this context is likely unintentional. This finding indicates that while deliberate bias occurs in certain sports, unintentional and subconscious factors may also play a role. This suggests a nuanced interplay between intentional and unintentional as well as subconscious influences and highlights the complexity of national bias in sports judging.

### 3.4. Proposed measures against national bias

Osório (2017, 2020) asserts that bias-free scoring systems are inherently unattainable under social choice theory, as judges can obscure biases through strategic voting behaviour. Nevertheless, increasing awareness of national bias has been suggested as an initial step toward its mitigation (Plessner & Haar, 2006), and various strategies have been proposed to address national bias in international sports judging.

#### 3.4.1. Structural modifications to judging panels

One approach prohibits judges from scoring athletes of their own nationality (Heiniger & Mercier, 2019; Myers et al., 2006; Soler, 2021; Whissell et al., 1993), a measure already implemented in acrobatic gymnastics finals (Heiniger & Mercier, 2021; Wettstone, 1968). However, this method is infeasible in many international competitions due to the large number of participating nations (Sandberg, 2018; Soler, 2021) and may also require costly training of judges from underrepresented countries (Soler, 2021).

Campbell and Galbraith (1996) proposed that balancing judging panels to include representatives from all medal-contending nations could allow for a reciprocal neutralization of national bias. However, empirical studies indicate that judges exhibit national bias to varying degrees (Heiniger & Mercier, 2019; Krumer et al., 2022; Lyngstad et al., 2020; Sala et al., 2007; Zitzewitz, 2006), which limits the effectiveness of reciprocal bias neutralization.

More reasonable propositions involve expanding the number of judges per panel to dilute national bias (Heiniger & Mercier, 2019; Landers, 1970; Myers et al., 2006; Soler, 2021; Zitzewitz, 2014). While this may also reduce the potential for collusion (Kirkbride, 2013), it raises economic costs (Myers et al., 2006) and increases the likelihood of co-national judging (Heiniger & Mercier, 2019; Soler, 2021), which might facilitate reciprocal bias compensation (Campbell & Galbraith, 1996) but also exacerbate distortions through indirect compensating or affirming bias.

#### 3.4.2. Exploiting judge demographics and psychological factors

Recent studies have explored individual judge characteristics as potential moderators of bias. Heiniger and Mercier (2021) reported that female judges exhibit less score variability and suggested they may be 'significantly more accurate than men judges in artistic gymnastics and in trampoline' (Heiniger & Mercier, 2021, pp. 2–3). However, the results could also reflect an aversion to providing outlier scores (Lee, 2008) rather than greater proficiency, as such aversion has been linked to the dominance, autonomy, and submissiveness personality traits (Scheer et al., 1983). Consequently, female judges may provide fewer outlier scores than male judges not due to higher accuracy, but because of a greater reluctance to deviate from the panel consensus, which may compromise independent evaluations as required by judging standards (FIG, 2022a, 2022b). Nonetheless, Heiniger and Mercier (2021) advocate for revisions in male judge selection, training, and evaluation processes in response to their findings.

Scholten et al. (2020) further suggested that younger judges may exhibit less patriotic bias due to reputational career concerns, although this could equally reflect age-related leniency. More broadly, career-related pressures may influence judging behaviors, leading some researchers to propose increased monitoring (FEI, 2018; Heiniger & Mercier, 2019; Scheer & Ansorge, 1987; Wettstone, 1968) and sanctions for judges with poor performance, including disqualification from future events (Heiniger & Mercier, 2019; Leskošek et al., 2012; Roetzheim & Muzyczko, 1986; Soler, 2021; Whissell et al., 1993). However, strict oversight may unintentionally increase affirming bias, as judges may align scores with their colleagues to avoid penalties (Brennan, 1996; Krumer et al., 2022; Lee, 2008).

To address career-related mechanisms underlying national bias, some scholars have proposed that judges should not be selected by national federations or event organizers, as such practices may incentivize biased judging as a signaling strategy to improve prospects for future appointments (Heuschmann,

2007; Kirkbride, 2013; Soler, 2021). Moreover, the centralization of judge selection and appointment processes may reduce the potential influence or pressure exerted by national federations and their leadership, particularly in federations with strong presidential authority (Brennan, 1996; Goodwin, 2004; Sala et al., 2007).

#### *3.4.3. Transparency and anonymity*

To combat vote trading, anonymous scoring has been implemented in some sports, but evidence suggests it is insufficient for reducing bias (Fang & Ho, 2024; Zitzewitz, 2014). In certain cases, anonymity may even encourage strategic or overt bias (Lee, 2008; Osório, 2017; Truchon, 2004), which suggests that this approach is rather superficial (Zitzewitz, 2014). By contrast, transparent scoring systems have been proposed as a countermeasure to bias (Roetzheim & Muzyczko, 1986; Weekley & Gier, 1989), as public score disclosure seems to reduce patriotic and affirming biases in ski jumping (Bouwens et al., 2022), while a survey among rhythmic gymnastics judges and fans suggests similar benefits (Soler, 2021). However, Fang and Ho (2024) report no similar effects in figure skating, although their study did not account for individual judge leniency.

#### *3.4.4. Enhancing scoring criteria and specialization*

Increasing the detail of scoring criteria has also been recommended to reduce subjectivity and the effects of political pressures (Soler, 2021; Wettstone, 1968). Similarly, task specialization, where judges evaluate only specific performance aspects, has been proposed to reduce cognitive load and improve scoring reliability (Bailie, 1965; Bar-Eli et al., 2011; Landers, 1970; Wolframm, 2023). Furthermore, reference judges have been introduced in some gymnastics disciplines to provide alternative reference scores. However, these reference judges have not demonstrated higher accuracy than regular judges in artistic gymnastics and their smaller numbers could intensify biases (Heiniger & Mercier, 2019, 2021), leading some researchers to advocate for their discontinuation (Soler, 2021).

Additionally, pre- (and post-)competition judges' meetings have been proposed as a means to foster a shared understanding of scoring criteria, thereby potentially reducing interpretative discrepancies among judges (Roetzheim & Muzyczko, 1986).

Furthermore, excluding judges from observing training sessions at competition venues has been proposed as a means of preventing unintentional patriotic bias due to the mere exposure mechanism (Ste-Marie, 1996, 2003; Ste-Marie et al., 2001; Ste-Marie & Lee, 1991).

Lastly, detailed checklists specifying intended elements may reduce cognitive load, thereby enhancing assessment accuracy and consistency while mitigating bias (Ste-Marie et al., 2001). However, such checklists may also amplify expectation effects, particularly when the listed elements vary in difficulty, potentially influencing the judges' assessments (Morgan & Rotthoff, 2014).

#### *3.4.5. Adjusting final score calculations*

Score truncation is another commonly proposed method to mitigate national bias (Bassett & Persky, 1994; Bring & Carling, 1994; Emerson, 2007; FEI, 2018; Heiniger & Mercier, 2019, 2021; Krumer et al., 2022; Landers, 1970; McFee, 2013; Russell, 2001; Sandberg, 2018; Whissell et al., 1993; Yang, 2006). However, aggressive truncation may discard valuable information due to high signal-to-noise ratios in judges' scores (Osório, 2017, 2020; Zitzewitz, 2006). Moreover, excessive truncation may have facilitated vote trading in figure skating, though causality remains unclear (Zitzewitz, 2006).

Furthermore, systematically excluding the scores of compatriot judges has been debated (Fenwick & Chatterjee, 1981). However, the exclusion of stricter judges might benefit their compatriots, while the removal of more lenient judges could disadvantage them, which leads to distortions and potential strategic adaptations. This risk also applies to the proposed random exclusion of some judges' scores (Kirkbride, 2013; McFee, 2013), as it may disproportionately benefit or penalize athletes, depending on the excluded judges' leniency (Emerson, 2007; Emerson & Arnold, 2011).

Mathematical models that adjust final scores based on national bias estimates and judge-specific factors such as leniency and proximity to average scores have also been proposed (Ansorge & Scheer, 1988; Osório, 2017, 2020; Roetzheim & Muzyczko, 1986; Scheer & Ansorge, 1987). However, these models face

limitations due to the variability of national bias across individual scoring cases (Campbell & Galbraith, 1996; Heiniger & Mercier, 2019, 2021; Zitzewitz, 2006), natural error margins in scores (Emerson et al., 2009) and potential strategic adaptations by judges (Scheer & Ansorge, 1987). Despite these limitations, analyses based on social choice theory (Balinski & Laraki, 2014; Truchon, 2004; Wu & Yang, 2004) have identified the lower-middlemost order function as a potential solution to further manipulation resistance (Balinski & Laraki, 2014).

