Rage against the machine: investigating conspiracy theories about the video assistant referee on Twitter during the 2018 FIFA World Cup

Paul Bertin, Sylvain Delouvée, Kathleen McColl & Jan-Willem van Prooijen

To cite this article: Paul Bertin, Sylvain Delouvée, Kathleen McColl & Jan-Willem van Prooijen (2022): Rage against the machine: investigating conspiracy theories about the video assistant referee on Twitter during the 2018 FIFA World Cup, Sport Management Review, DOI: 10.1080/14413523.2022.2127179

To link to this article: https://doi.org/10.1080/14413523.2022.2127179

Published online: 28 Nov 2022.
Rage against the machine: investigating conspiracy theories about the video assistant referee on Twitter during the 2018 FIFA World Cup

Paul Bertin a,b, Sylvain Delouvée c, Kathleen McColl c and Jan-Willem van Prooijen d,e,f

aCenter for Social and Cultural Psychology (CeScuP), Université libre de Bruxelles, Brussels, Belgium; bUniversité Côte d’Azur, LAPCOS, France; cUniv Rennes, LP3C (Laboratoire de Psychologie: Cognition, Comportement, Communication), Rennes, France; dVU Amsterdam, Amsterdam, the Netherlands; eThe Netherlands Institute for the Study of Crime and Law Enforcement (NSCR), Amsterdam, The Netherlands; fMaastricht University, Maastricht, the Netherlands

ABSTRACT
Conspiracy theories arise during important societal and political events, with negative consequences. Yet, conspiracy theories remain to be investigated in the context of sporting tournaments, in spite of the importance of such events in contemporary societies. During the 2018 FIFA World Cup, conspiracy theories alleging that the newly introduced video Assistant Referee (VAR) was used with malevolent motives gained popularity online. In this paper, we used a Twitter content analysis to explore VAR conspiracy theories during the World Cup (N = 2,768 tweets). Conspiracy tweets peaked after eliminations of some teams and were strongly associated with labels referring to self-categorization at the group-level, supporting the notion that conspiracy beliefs are triggered by contexts threatening one’s social identity. Conspiracy tweets were also correlated with tweets expressing skepticism or defiance against the VAR, suggesting that conspiracy beliefs might be related to sport fans’ other identity management strategies. Finally, we drew an inter-group typology of VAR conspiracy beliefs’ recurrent figures, highlighting that higher ordered categorization helped conspiracy narratives and content to adapt throughout the tournament. We discuss the identity management strategy status of sport fans’ conspiracy beliefs.

“At the international level sport is frankly mimic warfare.” Orwell (1945, para. 4)

The Fédération Internationale de Football Association (FIFA) World Cup tournament gathers billions of viewers1 and could be described as a theater of war for sport fans, following George Orwell’s quote. Indeed, it is a stressful and emotional time for national teams’ supporters, who repeatedly face the threat of seeing their team eliminated from the tournament. Literature derived from the Social Identity Approach (Reicher et al., 2010) has documented various identity management strategies used by sports fans to cope

CONTACT Paul Bertin paul.bertin@ulb.be

1The 2018 edition has been watched by more than 3.5 billion people (FIFA, 2018a).

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2022 Sport Management Association of Australia and New Zealand
such with team-based threats (see, Wann & James, 2018). In this paper, we propose to add conspiracy beliefs to this set of coping strategies. Conspiracy beliefs are defined as accusations of powerful outgroups secretly plotting to achieve goals that will directly or indirectly harm one or several groups (Bertin et al., 2022a), and can lead to numerous detrimental consequences (see, Douglas et al., 2019). These beliefs frequently arise following societal crises (Van Prooijen & Douglas, 2017) and perceived threat (Bertin et al., 2022b). Yet, despite the emotional burden sports’ fans might encounter, to the best of our knowledge conspiracy beliefs have never been investigated in the sport context, as highlighted by Van Prooijen and Douglas (2018). In this paper, we used a Twitter content analysis to explore conspiracy theories about the newly introduced Video Assistant Referee (VAR) in the context of the 2018 FIFA World Cup.

**Team-based threats and identity management strategies**

Sport fans are frequently confronted with team-based threats to their identity (Delia, 2017, 2019; Foster & Hyatt, 2007; Mansfield et al., 2020; Wann & James, 2018; Wegner et al., 2020). The most common threat is to the value and competence of the supported team, following a defeat, elimination, or relegation (Branscombe et al., 1999; Doyle et al., 2017). More frequently encountered by high identifiers (Bernache-Assollant, 2010), such threatening events are experienced as personal defeats (Hirt et al., 1992) and can lead to serious physical and psychological damage. For example, after the elimination of the Nottingham Forest Football Club from the FA cup, Steels (1994) reported an increase of self-poisoning records at the local hospital. Similarly, teams’ relegation from the English Premier League can impact supporters in a way close to post-traumatic stress syndrome (Banyard & Shevlin, 2001). Finally, Kirkup and Merrick (2003) five-year longitudinal study in England highlighted a relationship between home defeat and death rate due to circulatory disease among men. These findings underscore how heavily invested sports fans can be in their team’s performance, making it plausible that they display an array of identity management strategies to cope with team-based threats.