#### *3.4.6. Technological support*

Technological advancements, such as electronic scoring systems and slow-motion replay technologies, have been proposed as potential solutions to enhance scoring accuracy and reduce bias (Allen et al., 2021; Díaz-Pereira et al., 2014; Myers et al., 2006; Sato & Hopper, 2021; Soler, 2021).

In addition, an unconventional approach involves integrating live audience reactions, such as cheering, into the scoring process (Kirkbride, 2013). However, this proposal raises concerns regarding potential biases linked to the nationality distribution of spectators and the influence of untrained judgments. Additionally, it may incentivize exaggerated spectator reactions as a strategic attempt to influence outcomes (Kirkbride, 2013). Spectators could, in effect, become proxy participants in the competition, attempting to sway results through amplified vocal or physical expressions of support, often along partisan lines.

## 4. Discussion

This scoping review was conducted to provide a comprehensive and critically informed synthesis of the existing literature on national bias in international sports judging. The substantial heterogeneity and fragmentation within this body of research—including the wide variation in statistical approaches used to examine different forms of national bias across diverse sports and competitive contexts, the divergent theoretical perspectives on underlying mechanisms and the question of intentionality, and the broad spectrum of proposed mitigation strategies—underscore the necessity of such a review. It adheres to the current PRISMA-2000 guidelines (Page et al., 2021) and is structured around four principal domains of inquiry: (1) the forms of national bias examined in the literature; (2) the mechanisms assumed to contribute to such bias; (3) the often assumed but debatable intentionality of national bias; and (4) the range of countermeasures proposed in the extant body of literature, including their potential limitations and unintended consequences.

A systematic database search supplemented by iterative reference tracking identified 68 eligible contributions, encompassing both empirical investigations and theoretical discussions. Empirical studies explored various forms of national bias across multiple sports and applied a diverse range of statistical methods. More recent studies yielded the most robust findings, owing to statistical advancements such as employing multiple linear regression modeling to account for confounding factors such as judge leniency. These studies predominantly focused on patriotic bias, with comparatively less attention paid to other manifestations such as competitor bias, political bias, vote trading, and indirect national bias in either its compensatory or affirming form.

Patriotic bias was consistently documented across a wide range of sports, with estimated magnitudes ranging from 13% to as much as half a within-performance standard deviation. Nominally larger magnitudes of patriotic bias are observed for top-ranked athletes or medal contenders (Campbell & Galbraith, 1996; Heiniger & Mercier, 2019, 2021). At the level of individual judges, the extent of patriotic bias varies considerably: while some judges exhibit strong bias in favor of their compatriots, others show minimal bias, and a minority even assign lower scores to compatriot athletes (Heiniger & Mercier, 2019; Sala et al., 2007; Sandberg, 2018). Furthermore, patriotic bias is the most widely acknowledged form of national bias and explicitly addressed in the official rules and regulations of several international sports federations (FEI, 2018, 2023; FIS, 2020, 2025a; ISU, 2024). It also frequently features in public discourse, particularly in mass media coverage of aesthetic sports, where controversial judging decisions or competition outcomes are discussed (Higuchi, 2022; Pender, 2024; Templon & Adams, 2018).

In contrast, evidence for competitor bias remains limited, with only one of six studies providing robust evidence of a minute competitor bias while accounting for judge leniency (Heiniger & Mercier, 2021). A probable reason for the limited findings is the common assumption that athletes on adjacent ranks are perceived as close competitors, which may not accurately reflect true competitive dynamics (Braeunig, 2024). Political bias, including bloc judging during the Cold War, has been suggested in some studies (Sala et al., 2007; Seltzer & Glass, 1991; Zitzewitz, 2006), although certain studies lack stringent statistical testing (Emerson et al., 2009), and cultural relationships also appear to influence cross-national scoring patterns (Fenwick & Chatterjee, 1981; Seltzer & Glass, 1991), potentially confounding the results on political bias. On the contrary, vote trading remains largely theoretical, as clear empirical evidence is missing. Nevertheless, instances of alleged collusion among judges—particularly in the form of reciprocal score exchanges—have been reported in mass media, with multiple accounts suggesting potential score manipulation (Higuchi, 2022). Lastly, indirect forms of national bias, including affirming bias in figure skating and dressage (Sandberg, 2018; Zitzewitz, 2006), and compensating bias in ski jumping (Zitzewitz, 2006), have been occasionally observed. However, findings related to indirect national bias remain partly inconsistent and two studies suggest that such bias diminishes when spectator behavior and home advantage are accounted for (Bouwens et al., 2022; Krumer et al., 2022), highlighting the need for further research and improved control of confounding variables in future research.

Regarding the mechanisms underlying national bias, the existing evidence suggests the involvement of both deliberate and unintentional processes. For instance, some findings indicate that patriotic bias tends to be more pronounced when evaluating top-ranked athletes or medal contenders (Campbell & Galbraith, 1996; Heiniger & Mercier, 2019, 2021), which may indicate a strategic and deliberate application of patriotic bias. This behavior is potentially driven by role conflicts or considerations related to judges' career advancement (Lyngstad et al., 2020; Sala et al., 2007; Zitzewitz, 2006). Supporting this interpretation are reports of pressure exerted on judges by national federations, including anecdotal accounts from within the sports community of judges being encouraged—or even coerced—into certain scoring behaviors (Brennan, 1996; Goodwin, 2004). These concerns have also been echoed in broader media coverage (Becker, 2020; Mackay, 2002). In contrast, other studies highlight unintentional or subconscious processes, including temporary cohesion among judges on a panel (Sandberg, 2018; Wolframm, 2023), culturally embedded stylistic preferences (Fenwick & Chatterjee, 1981; Seltzer & Glass, 1991), and cognitive biases such as the mere exposure effect and fluency heuristic (Ste-Marie, 1996; Ste-Marie & Lee, 1991). While the term *bias* is typically understood as implying intentional misconduct (Emerson et al., 2009), and some interpretations reduce national bias to simple expressions of nationalism (FEI, 2018), such perspectives likely oversimplify the phenomenon. On the contrary, research suggests that national bias forms due to an interplay of intentional as well as unintentional factors. Consequently, interventions such as sanctions against biased judges may fail to address the full complexity of the issue (Emerson & Meredith, 2011).

Although numerous countermeasures have been proposed to address national bias in sports judging, each entails specific feasibility constraints and trade-offs, as specified in Table A6 of the Appendix.

Structural modifications to judging panels—such as excluding co-national judges, balancing nationalities within judging panels, or increasing panel size—may limit direct favoritism, enable reciprocal bias neutralization, or dilute national bias. However, these approaches risk fostering juror collusion, demand more trained judges, increase costs, and may inadvertently exacerbate indirect national bias. Moreover, judge-level analyses reveal substantial variation in patriotic bias, suggesting that panel balancing may be insufficient and risks escalating a competition by proxy, wherein judges compete for the maximum tolerated level of bias (Braeunig, 2024).

Measures targeting judge demographics and psychological factors are relatively easy to implement but show limited empirical support. Enhanced monitoring and sanctions may deter biased behavior but can also trigger affirmation bias (Brennan, 1996; Krumer et al., 2022; Lee, 2008). Centralized judge selection, independent of national federations, may reduce career-driven signaling incentives and external pressures.

Transparency and anonymity policies are administratively simple. While anonymity can reduce perceived bias, it hinders public and scientific monitoring. Conversely, transparency facilitates monitoring

and enjoys broad support among both judges and spectators (Soler, 2021), though it may heighten bias perceptions (Soler, 2021).

Reforms to scoring procedures—such as refining criteria, enhancing judge specialization, or employing reference judges—can reduce subjectivity and cognitive load, but often require more resources. Similarly, judge briefings before and after competitions may improve interpretative consistency. Notably, such meetings are already mandated in the official regulations of numerous international sports federations (FIG, 2022a, 2022b, 2024; FIS, 2025b; ISU, 2024) in order to uphold interpretative alignment (FIG, 2025a). Restricting judges from observing athlete training may reduce unconscious patriotic bias. However, in some sports or disciplines, the observation of training sessions is considered a beneficial practice that aligns with established judging standards (FIG, 2022b; FIS, 2025b; World Aquatics, 2025), while others explicitly prohibit judges from attending podium training sessions prior to competition (FIG, 2025b). Checklists outlining required elements, as implemented in diving and high diving (World Aquatics, 2025) may reduce judges' cognitive load during scoring, thereby potentially mitigating nationality-based heuristic biases; however, such checklists may introduce expectation effects, particularly when the listed elements vary in difficulty, potentially influencing judges' anticipatory evaluations (Morgan & Rotthoff, 2014).

Scoring adjustments, including score truncation or exclusion of co-national scores, are easy to implement but present multiple concerns—such as data loss or the introduction of new biases due to differential judge leniency. Mathematical score corrections also face limitations due to score variability and the risk of strategic adaptation (Osório, 2017; Scheer & Ansorge, 1987).

Technological support systems, such as electronic judging systems, may reduce cognitive demands but are costly and limited in capturing aesthetic aspects of sports performances. Lastly, spectator-based scoring, while offering broader evaluative input, may amplify biases rooted in audience composition and behavior.

In sum, although several countermeasures have been implemented across various sports and may contribute to reducing national bias, its complete eradication appears unlikely due to the intricate interplay of structural, psychological, and contextual factors.

Lastly, certain limitations of this scoping review must be acknowledged. Of the 68 contributions analyzed, only 23 were identified through a systematic database search, while the remainder were obtained through iterative reference tracking. Notably, three studies that had been initially identified through exploratory searches were not retrieved through either the systematic search or reference tracking—two of which are recent publications unlikely to have yet appeared in citation networks. Additionally, 11 sources identified through reference tracking could not be accessed, including nine conference papers cited in a single included empirical study and two brief articles published in sport-specific magazines. Despite these constraints, this review compiles a diverse body of empirical studies, theoretical analyses, and grey literature spanning the period from 1965 to 2024, thereby offering a broad and comprehensive perspective on national bias in international sports judging.

Potential limitations also lie in the risk of publication bias—specifically, the underrepresentation of studies reporting null results (Ekmekci, 2017; Hubbard & Armstrong, 1997; Sharma & Verma, 2019). Nevertheless, several included studies report null findings: Three of the six studies investigating competitor bias, four of ten studies on political bias, and three of six studies on indirect forms of national bias found no significant effects. Regarding patriotic bias—the most frequently studied form—eight of 34 studies (at least partially) reported null results. While the overall tendency of published research may lean toward significant findings, the proportion of null results among the included studies is relatively high within this field. Nonetheless, the potential influence of publication bias cannot be entirely dismissed and may result in an overrepresentation of positive findings. Another methodological limitation concerns the synthesis of highly heterogeneous contributions. Differences in study designs, statistical approaches, sample sizes, and contextual confounding variables limited the extent to which methodological nuances and small-scale findings could be fully explored.

Despite these limitations, this review consolidates key empirical findings, theoretical perspectives, and future research directions in the study of national bias in international sports judging—a highly complex and multifaceted phenomenon. Its investigation requires careful methodological consideration of critical confounding factors, such as judge leniency, home advantage, audience behavior, and

sport-specific characteristics including competition formats, regulatory frameworks, and scoring systems. These complexities significantly impede definitive conclusions about the origins of national bias or the efficacy of proposed countermeasures, as numerous alternative explanations must be considered and not overlooked (Emerson et al., 2009; Emerson & Meredith, 2011; Lyngstad et al., 2020; Ste-Marie, 1996).

To address these challenges, future research should systematically investigate distinct forms of national bias and their underlying mechanisms across various sports, considering different competition structures, institutional norms, and socialization contexts. In particular, studies on competitor bias would benefit from exploring alternative operationalizations of competitive relationships to generate robust findings. Furthermore, systematic comparative analyses of sports with differing scoring systems and evaluative standards may offer valuable insights into the socio-psychological mechanisms driving national bias, particularly those shaped by institutional regulations that could be addressed by rule changes. Thus, comparative analyses of national bias across different sports may inform the development and implementation of more effective countermeasures and policy interventions. In this regard, 'control scores' (Heiniger & Mercier, 2021, p. 5)—performance assessments determined by international sports federations for training and evaluating judges—could serve as valuable, albeit currently inaccessible, empirical tools. These scores represent the closest approximation of a 'true performance score' (Heiniger & Mercier, 2019, p. 8) and could significantly enhance the validity of future research by offering a standardized point of reference, thus enabling more precise estimates of national bias. While some studies have employed performance-related metrics (e.g., jump distance and speed in ski jumping) as proxies for the true performance quality (Krumer et al., 2022; Zitzewitz, 2006), such data are not consistently available across sports, and concerns about potential model overfitting must be considered. Thus, access to control scores would substantially benefit the empirical investigation of national bias, its underlying mechanisms, and the effectiveness of proposed countermeasures.

Beyond these challenges, research on national bias presents a valuable opportunity to explore a range of compelling questions across multiple scholarly disciplines. These include social choice theory (Balinski & Laraki, 2014; Truchon, 2004; Wu & Yang, 2004), organizational decision-making (Zitzewitz, 2006, 2014), social cognition (Graf, 2010; Plessner & Haar, 2006), the role of heuristics in complex judging processes (Ste-Marie, 1996, 2003; Ste-Marie et al., 2001), and broader discussions on in-group favoritism, particularly in political and cultural contexts (Sandberg, 2018).

## Notes

1. In the respective studies, patriotic bias is quantified and reported in relative terms, specifically as standard deviation units of the scores assigned by all judges on the panel for the same performance (and, where applicable, for individual performance aspects that are evaluated separately).
2. Judges may even be sworn to this role by taking an oath, as it is the case in the men's and women's artistic gymnastics (see Fédération Internationale de Gymnastique [FIG], 2022a, p. 18; FIG, 2022b, Section 5 p. 4).
3. For instance, international figure skating judges are evaluated on their scoring practices, with specific attention given to detecting national bias. If national bias is confirmed, sanctions may be applied, ranging from temporary suspension of officiating duties to permanent revocation of licensure, depending on the gravity of the violation (ISU, 2021, 2024).

## Acknowledgments

The author gratefully acknowledges PD Dr. Werner Pitsch (Saarland University) for his valuable guidance on ensuring transparency and reproducibility of the screening and reporting process of systematic literature search results. Further appreciation is extended to Christiane Speicher, Head of Document Delivery and Digitalization at Saarland University and State Library, for her valuable and generous support, time, and expertise on navigating (international) literature search processes. Finally, the author acknowledges the constructive feedback provided by two anonymous reviewers, whose comments significantly contributed to the clarity and overall quality of this scoping review.

## Author contributions

CRediT: **Patrick Alexander Braeunig**: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing.

## Disclosure statement

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author received no financial support for the research, authorship, and/or publication of this article.

## About the author

***Patrick Alexander Braeunig***, M.Sc., is a Research Associate in the Department of Sport Sociology and Sport Economics at Saarland University. His research focuses on social cognition and decision-making in sports judging.

## ORCID

Patrick Alexander Braeunig http://orcid.org/0009-0000-6922-0632

## Data availability statement

Data sharing is not applicable to this article as no data were created or analysed in this study. The review protocol, PRISMA-2000 checklist and the extensional PRISMA-ScR checklist, full reference tracking documentation and funding documentation for all included contributions is provided through the Open Science Framework (https://osf.io/8u6cm/overview?view_only=c8a4298433c14190865c62cce6cbf7a8).

## References

Allen, E., Fenton, A., & Parry, K. (2021). Computerized gymnastics judging scoring system implementation – An exploration of stakeholders' perceptions. *Science of Gymnastics Journal*, *13*(3), 357–370. https://doi.org/10.52165/sgj.13.3.357-370

Ansorge, C. J., & Scheer, J. K. (1984). Biasing factors affecting judging of gymnastics. *Technique*, *4*(2), 19–21.

Ansorge, C. J., & Scheer, J. K. (1988). International bias detected in judging gymnastic competition at the 1984 Olympic Games. *Research Quarterly for Exercise and Sport*, *59*(2), 103–107. https://doi.org/10.1080/02701367.1988.10605486

Bailie, S. (1965). We must have a new scoring system. *Modern Gymnast*, *7*(3), 12–13.