In social psychology, identity management strategies have been mostly documented based on the Social Identity Approach (Reicher et al., 2010). Applied to the sport context, this theory predicts that supporters will strategically use a set of reactions and behaviors to promote or maintain a positive social identification of the supported team. Historically, one of the first contributions investigating fans’ reactions to game outcomes (although not based on the Social Identity Approach) has been about Basking in Reflected Glory (BIRGing), occurring after a victory when sports fans boost their social identity by expressing their belonging (Cialdini et al., 1976). In contrast, Cutting Off Reflected Failure (CORFing) takes the form of distancing one’s self from the team following a defeat (Wann & Branscombe, 1990). When mobility is not possible due to the centrality of one’s attachment to the team, blasting might occur by denigrating other teams and fans (Cialdini & Richardson, 1980; Wann, 1993). While a strong body of evidence documents these identity management strategies as well as many others, here we argue that sports fans additionally endorse conspiracy theories to regulate team-based threats to identity.
An intergroup approach to sports conspiracy beliefs

It has been proposed that conspiracy beliefs might be rooted at a group level in the need to maintain a positive social identity (Biddlestone et al., 2021; Douglas et al., 2017; Krekó, 2015; Sternisko et al., 2020; Van Prooijen & van Lange, 2014). Conspiracy beliefs can fulfill this function by attributing negative outcomes for one’s ingroup to the covert actions of an alleged malevolent outgroup (Abalakina-Paap et al., 1999). Early findings showed that in the U.S., the belief that the U.S. government was conspiring against African Americans was predicted by race (Crocker et al., 1999). The authors proposed that conspiracy beliefs are “a construal that reflects a particular view of the situation facing one’s racial group and perhaps an attempt to cope with the threat to the self that situation poses” (p. 949). Similarly, after a terrorist attack by Indonesian Muslim extremists in Indonesia, a conspiracy theory spread in the population claiming that the perpetrators were in fact Western governments, thus protecting a positive Indonesian identity by blaming an outgroup for these events (Mashuri & Zaduqisti, 2015). Hence, perceiving a threat for one’s ingroup identity is a core feature of conspiracy beliefs at the group level.

Congruently, the existential threat model of conspiracy belief (Van Prooijen, 2020) predicts that endorsement of a conspiracy theory is likely to occur if people experience feelings of an existential threat, defined as the anxiety or insecurity that people feel when they, or the people around them, are harmed or expect to suffer losses. For example, Uscinski and Parent (2014) showed that conspiracy theories were more popular among partisans supporting an electoral loser than among partisans supporting an electoral winner in the United States. This asymmetry has been explained by partisans’ compensating for the threat of being out-of-power (the “conspiracy theories are for losers” argument, Smallpage et al., 2017). Similarly, the elimination of the team one supports from an international tournament might be an especially prominent existential threat. Another core feature to predict conspiracy beliefs according to this model is the saliency of an antagonistic outgroup. That is, trying to make sense of an existential threat by attributing the wrongdoing to a contextually or historically competing outgroup, such as (in the context of a sporting event) other national teams or officials. For instance, Finn (1991) described how Catholic Irish-Scottish players were accused of conspiring against the Scottish majority in the Scottish football championship, extending the intergroup conflict between the Scots and the Catholic Irish Scots.

Hence, investigating conspiracy theories in the sport context is of importance, as these reactions indicate distrust toward groups but also toward institutions (Van Prooijen et al., 2022a). Indeed, conspiracy beliefs have been shown to increase prejudice against groups (Jolley et al., 2020), in line with their description as an “outlet for hostility” (Abalakina-Paap et al., 1999). Applied to the sport context, these beliefs might lead to dysfunctional behaviors toward other groups of fans (Larkin & Fink, 2019) and officials. We might speculate that these beliefs also lead to collective expressions of anger and resentment harming brand images, and impeding attendance and engagement in fandom activities. As such, conspiracy theories that arose regarding the VAR during the 2018 FIFA World Cup are indicative of the harm these reactions might cause to institutions.
The case of VAR conspiracy theories during the 2018 FIFA World Cup

In this research, we focused on conspiracy theories surrounding the VAR during the 2018 FIFA World Cup. The VAR was used for the first time in a World Cup during this tournament (BBC, 2018). This system is aimed at improving referees’ decisions by allowing them to review litigious actions on a screen (FIFA, 2018b). However, from the supporters’ perspective, the VAR might provide fertile ground for mistrust and anger (Hamsund & Scelles, 2021; Kiuchi, 2022; Kolbinger & Knopp, 2020). Indeed, given the reviewed in-game situations’ being ambiguous, supporters might be quick to fall into engaging in ingroup bias (Wann & Grieve, 2005) and to perceive the right decision to be the one in favor of their team. Furthermore, supporters might be both frustrated by the VAR’s call not being in favor of their team, but also by an absence of a VAR call, when they believe it should have been necessary. Hence, these team-centered perceptions of the VAR, added to the emotional load of the event, created a fertile ground for conspiracy theories that we labeled VAR conspiracy beliefs.

VAR conspiracy theories claimed that VAR was introduced in the tournament with the malevolent goal of discarding non-European or “minor” teams (i.e., teams representing African, Arab, and “small” countries like Sweden or Serbia). One of the most prominent versions of the VAR conspiracy theory claimed that racist motives were behind the introduction of the VAR (i.e., to prematurely eliminate African and Arab teams). This latter motive was fueled by the fact that for the first time since 1982, no African or Arab teams out of the six initially qualified (Nigeria, Egypt, Senegal, Morocco, Tunisia, Iran) reached the knockout stage (Africanews, 2018). This VAR conspiracy theory led to massive online protests, in the form of online petitions, one of which, that called to reschedule the Morocco – Spain match, gathering more than 60,000 signatures. Lastly, players also played a role in the rise of this mistrust, as illustrated by the Moroccan player, Nordin Amrabat, who claimed in front of the cameras that “VAR is bullshit!”, or Younès Belhanda, who argued that “VAR just helps big teams”. In the following section, we investigate VAR conspiracy theories through the lens of a Twitter content analysis.

Overview of the present study

We aimed to examine conspiracy theories in the sport context through a case study of VAR conspiracy theories during the 2018 World Cup. Using a Twitter content analysis, our goals were to explore: (a) the temporal distribution of VAR conspiracy theories, (b) the match distribution of VAR conspiracy theories, (c) the relationship between conspiracy theories and tweets expressing skepticism or defiance against the VAR, (d) the potential differential popularity between VAR conspiracy theories and tweets expressing skepticism or defiance against the VAR, (e) whether fans expressed conspiracy tweets taking a group vs. individual stance, and to (f) establish an intergroup typology of VAR conspiracy theories based on the tweet’s content. Through these objectives, we sought to increase knowledge of conspiracy theories in the sport context expressed as reactions to identity-threats, by providing an initial insight into their nature, content, and context.
Method

The microblogging platform Twitter has been increasingly used for academic research purposes to study a wide range of topics (Williams et al., 2013). Some authors even argue that the archival research Twitter facilitates might improve social psychology, among other domains, by allowing large-scale investigations of socially sensitive phenomena (Heng et al., 2018). Twitter is very useful to study “breaking news” (i.e., highly mediatized events like natural disasters), due to the reactivity this kind of news elicits. The hashtag system interferes with traditional media coverage, allowing a comprehensive account of the diversity of reactions to an event (Bruns & Moe, 2014). In the present research, World Cup matches were the “breaking news”, which we analyzed for VAR conspiracy theories.

Twitter seems especially relevant to explore sports fans’ identity management strategies. Indeed, previous research highlighted the importance the internet plays in the development of sports fans’ identity, allowing a digital transposition of identity management strategies (for a review, see, Pegoraro, 2013). Notably, Berg and Harthcock (2008) observed similarities between online and offline sports fans’ identity management strategies. This might be due to sports fans’ identification and self-esteem being highly sensitive to mass online media coverage (Phua, 2010). Finally, online gatherings of sports fans have been reported to be larger and more heterogeneous than offline ones (Hutchins & Rowe, 2012). That is, research based on online gatherings allows us to investigate a bigger and more diverse audience of sport fans, which provides unique research opportunities.

Data collection

We collected tweets published in English during the 2018 FIFA World Cup (14 June to 15 July 2018), using the independent and open-source software, Orgneat (Orgneat, 2020). This software allows a collection of “historical” tweets based on keywords. To obtain tweets broadly referring to VAR conspiracy theories, we chose four different keyword associations. First, we targeted VAR conspiracy theories claiming that the FIFA used the VAR with racist motives using the keywords “var” and “racism”. This first research allowed us to identify 740 tweets. Second, we aimed to tap into conspiracy theories surrounding the anger movement launched by the Moroccan player, Nordin Amrabat, using the keywords “var” and “bullshit” (N = 9,607). Finally, to diversify our sample with VAR conspiracy theory tweets unrelated to the previously mentioned keywords, we ran two broader collections using the keywords “var” and “FIFA” (N = 10,849), and “var” and “world cup” (N = 26,387).

Then, we randomly sampled 740 tweets from each data set. We arbitrarily chose this subsample size to match the size of our smallest data set. Each of these subsamples was combined into a new data set (N = 2,960), from which redundant tweets were deleted (n = 192). The final data set was composed of 2,768 tweets.

Data and Supplementary Materials are openly available on the Open Science Framework repository for this project: https://osf.io/xeavt/?view_only=00f9f4a315a5433eba727c04736b1306
**Data analysis**

We drew from Bordia and Difonzo’s (2004) Rumor Interaction Analysis System (RIAS) to develop a matrix, aiming to categorize the sampled tweets. The RIAS was first designed to categorize and analyze the content of rumors on internet discussion groups (Bordia & Difonzo, 2004). More recently, the RIAS has been adapted to study conspiracy beliefs using content analysis of Twitter statements (Wood, 2018). By categorizing a statement’s content based on a set of established criteria, the RIAS is well suited to classify a small number of characters such as those involved in tweets. As did Wood (2018), we did not use RIAS in a strict way, but simplified it and adapted it for the purpose of the present research.