Balinski, M., & Laraki, R. (2014). Judge: Don't vote!. *Operations Research*, *62*(3), 483–511. https://doi.org/10.1287/opre.2014.1269

Ball, D. W. (1973). A politicized social psychology of sport: Some assumptions and evidence from international figure skating competition. *International Review of Sport Sociology*, *8*(3), 63–71. https://doi.org/10.1177/101269027300800304

Bar-Eli, M., Plessner, H., & Raab, M. (2011). *Judgement, decision making and success in sport*. John Wiley & Sons. https://doi.org/10.1002/9781119977032

Bassett, G. W., & Persky, J. (1994). Rating skating. *Journal of the American Statistical Association*, *89*(427), 1075–1079. https://doi.org/10.2307/2290937

Becker, L. (2020). *I call foul play: Figure skating is not a legitimate sport*. The Daily Free Press. https://dailyfreepress.com/02/18/21/164317/i-call-foul-play-figure-skating-is-not-a-legitimate-sport/

Berry, S. M. (2002). Gold medals, protests, and trading ratings: The 2002 Winter Olympics figure skating. *CHANCE*, *15*(2), 14–18. https://doi.org/10.1080/09332480.2002.10542322

Bouwens, J., Hofmann, C., & Lechner, C. (2022). Transparency and biases in subjective performance evaluation. *TRR 266 Accounting for Transparency Working Paper Series*. https://doi.org/10.2139/ssrn.4012905

Braeunig, A. P. (2024). Diskretionäre Spielräume und wie sie genutzt werden – Analysen zur nationalen Parteilichkeit von Juroren im Eiskunstlaufen. In M. Frenger, J. Haut, M. Klein, & L. Thieme (Eds.), *Artenvielfalt im Biotop der Wissenschaft. Sozioökonomische Perspektiven auf Sport und Gesellschaft* (pp. 251–285). Universaar.

Brennan, C. (1996). *The inside edge: A revealing journey into the secret world of figure skating*. Scribner.

Bring, J., & Carling, K. (1994). A paradox in the ranking of figure skaters. *CHANCE*, *7*(4), 34–37. https://doi.org/10.1080/09332480.1994.10542440

Bruine de Bruin, W. (2005). Save the last dance for me: Unwanted serial position effects in jury evaluations. *Acta Psychologica*, *118*(3), 245–260. https://doi.org/10.1016/j.actpsy.2004.08.005

Bruine de Bruin, W. (2006). Save the last dance II: Unwanted serial position effects in figure skating judgments. *Acta Psychologica*, *123*(3), 299–311. https://doi.org/10.1016/j.actpsy.2006.01.009

Callahan, B. P., Mulholland, S. E., & Rotthoff, K. W. (2016). Cultural bias: Gymnasts, judges, and bilateral trade agreements. *Journal of SPORT*, *5*(1), 25–45. https://doi.org/10.2139/ssrn.2140469
Campbell, B., & Galbraith, J. (1996). Nonparametric tests of the unbiasedness of Olympic figure-skating judgments. *Journal of the Royal Statistical Society. Series D (The Statistician)*, *45*(4), 521–526. https://doi.org/10.2307/2988550
Chien, P. M., Kelly, J. K., & Weeks, C. S. (2016). Sport scandal and sponsorship decisions: Team identification matters. *Journal of Sport Management*, *30*(5), 490–505. https://doi.org/10.1123/jsm.2015-0327
Criley, D. (1976). XXI Olympiad-questions, and answers. *International Gymnast*, *18*(9), 10–11.
Criley, D. (l972). Viewpoints. *Gymnast*, *15*(11), 8–9.
Daumann, F. (2015). *Grundlagen der Sportökonomie* (2nd ed.). UVK Verlag. https://doi.org/10.36198/9783838542898
Deuel, N. R. (1989). Nationalism in the scoring of dressage competitions at the 1988 Olympics. *Journal of Equine Veterinary Science*, *9*(5), 281–283. https://doi.org/10.1016/S0737-0806(89)80091-7
Díaz-Pereira, M. P., Gómez-Conde, I., Escalona, M., & Olivieri, D. N. (2014). Automatic recognition and scoring of olympic rhythmic gymnastic movements. *Human Movement Science*, *34*, 63–80. https://doi.org/10.1016/j.humov.2014.01.001
Dosseville, F., Laborde, S., & Garncerzyk, C. (2014). Current research in sports officiating and decision-making. In C. Mohiyeddini (Ed.), *Contemporary topics and trends in the psychology of sports* (pp. 13–38). Nova Science Publishers.
Ekmekci, P. E. (2017). An increasing problem in publication ethics: Publication bias and editors' role in avoiding it. *Medicine, Health Care, and Philosophy*, *20*(2), 171–178. https://doi.org/10.1007/s11019-017-9767-0
Emerson, J. W. (2007). Chance, on and off the ice. *CHANCE*, *20*(2), 19–21. https://doi.org/10.1080/09332480.2007.10722841
Emerson, J. W., & Arnold, T. B. (2011). Statistical sleuthing by leveraging human nature: A study of Olympic figure skating. *American Statistician*, *65*(3), 143–148. https://doi.org/10.1198/tast.2011.10165
Emerson, J. W., & Meredith, S. (2011). Nationalistic judging bias in the 2000 olympic diving competition. *Math Horizons*, *18*(3), 8–11. https://doi.org/10.4169/194762111X12954578042812
Emerson, J. W., Seltzer, M., & Lin, D. (2009). Assessing judging bias: An example from the 2000 olympic games. *American Statistician*, *63*(2), 124–131. https://doi.org/10.1198/tast.2009.0026
Emrich, E., & Pierdzioch, C. (2015). Theoretische Rahmung. In E. Emrich, C. Pierdzioch, & W. Pitsch (Eds.), *Falsches Spiel im Sport. Analysen zu Wettbewerbsverzerrung* (pp. 15–44). Universaar.
Fang, X., & Ho, C. Y. Verein für Socialpolitik/German Economic Association. (2024). The effect of transparency on subjective evaluations: Evidence from competitive figure skating. In *Beiträge zur Jahrestagung des Vereins für Socialpolitik 2024: Upcoming Labor Market Challenges*. ZBW – Leibniz Information Centre for Economics. https://www.econstor.eu/bitstream/10419/ 302427/1/vfs-2024-pid-108000.pdf
Fédération Équestre Internationale (FEI). (2018). *Dressage judging working group report of the FEI dressage judging working group.* https://inside.fei.org/system/files/ FinalReportDJWG.pdf
Fédération Équestre Internationale (FEI). (2023). *Dressage rules.* https://inside.fei.org/sites/ default/files/FEI_Dressage_Rules_2025_Clean_Version_Updated_18_June_2025.pdf
Fédération Internationale de Gymnastique (FIG). (2022a). *2022-2024 code of points men's artistic gymnastics*. https://www.gymnastics.sport/publicdir/rules/files/en_% 202022-2024%20MAG%20CoP.pdf
Fédération Internationale de Gymnastique (FIG). (2022b). *2022-2024 code of points women's artistic gymnastics.* https://www.gymnastics.sport/publicdir/rules/ files/en_2022-2024%20WAG%20COP.pdf
Fédération Internationale de Gymnastique (FIG). (2024). *Rhythmic gymnastics code of points 2025-2028*. https://www.gymnastics.sport/publicdir/rules/files/en_ 1.1%20-%20RG%20Code%20of%20Points%202025-2028%20-%20valid%20from%201st%20April%202025.pdf
Fédération Internationale de Gymnastique (FIG). (2025a). *Technical regulations 2025.* https://www.gymnastics.sport/publicdir/rules/files/en_1.1%20-%20Technical% 20Regulations%202025.pdf
Fédération Internationale de Gymnastique (FIG). (2025b). *FIG general judges' rules.* https://www.gymnastics.sport/publicdir/rules/files/en_1.2%20-%20General% 20Judges'%20Rules%202025-2028.pdf
Fédération Internationale de Ski (FIS). (2020). *Guidelines for FIS jumping judges.* https://assets.fis-ski.com/f/252177/1a0a773d6e/guidelines-for-jumping-judges_ 2020.pdf
Fédération Internationale de Ski (FIS). (2025a). *The international snowboard/freestyle/freeski/ski cross competition rules (ICR). Book vi. Joint regulations for snowboard/freestyle ski/freeski.* https://assets.fis-ski.com/f/252177/x/b5f88085ce/sbfsfk-new-icr-spring-2025_ clean.pdf
Fédération Internationale de Ski (FIS). (2025b). *Snowboard & Freeski park & pipe judges handbook.* https://assets.fis-ski.com/f/252177/7c81eac52f/ fis_sb_fk-judgeshandbook_update_spring-2022.pdf
Fenwick, I., & Chatterjee, S. (1981). Perception, preference, and patriotism: An exploratory analysis of the 1980 winter olympics. *American Statistician*, *35*(3), 170–173. https://doi.org/10.2307/2683998
Goodwin, J. (2004). *The second mark: Courage, corruption, and the battle for olympic gold*. Simon and Schuster.
Graf, K. (2010). *Der Reihenfolge-Effekt bei der Beurteilung von schnellen Bewegungen am Beispiel von Kampfrichterurteilen im Gerätturnen* [Doctoral dissertation]. Universität Stuttgart. https://elib.uni-stuttgart.de/bitstream/11682/5530/1/Dissertation_KG_Publikation.pdf
Gymmedia International. (2004). *Wertungseklat: Arkajew spricht von Mafia der Kampfrichter.* Gymmedia International. https://www.gymmedia.de/ Geraetturnen/Wertungseklat-Arkajew-spricht-von-Mafia-der-Kampfrichter