Specifically, based on our goals and on preliminary observations of the tweets’ content, we developed four categories to study conspiracy tweets about the VAR, as well as other forms of distrust toward the VAR. The first category was labelled “conspiracy tweets”. It referred to tweet content that linked the use of the VAR to hidden motives of a powerful outgroup harming a victimized ingroup (e.g., “Football as we know it managed for 150 + years without the gimmick of #VAR and we see now how it is used to ensure FIFA’s favorites (those who pay the biggest bribes) go through. God curse those who are conspiring to rob #IRN of their rightful victory”). Second, skeptical tweets referred to tweets discussing or criticizing the use of the VAR in an argumentative way, without referring to the conspiracist features of the first category (e.g., “If the 7 VAR refs are all FIFA accredited and make a choice at all angles why does another ref have to make the final decision #greyareaa”). Third, defiant tweets referred to tweets criticizing the use of the VAR in a non-argumentative way, hence without falling into the conspiracy or skeptic features of categories 1 and 2 (e.g., “VAR is bullshit”). This latter category might be compared to blasting, as previously described in the literature (Cialdini & Richardson, 1980; Wann, 1993). The last category comprised unclassified tweets (i.e., tweets not matching criteria of either of the above-mentioned categories, e.g., “well the same Japan team got eliminated last night and since then i haven’t seen anyone on social media complaining about racism against Asians. That was a pure game of good football; VAR only called in once i think. Both teams going for the win. Anybody got anything to say???”)

Before coding the whole sample of tweets, two independent judges (i.e., who worked separately and were blind to the other judge’s coding) coded 400 tweets (i.e., 100 tweets sampled from each of the four databases), to estimate the adequacy of the categorization matrix. We calculated Cohen’s Kappa to test the interrater reliability of this first set of categorizations. According to Landis and Koch’s (1977) benchmark, the level of agreement was substantial, κ = .67, 95% CI [.61, .72], p < .001. Then, the entire dataset was coded. The level of agreement increased, compared with the first wave of coding, and the overall interrater reliability remained substantial, κ = .70, 95% CI [.68, .72], p < .001. A third judge then categorized the tweets for which a disagreement remained (n = 618).

Summaries of the categorization depending on days of the tournament can be found in Table S1 (Supplementary Materials). We also summarized the categorization depending on specific matches of the tournament (Table S2). Note that we duplicated in the dataset tweets referring to more than one match (n = 65). For tweets not explicitly stating to

---

2Note that the following reproduced tweets were not altered in spelling nor in capitalization, as advised by the APA guidelines (American Psychological Association, 2020).
which matches they referred to (e.g., by not using an official FIFA World Cup hashtag), we attempted to retrieve this information based on the content and date of the tweets. In total, 1,218 tweets referred to at least one match, with 51 out of 64 matches played during the tournament.

**Results**

**Data visualization**

**Distribution by days of the tournament**

First, we used data visualization to observe the temporal distribution of the coded tweets for each day of the World Cup. As can be seen in Figure 1, conspiracy tweets slowly increased from June 19 (when the first team was eliminated) to June 24, before peaking between June 25 and June 28 (end of the group stages when half of the teams did not qualify for the knockout stage). Interestingly, defiant tweets seemed to increase more than conspiracy tweets, before this pattern changed on June 26 and after. The distribution then sharply decreased from July 29 to July 15 (knockout stage), before slightly increasing on July 15 (Final match). Finally, whereas skeptical tweets were quite equally distributed with other categories of tweets, they did not peak as much as conspiracy and defiant tweets, which might indicate that this category is unrelated to threatening contexts.

To further explore the categorization of tweets, we next moved from a day to a match distribution.

**Distribution by retrieved matches**

For each match, categories were positively correlated, such that the more conspiracy tweets were expressed, the more they were also skeptical, \( r = .57, p < .001 \), and defiant tweets, \( r = .53, p < .001 \). Skeptical and defiant tweets were also correlated, \( r = .52, p < .001 \).

![Figure 1](image-url)  
**Figure 1.** Distribution of the number of tweets per category for each World Cup days.  
Note. \( N = 1,974 \) tweets. Uncategorized tweets were excluded to improve clarity. * = crucial referee decision based on a VAR review. † = team eliminated from the tournament. Note that we did not consider the defeat during the third-place match between Belgium and England as an elimination, as these two teams were eliminated during semi-final matches. See Figure S1 in the Supplementary Materials for the figure including the uncategorized tweets category.
As can be seen in Figure 2, the classification of conspiracy tweets was highly asymmetrical between matches. Many matches had a small number of tweets, independent of the categories. About seven matches had a substantial number of tweets, but nothing compared to the three leading matches. Indeed, Spain against Morocco (at the end of which Nordin Amrabat said that “VAR is bullshit”), Nigeria against Argentina (elimination of Nigeria), and Senegal against Poland (elimination of Senegal), generated the greatest number of tweets. Interestingly, a consequence of Nordin Amrabat’s words appeared to be that conspiracy and defiant tweets were rather equally distributed for the Spain – Morocco match, whereas conspiracy tweets were much more prevalent for the two other leading matches. As for the daily distribution, skeptical tweets were overall less represented than the two other categories.

Hence, conspiracy tweets seemed to rise at the end of the group stage, presumably as a consequence of increased team eliminations. Conspiracy tweets might also have been amplified by the “VAR is bullshit” movement initiated by Nordin Amrabat’s words, which at first caused a wave of defiant tweets. Interestingly, the increase of tweets during the World Cup Final had a unique distribution, with only a minority of conspiracy tweets compared to skeptic and defiant tweets. Next, we explored whether VAR conspiracy tweets were more popular than other categories of tweets expressing skepticism or defiance against the VAR.

**Exploration of popularity indexes**

We aimed to investigate whether conspiracy tweets varied in popularity, compared with other categories of tweets. To do so, we relied on popularity indexes, that is replies, retweets, and likes given to each category of tweets. However, large differences of variance within and between popularity indexes did not allow us to run inferential
statistics to test whether categories of tweets differed in popularity. See Table S4 for descriptive statistics of popularity indexes. We next present advanced coding of our database designed to investigate further the identity status of conspiracy tweets.

**Testing the identity status**

Our claim is that conspiracy tweets are expressed as identity management strategies by fans facing an identity threat. However, data visualization does not allow for a comprehensive testing of this hypothesis. Hence, we hand-coded the sample of 2,778 tweets depending on whether the author of the tweet referred to himself as an individual (e.g., using labels such as “I” or “my”) or as a group member (e.g., using labels such as “we” or “our”). In total, 241 tweets were categorized as expressed at the individual-level and 192 at the group-level.

In absolute terms, both individual and group-level tweets were more numerous among conspiracy than among skeptical and defiant tweets (see, Table 1 for details). A chi-squared test reported that differences in counts between categories were significant, \( \chi^2 (2, N = 433) = 24.6, p < .001 \).

Of importance, this over-representation among conspiracy tweets was more pronounced for group-level tweets (see, Figure 3). That is, whereas individual tweets were almost similarly numerous among conspiracy and skeptical tweets, group-level tweets appeared as rather specific to conspiracy tweets.

In the following, we sought to analyze and visualize individual and group-level tweets relative to the entire sample of categorized tweets (i.e., conspiracy, skeptic, or defiant), rather than in absolute count. To observe these proportions, we had to account for categorized tweets for which identity status could not be established \((N = 1541\) tweets).\(^3\) The proportions of individual and group-level tweets are displayed in Figure 4.

<table>
<thead>
<tr>
<th>Table 1. Contingency table of the number of individual and group-level tweets for each category.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category of Tweets</strong></td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td><strong>Conspiracy</strong></td>
</tr>
<tr>
<td>Observed N (expected N)</td>
</tr>
<tr>
<td>% Within row</td>
</tr>
<tr>
<td>% Within column</td>
</tr>
<tr>
<td><strong>Skeptic</strong></td>
</tr>
<tr>
<td>Observed N (expected N)</td>
</tr>
<tr>
<td>% Within row</td>
</tr>
<tr>
<td>% Within column</td>
</tr>
<tr>
<td><strong>Defiant</strong></td>
</tr>
<tr>
<td>Observed N (expected N)</td>
</tr>
<tr>
<td>% Within row</td>
</tr>
<tr>
<td>% Within column</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
<tr>
<td>% Within row</td>
</tr>
<tr>
<td>% Within column</td>
</tr>
</tbody>
</table>

\(^3\)The Chi-squared test remained significant when accounting for uncategorized tweets, \( \chi^2 (2, N= 1974) = 122, p < .001.\)
As for counts, the proportion of defiant tweets marked at the identity level was small, be it at the individual (4.7%), or at the group-level (4.9%). By contrast, the proportion of skeptical tweets expressed at the individual-level (23.4%) was much higher than that expressed at the group-level (8.2%). This balance was more nuanced regarding conspiracy tweets: the proportion of tweets expressed at the group-level (13.7%) was higher than in
other categories and represented a bigger part than individual-level conspiracy tweets (12.5%). These observations suggest that fans expressing conspiracy tweets have greater tendencies to self-categorize at the group-level, which is congruent with our conceptualization of sport fans’ conspiracy beliefs as identity management strategies expressed in the face of team-based threats.

Whereas the previous analyses allow us to better understand the conditions under which conspiracy tweets are expressed, it would be erroneous to draw VAR conspiracy tweets as a uniform phenomenon. Indeed, conspiracy theories are too often described as abstractions, without accounting for contexts and contingencies. Conspiracy theories are not self-contained narratives but are deeply embodied in social dynamics, which ultimately impact their shape. Hence, in the following, we take a qualitative stance – i.e., analyzing in-depth the content of tweets – and draw a typology of VAR conspiracy theories, based on our sampled tweets.

**Content of conspiracy tweets**

We identified a typology of VAR conspiracy theories based on the intergroup approach to conspiracy beliefs (Nera et al., 2021; Sternisko et al., 2020; Van Prooijen & van Lange, 2014). That is, we differentiated VAR conspiracy theories depending on the categorization made of portrayed ingroups and outgroups. Below is the proposed typology, with examples for each kind of VAR conspiracy theories (Table 2).

This typology highlights that the content of conspiracy tweets has been changing through the tournament. Below we test whether this content might have been evolving following a higher ordered recategorization process.

**Data visualization of African teams’ recategorization at the continental level**

As previously described, the continental level “Africa” as an ingroup in tweets seemed to be especially prevalent among supporters of African teams. To test the hypothesis that this higher ordered recategorization occurred following successive eliminations, we calculated the sum of the number of occurrences of the words “Africa” (and “African”), and the African team’s names in our sampled tweets for each day of the tournament (Figure 5). We observed that the first peak of occurrences was on June 25 for “Morocco”, while the “African” occurrences were much smaller. On June 26, the peak was similar in intensity for “Nigeria” and “Africa”. Then, from June 27 to July 1, “Africa” remained more prevalent than names of individual African country teams, which we interpreted as the recategorization of eliminated teams on a higher ordered level.

The pattern of the distribution was similar overall when only conspiracy tweets were considered (n = 935; Figure 6). Last, occurrences of “Africa” were much more prevalent than the “Arabian” (and “Arab”), which could have been an alternative super-ordinate categorization (see Figures S3 and S4 in the Supplementary Materials).