Hawson, L. A., McLean, A. N., & McGreevy, P. D. (2010). Variability of scores in the 2008 olympic dressage competition and implications for horse training and welfare. *Journal of Veterinary Behavior*, *5*(4), 170–176. https://doi.org/10.1016/j.jveb.2009.12.010

Heiniger, S., & Mercier, H. (2019). *National bias of International Gymnastics Judges during the 2013-2016 olympic cycle*. ArXiv. https://arxiv.org/abs/1807.10033

Heiniger, S., & Mercier, H. (2021). Judging the judges: Evaluating the accuracy and national bias of international gymnastics judges. *Journal of Quantitative Analysis in Sports*, *17*(4), 289–305. https://doi.org/10.1515/jqas-2019-0113

Heuschmann, G. (2007). *Tug of war: Classical versus "modern" dressage*. Trafalgar Square.

Higuchi, K. (2022). *Timeline of figure skating controversies from 1902 to 2022*. CNN. https://edition.cnn.com/2022/02/19/sport/figure-skating-controversies-timeline-spt-intl

Hill, R., & Barton, R. (2005). Red enhances human performance in contests. *Nature*, *435*(7040), 293–293. https://doi.org/10.1038/435293a

Hubbard, R., & Armstrong, J. S. (1997). Publication bias against null results. *Psychological Reports*, *80*(1), 337–338. https://doi.org/10.2466/pr0.1997.80.1.337

International Skating Union (ISU). (2021). *Communication 2423*. Rules of Procedure for Officials Assessment Commission – Evaluation of Judging – Assessments for the Figure Skating Branch. https://www.isu.org/figure-skating/rules/fsk-communications/26068-isu-communication-2423/file

International Skating Union (ISU). (2024). *Special regulations & technical rules*. Single & pair Skating and Ice Dance 2024. https://www.isu.org/figure-skating/rules/fsk-regulations-rules/file

Kestnbaum, E. (2003). *Culture on ice: Figure skating and cultural meaning*. Wesleyan University Press. https://archive.org/details/cultureonicefigu00kest/page/n5/ mode/2up

Kirkbride, A. N. (2013). Scoring/judging applications. In T. McGarry, P. O'Donoghue, & J. Sampaio (Eds.). *Routledge handbook of sports performance analysis* (pp. 140–152). Routledge.

Klein, M., Frenger, M., & Sommer, T. (2014). Interrater-Reliabilität der Punktrichterentscheidungen im Taekwondo-Formenlauf. In S. Liebl, & P. Kuhn (Ed.*), Menschen im Zweikampf – Kampfkunst und Kampfsport in Forschung und Lehre 2013. 3. Internationales Symposium „Kampfkunst und Kampfsport" vom 7.-9. November 2013 in Erlangen* (pp. 209–215). Feldhaus.

Krumer, A., Otto, F., & Pawlowski, T. (2022). Nationalistic bias among international experts: Evidence from professional ski jumping. *Scandinavian Journal of Economics*, *124*(1), 278–300. https://doi.org/10.1111/sjoe.12451

Landers, D. M. (1970). A review of research on gymnastic judging. *Journal of Health, Physical Education, Recreation*, *41*(7), 85–88. https://doi.org/10.1080/00221473.1970.10610644

Lee, J. (2008). Outlier aversion in subjective evaluation: Evidence from world figure skating championships. *Journal of Sports Economics*, *9*(2), 141–159. https://doi.org/10.1177/1527002507299203

Leskošek, B., Čuk, I., Pajek, J., Forbes, W., & Bučar-Pajek, M. (2012). Bias of judging in men's artistic gymnastics at the European Championship 2011. *Biology of Sport*, *29*(2), 107–113. https://doi.org/10.5604/20831862.988884

Looney, M. A. (2004). Evaluating judge performance in sport. *Journal of Applied Measurement*, *5*(1), 31–47.

Lyngstad, T. H., Härkönen, J., & Rønneberg, L. T. S. (2020). Nationalistic bias in sport performance evaluations: An example from the ski jumping world cup. *European Journal for Sport and Society*, *17*(3), 250–264. https://doi.org/10.1080/16138171.2020.1792628

Mackay, D. (2002). *French officials banned over gold scandal*. The Guardian. https://www.theguardian.com/sport/2002/may/01/olympicgames.winterolympics2002

McFee, G. (2013). Officiating in aesthetic sports. *Journal of the Philosophy of Sport*, *40*(1), 1–17. https://doi.org/10.1080/00948705.2012.725910

Moreland, R. L., & Zajonc, R. B. (1977). Is stimulus recognition a necessary condition for the occurence of exposure effects? *Journal of Personality and Social Psychology*, *35*(4), 191–199. https://doi.org/10.1037/0022-3514.35.4.191

Morgan, H. N., & Rotthoff, K. W. (2014). The harder the task, the higher the score: Findings of a difficulty bias. *Economic Inquiry*, *52*(3), 1014–1026. https://doi.org/10.1111/ecin.12074

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, *18*(1), 143. https://doi.org/10.1186/s12874-018-0611-x

Myers, T. D., Balmer, N. J., Nevill, A. M., & Nakeeb, Y. A. (2006). Evidence of nationalistic bias in Muaythai. *Journal of Sports Science & Medicine*, *5*(CSSI), 21–27.

Osório, A. (2017). Judgement and ranking: Living with hidden bias. *Annals of Operations Research*, *253*(1), 501–518. https://doi.org/10.1007/s10479-016-2341-y

Osório, A. (2020). Performance evaluation: Subjectivity, bias and judgment style in sport. *Group Decision and Negotiation*, *29*(4), 655–678. https://doi.org/10.1007/s10726-020-09672-4

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, *10*(1), 89. https://doi.org/10.1186/s13643-021-01626-4

Pender, K. (2024). *Australian surfing judge removed from Olympics over photo with Ethan Ewing*. The Guardian. https://www.theguardian.com/sport/ article/2024/aug/02/australian-surfing-judge-ben-lowe-removed-olympics-2024-teahupoo-tahiti

Peters, M. D. J., Marnie, C., Colquhoun, H., Garritty, C. M., Hempel, S., Horsley, T., Langlois, E. V., Lillie, E., O'Brien, K. K., Tunçalp, Ö., Wilson, M. G., Zarin, W., & Tricco, A. C. (2021). Scoping reviews: Reinforcing and advancing the methodology and application. *Systematic Reviews*, *10*(1), 263. https://doi.org/10.1186/s13643-021-01821-3

Plessner, H., & Haar, T. (2006). Sports performance judgments from a social cognitive perspective. *Psychology of Sport and Exercise*, *7*(6), 555–575. https://doi.org/10.1016/j.psychsport.2006.03.007

Popović, R. (2000). International bias detected in judging rhythmic gymnastics competition at Sydney-2000 Olympic Games. *Facta universitatis series*. *Physical Education and Sport*, *1*(7), 1–13.

Richards, D. (2006). Protocol-driven search strategies and systematic reviews of complex evidence. *Evidence-Based Dentistry*, *7*(1), 19–19. https://doi.org/10.1038/sj.ebd.6400383

Roetzheim, B., & Muzyczko, T. (1986). Bias correction factors: A proposal to minimize unwanted pattern bias in international competitions. *Technique*, *6*(1), 14–17.

Rotthoff, K. W. (2015). (Not finding a) sequential order bias in elite level gymnastics. *Southern Economic Journal*, *81*(3), 724–741. https://doi.org/10.2139/ssrn.2230038

Russell, G. W. (2001). *Sport science secrets: From myth to facts*. Trafford Publishing.