**Discussion**

The present work aimed to investigate the nature, content, and context of conspiracy theories about the VAR that spread during the 2018 World Cup on Twitter. Our findings suggest that conspiracy tweets peaked at the end of the group stages, and then abruptly
Table 2. Typology of the content of conspiracy tweets.

<table>
<thead>
<tr>
<th>Portrayed Ingroup</th>
<th>Description</th>
<th>Sample tweet</th>
</tr>
</thead>
<tbody>
<tr>
<td>National</td>
<td>Most of observed conspiracy theories focused on one’s national team interest and victimhood. As such, the national ingroup level was the most basic and widespread ingroup categorization.</td>
<td>“I know one thing; in every World Cup I have seen I have come to the conclusion that the Refs are paid to ensure we lose. They have probably hacked the VAR as well under the orders of England hater; Sepp Blatter who remains plotting in the shadows!”</td>
</tr>
<tr>
<td>Continental</td>
<td>Interestingly, and congruently with the Self-Categorization Theory (Turner et al., 1987), supporters switched to higher ordered self-categorization depending on the context (i.e., match played), and VAR conspiracy theories were adapted consequently. The most observed recategorization occurred at the African level at the end of the group stages, when many African teams were successively eliminated (see, Figure 5).</td>
<td>“This var was actually put in place for the African teams due to racism. African teams have been getting robbed at this major event majority of the time due to bad call against them”</td>
</tr>
<tr>
<td>Racial</td>
<td>Another higher ordered self-categorization encountered was observed at the racial group level. Related tweets were similar to the continental category and arose following the African teams’ elimination. Congruently, the main alleged conspirators’ motive was racism.</td>
<td>“This is nonsense! Height of racism was shown today by FIFA and the ref! Keep Violating African Right with the nonsense VAR! I don’t know when all these hate against the black will stop! Total rubbish”.</td>
</tr>
<tr>
<td>Minor Teams</td>
<td>The nature of ingroups in these VAR conspiracy theories was similar to that regarding national groups previously described but differed in that supporters categorized their national team as being a “minor” team, less lucrative than “major” teams. Hence, alleged conspirators’ motive was mainly pecuniary. Note: higher ordered self-categorization also seemed to happen concerning minor national teams, with victimhood shared by several minor teams together (i.e., inclusive victimhood, see, Szabó, 2020).</td>
<td>“@FIFAcom VAR showing they choose who they want to benefit. You change de president but don’t change your old and disgusting mind. Sweden had to win but FIFA didn’t want. Money before soccer”.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Portrayed Outgroup</th>
<th>Description</th>
<th>Sample tweet</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIFA</td>
<td>The alleged conspirator that was the most common in VAR conspiracy theories we observed was FIFA, who organized the tournament and implemented the VAR system. Underlying harmful motives varied, from pecuniary to racism, as previously mentioned.</td>
<td>“Pure act of conspiracy by @FIFAcom To ensure that Argentina qualify. Suddenly the VAR is not effective to see the foul on @67Kelechi in the box. I can see Racism playing itself here”.</td>
</tr>
<tr>
<td>Switzerland</td>
<td>Switzerland appeared to be the national embodiment of the animosity against FIFA, as the FIFA headquarters is based in Zurich. Consequently, this outgroup appeared in conspiracy theories about matches involving the Swiss national team.</td>
<td>“@FIFAcom STILL corrupt i see. Serbia don’t get a penalty review from #VAR (which WAS a penalty) against SWITZERLAND. It just so happens the FIFA President (Gianni Infantino) is SWISS! Who’d have thought. Shocking. #FifaWorldCup2018 #FIFAWC2018 #SERSUI”</td>
</tr>
<tr>
<td>Russia</td>
<td>The host country of the tournament also gathered mistrust and was accused of conspiring. As for Switzerland, this outgroup mostly appeared in VAR conspiracy theories when the Russian national team was playing.</td>
<td>“The VAR all awarding Russia the calls; the same reason FIFA awarded Putin the World Cup?”</td>
</tr>
<tr>
<td>Western Countries</td>
<td>Higher ordered categorization also occurred about alleged conspiring outgroups, the most common being about western countries or “rich” countries.</td>
<td>“I will never forget that one day in #Russia2018 #FIFA created the #VAR to help the rich countries against Africa Arab and Muslim countries …”</td>
</tr>
</tbody>
</table>

(Continued)
decreased. A match-by-match distribution further indicated that only a few matches generated most of the conspiracy tweets from our sample. The occurrence of conspiracy tweets was positively correlated with tweets expressing skepticism and defiance towards the VAR. Importantly, conspiracy tweets were more frequently associated with labels of group-level self-categorization of fans than other categories of tweets. Finally, we proposed a typology of the different actors observed in the different versions of the VAR conspiracy theory from an intergroup approach. The main finding here was that higher ordered categorization was associated with the content of the conspiracy theories, which was confirmed by data visualization.

**Theoretical implications**

The rise of the conspiracy tweets echoed an increase in team eliminations and use of the VAR. Following Van Prooijen (2020), we proposed that people experienced the elimination of their team from the tournament as an existential threat and blamed this on the use
of the VAR by FIFA, leading to the observed expression of conspiracy beliefs. This expression was marked by an abrupt decline in the number of conspiracy tweets, which was rather intriguing. Here, we might speculate that conspiracy beliefs may serve as short-term identity management strategies in the sporting context. Once the instant negative emotions associated with the elimination of their team from the tournament dissipated, supporters no longer felt the need to tweet conspiracy theories about VAR. This transitory aspect leaves an important unanswered question: what remains of belief after the identity threat ends? Furthermore, it might also be worthwhile to investigate this speculated short-lived feature to further and better understand why some conspiracy beliefs last for decades (e.g., the JFK assassination), while others seem to last only for a few days.