Sala, B. R., Scott, J. T., & Spriggs, J. F. (2007). The cold war on ice: Constructivism and the politics of olympic skating judging. *Perspectives on Politics*, *5*(1), 17–29. https://doi.org/10.1017/S153759270707003X

Sampaio, B. R. (2012). *Three essays on applied microeconomics* [Doctoral dissertation]. University of Illinois. https://core.ac.uk/download/pdf/4837557.pdf

Sandberg, A. (2018). Competing identities: A field study of in-group bias among professional evaluators. *Economic Journal*, *128*(613), 2131–2159. https://doi.org/10.1111/ecoj.12513

Sato, N., & Hopper, L. S. (2021). Judges' evaluation reliability changes between identifiable and anonymous performance of hip-hop dance movements. *PloS One*, *16*(1), e0245861. https://doi.org/10.1371/journal.pone.0245861

Scheer, J. K., & Ansorge, C. J. (1987). Test of a model scoring system for the self-regulation of international bias in gymnastic judging. *Technique*, *7*(2), 4–7.

Scheer, J. K., Ansorge, C. J., & Howard, J. (1983). Judging bias induced by viewing contrived videotapes: A function of selected psychological variables. *Journal of Sport Psychology*, *5*(4), 427–437. https://doi.org/10.1123/jsp.5.4.427

Scholten, H., Schneemann, S., & Deutscher, C. (2020). The impact of age on nationality bias: Evidence from ski jumping. *Journal of Institutional and Theoretical Economics*, *176*(4), 708–734. https://doi.org/10.1628/jite-2020-0040

Seltzer, R., & Glass, W. (1991). International politics and judging in Olympic skating events. *Journal of Sport Behaviour*, *14*(3), 189–200.

Sharma, H., & Verma, S. (2019). Is positive publication bias really a bias, or an intentionally created discrimination toward negative results? *Saudi Journal of Anaesthesia*, *13*(4), 352–355. https://doi.org/10.4103/sja.SJA_124_19

Soler, M. P. (2021). *Evolution of nationalistic bias in rhythmic gymnastics judging at different competition levels* [Bachelor's Thesis]. ESCI-UPF. https:// repositori.upf.edu/bitstream/handle/10230/52343/TFG2021GNMlMartinezEvolution.pdf?sequence=1&isAllowed=y

Sorokowski, P., Szmajke, A., Hamamura, T., Jiang, F., & Sorokowska, A. (2014). "Red wins", "black wins" and "blue loses" effects are in the eye of beholder, but they are culturally universal: A cross-cultural analysis of the influence of outfit colours on sports performance. *Polish Psychological Bulletin*, *45*(3), 318–325. https://doi.org/10.2478/ppb-2014-0039

Ste-Marie, D. M. (1996). International bias in gymnastic judging: Conscious or unconscious influences? *Perceptual and Motor Skills*, *83*(3), 963–975. https://doi.org/10.2466/pms.1996.83.3.963

Ste-Marie, D. M. (1999). Expert-novice differences in gymnastic judging: An information-processing perspective. *Applied Cognitive Psychology*, *13*(3), 269–281. 3.0.CO;2-Y https://doi.org/10.1002/(SICI)1099-0720(199906)13:3<269::AID-ACP567>

Ste-Marie, D. M. (2003). Memory biases in gymnastic judging: Differential effects of surface feature changes. *Applied Cognitive Psychology*, *17*(6), 733–751. https://doi.org/10.1002/acp.897

Ste-Marie, D. M., & Lee, T. D. (1991). Prior processing effect on gymnastic judging. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(1), 126–136. https://doi.org/10.1037/0278-7393.17.1.126

Ste-Marie, D. M., Valiquette, S. M., & Taylor, G. (2001). Memory-influenced biases in gymnastic judging occur across different prior processing conditions. *Research Quarterly for Exercise and Sport*, *72*(4), 420–426. https://doi.org/10.1080/02701367.2001.10608979

Stern, M. (2010). *Stil-Kulturen. Performative Konstellationen von Technik, Spiel und Risiko in neuen Sportpraktiken*. Transcript.

Tang, L. J. (2013). The role of nationalism in the olympics: Reflecting on the 2012 London Games. *Sociological Research Online*, *18*(2), 113–117. https://doi.org/10.5153/sro.2976

Templon, J., & Adams, R. (2018). *Two Chinese skating judges have been suspended for favoring their country's skaters*. BuzzFeed News. https://www.buzzfeednews.com/article/johntemplon/figure-skating-authorities-suspend-two-chinese-judges-for

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

Truchon, M. (2004). *Aggregation of rankings in figure skating*. CIRPEE Working Paper No. 04-14. https://doi.org/10.2139/ssrn.559981

von Elm, E., Schreiber, G., & Haupt, C. C. (2019). Methodische Anleitung für Scoping Reviews (JBI-Methodologie). *Journal of Evidence and Quality in Health Care*, *143*, 1–7. https://doi.org/10.1016/j.zefq.2019.05.004

Weekley, J. A., & Gier, J. A. (1989). Ceilings in the reliability and validity of performance ratings: The case of expert raters. *Academy of Management Journal*, *32*(1), 213–222. https://doi.org/10.2307/256428

Wettstone, G. (1968). The judging at Mexico City. *Modern Gymnast*, *10*(11&12), *32*.

Whissell, R., Lyons, S., Wilkinson, D., & Whissell, C. (1993). National bias in judgments of olympic-level skating. *Perceptual and Motor Skills*, *77*(2), 355–358. https://doi.org/10.2466/pms.1993.77.2.355

Wirtz, M., & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. In *Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Hogrefe.

Wolframm, I. (2023). Let them be the judge of that: Bias cascade in elite dressage judging. *Animals: An Open Access Journal from MDPI*, *13*(17), 2797. https://doi.org/10.3390/ani13172797

World Aquatics. (2025). *Competition regulations.* https://resources.fina.org/ fina/document/2025/07/01/ ed3110a4-2291-411d-8526-6f641bd9237a/ Competition-Regulations_June-2025_Clean-updated-01.07.2025-.pdf

Wu, S., & Yang, M. (2004). Evaluation of the current decision rule in figure skating and possible improvements. *American Statistician*, *58*(1), 46–54. https://doi.org/10.1198/0003130042818

Yang, M. (2006). *Cultural biases in olympic figure skating judgments* [Senior Thesis]. Northwestern University. https://mmss.wcas.northwestern.edu/thesis/articles/ get/591/Yang%202006.pdf

Zitzewitz, E. (2006). Nationalism in winter sports judging and its lessons for organisational decision making. *Journal of Economics & Management Strategy*, *15*(1), 67–99. https://doi.org/10.1111/j.1530-9134.2006.00092.x

Zitzewitz, E. (2014). Does transparency reduce favoritism and corruption? Evidence from the reform of figure skating judging. *Journal of Sports Economics*, *15*(1), 3–30. https://doi.org/10.1177/1527002512441479

## Appendix

**Table A1.** *Web of Science* search

**Initial Search Strategy:**

**All Fields:** Sport AND (judg* OR rat* OR scor*) AND ((Bias AND (National* OR Patriot*)) OR Error OR Corrupt* OR Fav* OR Patriot*)

**Systematic and consecutive Exclusion of false positives:**

**All Fields:** NOT (Health OR Injur* OR Physic* OR Clinic* OR Illness OR Concuss* OR Patient* OR Body OR Heart OR joint OR neuro* OR Biomechanic* OR Mass OR Material* OR Structure OR Metabol* OR Speech OR Consum* OR Nutri* OR Protein* OR Orthopedic OR Operation OR Child* OR Stress OR Disease OR Trauma OR Fracture OR ACL OR Arthroscopi* OR Anxiety OR Learning OR Balance OR Algorithm OR Endurance OR Questionnaire OR Survey OR Track* OR Specie* OR Player* OR Baseball OR Soccer OR Hockey OR Volleyball OR Tennis OR Rugby OR Cricket OR Cycl* OR Shoot* OR shot* OR Fish* OR sensor OR Radio* OR Antenna OR evolution OR Hydro* OR linguist* OR COVID OR Syndrome OR Disorder* OR Resilience OR Doping OR Accident OR History OR Stocl OR Plan* OR Sex OR Mental OR Domain OR Transmission OR Sampling OR kine* OR Customer* OR Muscle OR Touris* OR Golf* OR Genetic* OR aerodynamic* OR Runner* OR Vehicle OR Smartphone OR Polyurethane OR audio* OR iso* OR Companie* OR atmospher* OR Handball)

**Abstract:** NOT (Peer OR Nazi OR earlobe OR Mesocycles OR distillation OR shore OR poker OR host OR mascots OR crystals OR inhibit* OR CPI OR RM OR NSO OR USD OR CRT OR ARFIMA OR VWM OR X-ray OR VO2max OR NCAA OR Leader* OR Marriage OR Electromagnet* OR Affordances OR Visuospatial OR Breaststroke OR Damage OR Culture OR Elo OR motor* OR catch* OR Predicti* OR Utility OR Industry OR medic* OR Diabet* OR power* OR Fenc* OR Prism* OR News OR Entropy OR Intelligence OR Laparoscop* OR Football OR Bio* OR Metal OR Load OR Bone OR Geographic OR Financ* OR Lactate OR Business OR School OR GPS OR Suicide OR Mone* OR Forecasting OR Coding OR Network* OR Blown OR allyship OR tactics OR white OR Index OR Australia OR Self)

**Title:** NOT (2-D OR Landing Error Scoring System OR Ratio Variables OR Anonymous Juries OR Push-Pull OR Treadmill OR Weight Change OR Intuition)

**Table A2.** *PubMed* search

**†Initial Search Strategy:**

**All Fields:** Sport AND (judg* OR rati* OR rate* OR scor*) AND ((Bias AND (National* OR Patriot*)) OR Error OR Corrupt* OR Favo* OR Patriot*)

†Initial Search Strategy had to be slightly altered as PubMed does not support less than four letter search terms with the use of an asterix, as in the original „rat*“ and „Fav*“ Terms

**Systematic and consecutive Exclusion of false positives:**

**All Fields:** NOT (Health OR Injur* OR Clinic* OR Concuss* OR Patient* OR Body OR Heart OR Joint OR Neuro* OR Biomechanic* OR Metabol* OR Nutri* OR Protein* OR Consum* OR Orthopedic OR Operation OR Disease OR Trauma OR ACL OR Syndrome OR Disorder* OR Accident OR Muscle OR Genetic* OR Child* OR Stress OR Anxiety OR Resilience OR Doping OR COVID OR Structure OR Material* OR Mass OR Mental OR Endurance OR Learning OR Balance OR Survey OR Algorithm OR Player* OR Track* OR Baseball OR Soccer OR Volleyball OR Tennis OR Cricket OR Shoot* OR Shot* OR Fish* OR Golf* OR Runner* OR Touris* OR Sensor OR Evolution OR Linguist* OR History OR Plan* OR Smartphone OR Radio* OR Specie* OR Sex OR Domain OR Transmission OR Sampling OR atmospher* OR Speech OR Walking OR Crossfit OR Swimming OR Dogs OR Sleep OR Healing OR AI OR Diagnos* OR Oculo* OR Screen OR Trajector* OR Fitness OR Hamstring OR Microscop* OR Aerob* OR Rowing OR Intensity OR Motor OR Boxer OR Vaso*)

**Title/Abstract:** NOT (Strength[Title/Abstract]) NOT (Lipoprotein[Title/Abstract]) NOT (Football[Title/Abstract]) NOT (Sprint[Title/Abstract]) NOT (Organizational[Title/Abstract]) NOT (Wahoo[Title/Abstract]) NOT (Quad*[Title/Abstract]) NOT (GPS[Title/Abstract]) NOT (Psychophysics[Title/Abstract]) NOT (Suggestibility[Title/Abstract]) NOT (devices[Title/Abstract]) NOT (committee[Title/Abstract]) NOT (scientists[Title/Abstract]) NOT (self[Title/Abstract]) NOT (medicine[Title/Abstract])

**Table A3.** *PsychArticles* search

**Initial Search Strategy:**
**All Fields:** Sport AND (judg* OR rat* OR scor*) AND ((Bias AND (National* OR Patriot*)) OR Error OR Corrupt* OR Fav* OR Patriot*)
**Systematic and consecutive Exclusion of false positives:**
**All Fields:** NOT (Injur* OR Clinic* OR Patient* OR Illness OR OR Trauma OR Heart OR Neuro* OR Anxiety OR Disorder* OR Syndrome OR Resilience OR Health OR Speech OR Stress OR Child* OR Body OR joint OR Mass OR linguist* OR Accident OR Operation OR Student OR Domain OR Transgression OR Hockey OR Baseball OR Framing OR Bayes Objects OR Covariance OR exploratory OR "four-culture" OR Calibration OR "forced-choice")
Search mode used: "Proximity". Search settings were set to apply related words, also search within the full text of the articles and to apply equivalent subjects. Contributions in a language other than English were excluded as well as contributions with no full text available

**Table A4.** *PsychInfo* search

**Initial Search Strategy:**
**All Fields:** Sport AND (judg* OR rat* OR scor*) AND ((Bias AND (National* OR Patriot*)) OR Error OR Corrupt* OR Fav* OR Patriot*)
**Systematic and consecutive Exclusion of false positives:**
**All Fields:** NOT (Injur* OR Clinic* OR Patient* OR Illness OR Trauma OR Heart OR Neuro* OR Anxiety OR Syndrome OR Resilience OR Health OR Speech OR Child* OR Body OR joint OR Mass OR linguist* OR Accident OR Student OR Structure OR Algorithm OR Survey OR History OR Metabol* OR Material* OR Biomechanic* OR Endurance OR Balance OR sensor OR Radio* OR COVID OR Doping OR Plan* OR Sex OR Mental* OR Domain OR Sampling OR Customer* OR Muscle OR Touris* OR Genetic* OR Vehicle OR audio* OR Companie* IR iso* OR Track* OR Specie* OR Player* OR Baseball OR Soccer OR Hockey OR Volleyball OR Tennis OR Rugby OR Cricket OR Cycl* OR Shoot* OR shot* OR Golf* OR Handball OR Football OR FIFA OR inhibit* OR poker OR host OR NCAA OR Leader* OR Affordances OR Visuospatial OR Culture OR medic* OR power* OR News OR Load OR Business OR Suicide OR Forecasting OR Erratum)
**Abstract:** NOT (Adolescent* OR Abuse OR ESport* OR Commentator* OR teaching OR substance OR Swimmer OR sponsor OR coping OR visuomotor OR ARFIMA OR employee OR „Air Force" OR Stroop OR Dissocia* OR Doctrine OR Illusory OR sorting* Fitts Or momentum Noncontingent OR spatial OR correction OR Canadian OR Brasile OR German OR practitioner OR nonfan OR acquisition OR fan OR Service)
Search mode used: "Proximity". Search settings were set to apply related words, also search within the full text of the articles and to apply equivalent subjects. Contributions in a language other than English were excluded as well as contributions with no full text available

**Table A5.** *ScienceDirect* search

**Initial Search Strategy:**
**All Fields:** Sport (judge OR rate OR score) (Bias OR Error) NOT Health NOT Child NOT Patient NOT Ocean NOT Learning
**Systematic and consecutive Exclusion of false positives:**
**All Fields:** (judge OR rate OR score) (Bias OR Error) NOT Health NOT Child NOT Patient NOT Ocean NOT Learning
**Title, abstract or author-specified keywords:** *"Sport"** NOT (Pain OR Network OR Sale OR Corporate OR Organism OR Biology OR Algorithm)
*As ScienceDirect allows for a maximum of only eight Boolean operators, an alternative search strategy was employed. The Term *"Sport"* was substituted by concrete sports in subsequent separate searches. The number of articles identified are presented in parentheses.
**Substitutes used for the term "*Sport*":**

| | | | | |
|---|---|---|---|---|
| "Figure Skating" (n = 5) | "Synchronized Skating" (n = 0) | "Synchronized Swimming" (n = 1) | "Artistic Swimming" (n = 0) | "Ski Jumping" (n = 3) |
| Snowboard (n = 7) | "Freestyle Ski" (n = 0) | "Free Ski" (n = 0) | "Ski Freestyle" (n = 0) | Freeski (n = 0) |
| Equestrian (n = 5) | "Hip Hop" (n = 7) | Parcour (n = 1) | "Break Dance" (n = 0) | Taekwondo (n = 4) |
| Wushu (n = 0) | Karate (n = 5) | Judo (n = 9) | Boxing (n = 7) | Gymnastics (n = 12) |
| Surfing (n = 36) | Diving (n = 73) | Dance (n = 63) | Dressage (n = 1) | Skateboard (n = 4) |
| Skating (n = 25) | | | | |

The search included only *review* and *research articles* written in the *English* language. Additionally, results from the following subject areas were excluded depending on the term that substituted †Sport†: Agriculture and Biological Sciences; Biochemistry, Genetics and Molecular Biology; Veterinary Science and Veterinary Medicine; Physics and Astronomy; Environmental Science; Earth and Planetary Science; Material Science; Chemistry; Chemical Engineering; Engineering; Energy; Nursing and Health Professions.