An alternative interpretation could be that believing in conspiracy theories in the context of sport events is not an identity management strategy in itself but rather the expression of another documented identity management strategy, ingroup bias. Previous research has highlighted that ingroup bias uncovers very different perceptions of the same game, depending on whether one self-categorizes at the ingroup or outgroup-level (Hastorf & Cantril, 1954; Molenberghs & Louis, 2018). Furthermore, as for conspiracy beliefs, ingroup bias is contingent upon perceptions of team-based threats and prevalent among higher identifiers (Wann & Grieve, 2005). Hence, future research might want to disentangle conspiracy beliefs from ingroup bias, or alternatively specify under which circumstances ingroup bias might take the form of conspiracy beliefs.

A key finding is the over-representation of group-level tweets among conspiracy tweets, which is in line with our conceptualization of conspiracy beliefs as identity management strategies. However, it is also important to note that individual-level tweets were also numerous among conspiracy tweets. These conspiracy tweets expressed at an individual level might be an expression of the conspiracy mentality of sport fans, which refers to one’s stable propensity to endorse conspiracy theories (Bruder et al., 2013; Enders et al., 2021; Imhoff et al., 2022). As individuals with a high level of conspiracy
mentality are prone to perceive plots in general, it is not surprising that they endorsed conspiracy theories about the VAR.

This interpretation of our results is congruent with Sternisko et al.’s (2020) theoretical framework which differentiates between individual and group-based motives behind conspiracy beliefs. This framework posits that an individual’s motives are primarily driven by conspiracy theory *qualities* (e.g., feeling unique or entertained; see, Imhoff & Lamberty, 2017; Lantian et al., 2017; Van Prooijen et al., 2022b). By contrast, group-level motives aim at protecting one’s social identity, congruent with the Social Identity Approach (Reicher et al., 2010). Hence, one may argue that the high levels of individual *and* group-level markers among conspiracy tweets indicated that these two components – content and qualities – were at stake in making these narratives appealing.

Sports fans’ conspiracy beliefs might also be the manifestation of a collective conspiracy mentality, that is a collective tendency to perceive conspiracies due to past suffering (Soral et al., 2018). Indeed, the countries from which supporters expressed the most conspiracy tweets were African and Arabian countries, for which neo-colonialism is still a vivid memory. Congruent with our findings, a study based on YouTube comments on the FIFA official channel found that viewing the VAR as a neo-colonialist biased tool was widespread (Petersen-Wagner & Ludvigsen, 2019). Another explanation could be that status matters, and that only fans perceiving their team as a *low*-status team suffering from power asymmetries (see, Nera et al., 2022a) endorse conspiracy theories as an explanation of failure. This would be congruent with research showing that members of disfavored groups are more inclined to explain the situation of their ingroup using conspiracy theories (Nera et al., 2022b). Hence, whereas African teams frequently underperformed during the FIFA World Cup compared with European teams, according to performance indicators (Kubayi & Toriola, 2020), conspiracy narratives might help to uphold a favorable perception of their teams.

The typology of VAR conspiracy theories we propose attempts to emphasize the core intergroup feature of such beliefs. Both ingroup and outgroups portrayed were context-dependent and evolved following advances in the tournament and team eliminations. Importantly, we observed that the level of categorization of the victimized ingroup similarly moved from a national to a continental level. This finding is especially interesting to analyze in light of recent research on comparative victimhood (Vollhardt et al., 2021). Whereas congruently with ingroup bias, we might at first speculate that competitive victimhood would be at play after one’s team’s elimination (i.e., claiming that one’s team’s elimination is more undeserved than elimination of other teams), we might have instead observed inclusive victimhood (i.e., claiming that the ingroup suffered in a way similar to what other groups experienced). As such, the conspiracy theories that we observed might reflect a narrative bonding of supporters of eliminated teams at a higher level of categorization. We nonetheless probably observed what might also be described as competitive victimhood, especially by English supporters (e.g., “How the F does VAR get used for everything in every other world cup match apart from in Englands game? Surely there is something dodgy there”). Here, we might hypothesize that “major” teams’ supporters use competitive victimhood to claim their superior status, whereas “minor” teams’ supporters use inclusive victimhood to gather with other “underdogs”.

Collective narcissism of sport fans could also play a role in explaining conspiracy beliefs in the sport context. Collective narcissism (i.e., a defensive identification rooted in
a perceived lack of recognition; Golec de Zavala et al., 2009; Larkin et al., 2021) is a robust predictor of conspiracy beliefs according to empirical (Bertin et al., 2022a, 2022b; Cichocka et al., 2016; Marchlewksa et al., 2019; Sternisko et al., 2021), and meta-analytic evidence (Biddlestone et al., 2022). The relationship between collective narcissism and conspiracy beliefs blaming external factors for a situation threatening the ingroup image has been proposed to be mediated by the endorsement of a victim status (Bertin & Delouvée, 2021). This phenomenon might similarly apply to the sport context: team narcissists feeling threatened by their team’s elimination might blame conspiracies, endorsing a victim status to legitimize these accusations. Hence, it would be worthwhile to test empirically the relationships between team narcissism, victimized perceptions of one’s team failure, and conspiracy beliefs about this failure, to further gain insights into sport fans’ reactions to threat.