**Table A6.** Summary of proposed measures against national bias

| | Efficacy | Feasibility | Tradeoffs |
|---|---|---|---|
| **Structural modifications to judging panels** | | | |
| Prohibiting same nationality judging | Eliminates both patriotic and competitor forms of national bias but may inadvertently increase the likelihood of vote trading. | Requires amendments to scoring rules and regulations but is likely impractical in competitions with a large number of participating nations.<br>This approach has been formally implemented in several sports. | + Prevents direct scoring benefits from compatriot judges.<br>– May foster collusion among judges.<br>– May necessitate extensive training for officials from underrepresented countries. |
| Balancing judging panels | Exhibits limited effectiveness due to substantial variation in judges' individual proficiency in navigating scoring margins. | Requires amendments to scoring rules and regulations, with feasibility diminishing as the number of nations to be balanced increases. | + Enables reciprocal neutralization of national bias<br>– Effectiveness is contingent upon judges' ability to strategically exploit scoring margins.<br>– Risks amplifying distortions through indirect compensating or affirming biases.<br>– Increases operational costs.<br>– May necessitate extensive judge training. |
| Expanding judging panels | Reduces the influence of individual jurors on competition outcomes, with the dilution effect increasing proportionally to the number of judges. | Requires amendments to scoring rules and regulations. | + Attenuates the impact of bias on final scores.<br>+ May facilitate reciprocal bias compensation.<br>+ Potentially decreases the effectiveness of collusion among jurors, thereby reducing vote trading.<br>– Increases the probability of matched nationalities between judges and athletes.<br>– Risks amplifying distortions through indirect compensating or affirming biases.<br>– Increases operational costs. |
| **Exploiting judge demographics and psychological factors** | | | |
| Revisions in male juror selection, training, and evaluation processes | The effectiveness of this measure remains uncertain, as it is predicated on conclusions that are subject to debate. | Requires modifications to the selection, training, and evaluation procedures for male judges. | – Lacks robust empirical evidence supporting the existence of the underlying issue it aims to address. |
| Increased monitoring and sanctions | The effectiveness of this measure remains uncertain, given that systematic monitoring and sanctioning mechanisms already exist in numerous sports where national bias remains evident. | Requires modifications to the evaluation and sanctioning procedures for judges.<br>Most sports already employ judge evaluations and impose sanctions for substandard performance. | + May exert a deterrent effect on biased judging.<br>– Strict oversight may amplify affirmation bias. |
| No judge selection by national federations | Reduces the potential influence of national federations and their leadership over judges, thereby also diminishing the strategic value of national bias as a signaling mechanism for securing future assignments. | Requires amendments to procedural rules and regulations.<br>This approach has been formally implemented in several sports. | + Limits the influence of national federations and their leadership on judges.<br>+ Reduces the incentive to employ national bias as a signaling strategy. |
| **Transparency and anonymity** | | | |
| Anonymization of scores | Anonymizing judges' scores prevents the direct observation of national bias; however, empirical evidence indicates limited effectiveness in reducing the actual occurrence of such bias. | Requires policy changes.<br>This approach has been implemented in several sports. | + Conceals potential national bias from participants and stakeholders, as individual scores are not publicly disclosed.<br>– May encourage strategic or overt bias due to reduced public scrutiny.<br>– Restricts opportunities for external monitoring. |
| Increasing transparency of scoring systems | Some empirical evidence suggests that enhancing the transparency of scoring systems may effectively reduce the manifestation of national bias. | Requires policy changes.<br>This approach has been implemented in several sports. | + May mitigate both national and affirming biases.<br>– Enables external oversight and facilitates academic research.<br>– Potential national bias may be perceived by participants and stakeholders. |

(*Continued*)

**Table A6.** Continued.

| | Efficacy | Feasibility | Tradeoffs |
|---|---|---|---|
| **Enhancing scoring criteria and specialization** | | | |
| Increasing the level of detail of scoring criteria | The efficacy of this measure remains uncertain due to a lack of empirical evidence. | Requires amendments to scoring rules and regulations. | + May reduce subjectivity in scoring.<br>+ May diminish the influence of political pressure on judges.<br>– May increase procedural complexity and cognitive load on judges.<br>– May promote reliance on heuristic processing. |
| Task specialization | Research indicates that a reduced cognitive load decreases the likelihood of biases arising from heuristic processing, including those based on nationality. | Requires amendments to scoring rules and regulations. | + Reduces complexity and cognitive demands on judges.<br>– May mitigate reliance on heuristic decision-making.<br>– May increase operational costs. |
| Employing Reference Judges | The efficacy of employing reference judges remains contested, as evidence suggests they do not consistently exhibit greater accuracy than regular judges. | Requires amendments to scoring rules and regulations.<br>This approach has been implemented in some sports. | + Employing a small number of reference judges may inadvertently amplify existing biases.<br>– Increases operational costs.<br>– May necessitate extensive judge training. |
| Pre- and post-competition judge meetings | The efficacy of pre- and post-competition judge conferences remains unclear due to a lack of empirical evidence. | Requires policy changes.<br>This approach has been implemented in several sports. | + Fosters a shared understanding of scoring criteria.<br>+ Potentially reduces interpretative inconsistencies among judges. |
| Excluding judges from observing training sessions | The efficacy of excluding judges from training session observations remains uncertain due to insufficient empirical evidence. | Requires policy changes. | + May mitigate unintentional patriotic bias arising from mere exposure and availability heuristic mechanisms.<br>+ Incurs no additional costs.<br>– In certain sports or disciplines, observing training sessions is regarded as a beneficial and established practice. |
| Checklists of required elements | The efficacy of implementing checklists remains unclear due to a lack of empirical evidence. | Requires policy changes and amendments to scoring rules and regulations. | + Reduces cognitive demands on judges.<br>+ May mitigate reliance on heuristic processing.<br>– May introduce bias through expectation effects. |
| **Adjusting final score calculations** | | | |
| Score truncation | Mitigates the impact of outlier scores influenced by national bias on performance evaluations and competition outcomes. | Requires amendments to scoring rules and regulations.<br>Score truncation is widely practiced across most sports. | + Reduces the effect of biased outlier scores on final results.<br>+ Incurs no additional costs.<br>– Potentially results in the loss of valuable evaluative information. |
| Excluding compatriot judges' scores from final score calculations | This approach is prone to introducing substantial bias arising from variations in individual judges' leniency and is highly susceptible to manipulation. | Requires amendments to scoring rules and regulations. | – May substantially distort competition outcomes as a result of variability in individual judges' leniency.<br>– Vulnerable to strategic exploitation by judges. |
| Mathematical models adjusting for national bias estimates | Mathematical models adjusting for estimated national bias are likely compromised by variability in bias across individual scoring instances. | Requires amendments to scoring rules and regulations. | • Inaccuracies stemming from case-specific bias fluctuations.<br>– Vulnerable to strategic exploitation by judges. |
| **Technological support** | | | |
| Electronic scoring systems | The efficacy of electronic scoring systems remains uncertain due to a lack of empirical evidence. | Involves substantial financial investment and extended development periods, alongside necessary policy and regulatory modifications. | + May reduce perceptual and cognitive demands on judges.<br>+ May improve overall scoring accuracy and mitigate certain biases.<br>– Faces challenges in adequately capturing the aesthetic quality of performances. |
| Integrating the audience into the scoring process | The efficacy of integrating audience participation in the scoring process remains uncertain and is likely to introduce substantial bias. | Involves substantial financial investment and extended development periods, alongside necessary policy and regulatory modifications. | + May capture a broader and more holistic assessment of aesthetic performance quality beyond that of trained judges alone.<br>– May introduce biases related to the nationality composition of the audience.<br>– Untrained spectator judgments may adversely affect competition outcomes.<br>– May encourage exaggerated audience reactions aimed at strategically influencing results. |