Finally, the defiant tweets we observed might be viewed as an expression of blasting (Bernache-Assollant et al., 2007). This online manifestation of blasting would be in line with research showing that another identity management strategy, BIRGing, could be observed on Twitter during the 2018 World Cup (Fan et al., 2020). Future research might seek to account more comprehensively for multiple identity management strategies when conducting online content analysis.

**Practical implications**

Conspiracy theories are consequential, which has been documented in a variety of contexts (Jolley et al., 2022; Van Prooijen & Douglas, 2018). In the case of VAR conspiracy theories, the associated consequences we observed were online petitions, protests, and distrust toward various groups. More dramatic consequences might occur in the future due to VAR conspiracy theories, especially between supporters of opposing teams since conspiracy beliefs have been related to violent behavioral intentions (Jolley & Paterson, 2020).

It is undeniable that the VAR had a concrete impact on 2018 FIFA World Cup games (e.g., by increasing the number of penalties; Kubayi et al., 2021), and this change in the tournament’s rules should have been closely monitored and explained. Instead, its introduction may have suffered from a lack of communication and transparency by FIFA, especially regarding VAR-based decisions, which probably provided a fertile setting for breeding VAR conspiracy theories. Hence, because the spread of this form of mistrust is not without risk, it would be worthwhile investigating ways to diminish VAR conspiracy beliefs, for example, by increasing public transparency about communication between the VAR team and the main referee.

**Limitations**

Our research has some limitations that must be acknowledged. First, we collected tweets using keywords in English. This linguistic constraint led to inequalities between represented teams depending on the prevalence of English-speaking communities in these countries. For example, the biggest peak in conspiracy tweets we observed was related to the Nigerian team, which has a broad English-speaking community. By contrast, we are likely to have missed conspiracy tweets posted by Asian, Arab, or European teams’
supporters who did not tweet in English. Future research might attempt to look at tweets posted in more diverse languages when investigating supporters’ reactions to threat and conspiracy beliefs. Relatedly, we determined the identity status of the tweets based on pronouns, while we could have extended this analysis to more complex identity markers (see, Wegner et al., 2020). However, it is worth noting that the constraint of character limit might restrain the potential to find a great variety of identity markers when analyzing tweets.

Second, our results focused on VAR conspiracy theories mainly regarding alleged racist motives, which was by far the more prevalent theme in our sample. Future research may further investigate VAR conspiracy theories about monetary issues or conflict of interest with FIFA, as highlighted in our typology. Conspiracy theories also exist in the football and sport contexts outside the realm of the VAR case (e.g., Finn, 1991), but more research is needed here as current knowledge is scarce. For example, it would be of interest to explore conspiracy theories about sport events in a broader way, especially regarding allegations of match-fixing, which we did not focus on outside narratives about the VAR.

Third, we inferred those who expressed conspiracy tweets were fans identifying with teams, but this point remained speculative in the present research. Ultimately, the assumption of conspiracy theories as sport fans’ identity management strategies would benefit from empirical research. Specifically, the use of longitudinal designs would allow to test hypotheses regarding conspiracy beliefs as reactions to threats, as well as their potential short-term nature.

Conclusion

As far as we know, the present contribution provides the first attempt to investigate empirically sport conspiracy theories. Through a Twitter content analysis, we explored the context and content of the occurrence of VAR conspiracy beliefs during the 2018 FIFA World Cup. Our findings suggest that conspiracy tweets arose at the end of the group stages, when an increase in team eliminations occurred, together with several crucial VAR-based decisions. Notably, conspiracy tweets were expressed in reaction to a small fraction of matches. In line with the claim that sport fans’ conspiracy beliefs are identity management strategies, labels referring to group-level self-categorization were over-represented among conspiracy tweets. Finally, the content of conspiracy beliefs used as identity-management strategies might be contingent upon perceived stakeholders of the match played by the supported team. Congruently, groups portrayed in conspiracy tweets’ narratives evolve during the tournament, through recategorization, as highlighted by our typology and graphs. Improving transparency of VAR calls could help tackle this phenomenon, which denotes distrust toward institutions and a potentially deteriorated brand image.

Acknowledgments

We are thankful to Romain Raymondie for suggesting the pre-title and to Claire Boisserie for her help in coding the data.
Author contribution

P.B. conceptualized the project. P.B., K.M., and S.D. created the materials. P.B. conducted the data collections. P.B. conducted the analysis. P.B., K.M., and J.V. wrote and revised the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by a grant (PDR 0253.19) from the Belgian National Fund for Scientific Research (FRS-FNRS; Fonds De La Recherche Scientifique - FNRS

ORCID

Paul Bertin http://orcid.org/0000-0002-6700-460X
Sylvain Delouvée http://orcid.org/0000-0002-4029-597X
Kathleen McColl http://orcid.org/0000-0001-9920-3112
Jan-Willem van Prooijen http://orcid.org/0000-0001-6236-0819

Ethical statement

The authors declare that there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

BBC. (2018). VAR: Video assistant referees will be used at 2018 World Cup in Russia. https://www.bbc.com/sport/football/43438344


