ORIGINAL RESEARCH

Journal of Sport & Exercise Psychology, 2010, 32, 3-22
© 2010 Human Kinetics, Inc.

How Embodied Cognitions Affect Judgments: Height-Related Attribution Bias in Football Foul Calls

Niels van Quaquebeke and Steffen R. Giessner
Erasmus University Rotterdam

Many fouls committed in football (called *soccer* in some countries) are ambiguous, and there is no objective way of determining who is the "true" perpetrator or the "true" victim. Consequently, fans as well as referees often rely on a variety of decision cues when judging such foul situations. Based on embodiment research, which links perceptions of height to concepts of strength, power, and aggression, we argue that height is going to be one of the decision cues used. As a result, people are more likely to attribute a foul in an ambiguous tackle situation to the taller of two players. We find consistent support for our hypothesis, not only in field data spanning the last seven UEFA Champions League and German Bundesliga seasons, as well as the last three FIFA World Cups, but also in two experimental studies. The resulting dilemma for refereeing in practice is discussed.

Keywords: dominance, power, refereeing, decision making, decision cue, information processing

Whereas some popular team sports, such as rugby or American football, use video replays in slow motion and from different angles to assist referees in making judgments in ambiguous situations, the Fédération Internationale de Football Association (FIFA)—as the overarching ruling body for football (also called *soccer*, a colloquial abbreviation of the term *association football*)—still takes a conservative stand on the use of referee-assisting technology (FIFA, 2007; Warshaw, 2006). Their argument is that such means provide an illusion of technical objectivity when in fact they might induce other kinds of errors or distortions to the game. Thus, to this date, football referees often rely on a variety of (sometimes foul-unrelated) decision cues to make sense of ambiguous contact situations (Mascarenhas, O'Hare, & Plessner, 2006). Among such cues found to influence foul judgments in professional sports are, for example, the color of a player's uniform (Frank & Gilovich, 1988; Tiryaki, 2005), preceding foul judgments (Plessner & Betsch, 2001), a team's reputation (Jones, Paull, & Erskine, 2002), and crowd noise (Nevill, Balmer, & Williams, 2002). Another potential cue that might affect foul judgments has yet to be scientifically investigated: the height of the players involved in a foul.

Van Quaquebeke and Giessner are with the Rotterdam School of Management, Erasmus University, Rotterdam, Netherlands.

In this respect, embodiment theories of cognition provide a useful theoretical framework to argue for an automatic association between perceptions of height and dominance. These theories are based on recent advances in cognitive psychology demonstrating that humans ground their conceptual thinking perceptually (Barsalou, 1999, 2008; Glenberg, 1997; Niedenthal, Barsalou, Winkielman, Kraut-Gruber, & Ric, 2005). In other words, people's mental representations of abstract concepts (such as dominance and power) are embodied in modal information about space and the body (such as height). Hence, thoughts about power and dominance are automatically connected to a vertical dimension in space. When people think about power, they automatically interpret up, above, and large as cues for power, and down, below, and small as cues for powerlessness (Lakoff & Johnson, 1980). Such an association of physical height and dominance seems so deeply rooted in basic human cognitions that it is even present across cultures (Fiske, 1992). Recent empirical evidence provides additional support for these arguments by showing that concept and modality congruencies influence the ease of judgmental processes (Meier, Hauser, Robinson, Friesen, & Schjeldahl, 2007; Schubert, 2005; Schubert, Waldzus, & Giessner, 2009) as well as the memory of events and judgments (Giessner & Schubert, 2007).

By and large, such embodiment theorizing draws on an evolutionary perspective when arguing that height is associated with dominance and power (Cawley, Joyner, & Sobal, 2006; Goodall, 1986; Schuett, 1997). Indeed, this more or less instinct-driven association is a vital survival mechanism for many animal species. A decision to fight or flee, for example, is often determined by height comparison (Archer, 1988; Parker, 1974). Moreover, among humans the instinctive association between height and dominance (Schwartz, Tesser, & Powell, 1982; Weinraub & Putney, 1978) is reinforced through socially driven learning (Chaiken, 1986; Zebrowitz, 1996). Children, for instance, learn to associate the superiority of adults with adults' advantage in height (Montepare, 1995). The same height-dominance association is further consolidated through experience when taller children use their physical superiority to coerce smaller children (Olweus, 1978), or when taller adults use their physical advantage to demonstrate dominance and gain superiority (Felson, 1996; Judge & Cable, 2004).

The theoretical account described above thus provides a solid basis for the claim that the association between physical height and dominance is automatic and can influence judgment processes. However, while the cognitive association of height and dominance has received considerable attention in empirical research, it remains yet to be shown how such a cognitive association might affect individual decision making, in particular, decision making that is of practical relevance (see Baumeister, Vohs, & Funder, 2007).

The aim of the present set of studies was thus to investigate whether people's cognitive association of height and dominance will affect their judgments in sports. Particularly, we hypothesize an attribution bias in foul calls such that taller people are likely to be perceived as foul perpetrators and their smaller opponents as foul victims. We refer to this phenomenon as an "attribution bias" because it is an asymmetric cognitive association that we propose to systematically affect people's foul decisions. However, please note that such labeling only describes a decision tendency and not whether such tendency is erroneous or correct (which we will come back to in the discussion).

We chose football as the context of our studies because the sport often yields ambiguous foul situations in which it is difficult to determine the perpetrator (cf. Andersen, Engebretsen, & Bahr, 2004; Fuller, Junge, & Dvorak, 2004). In such situations, people must rely on their "instincts" to make a call, which should increase the use and thus the detectability of a height as an additional decision cue. Furthermore, the use of referee assistance technology and adequate referee training is frequently debated in association football. Thus, by providing scientific insights on potential biases in refereeing, our work might help officials weigh the options.

Overview of the Studies

To test our hypothesis that the association of height and dominance leads to a height-related attribution bias when judging football calls, we acquired professional football refereeing data (i.e., all recorded fouls in the last seven years of German Bundesliga, in the last seven years of the Union of European Football Associations (UEFA) Champions League, and in the last three FIFA World Cups). To triangulate these findings and provide more controlled evidence for the hypothesized effect, we additionally conducted two experiments in which we presented participants with ambiguous foul scenarios and asked for their foul judgments.

Study 1

Methods

Sample. To obtain evidence whether the proposed effect is at all visible in the field, we obtained official football foul statistics for seven German Bundesliga seasons, spanning from 2000/01 to 2006/07 (N = 85,262); for seven UEFA Champions League seasons, spanning from 2000/01 to 2006/07 (N = 32,142); and the last three FIFA World Cup tournaments—in 1998, 2002, and 2006 (N = 6,440). We purchased these records from Impire AG, a commercial organization that professionally collects and codes statistical data related to football. The data included individual fouls with information on the physical height of the perpetrator and victim of each foul. However, in 165 cases of the UEFA Champions League data set and in 369 cases of the FIFA World Championship data set, physical height information was missing.

Results

We conducted separate analyses on the Bundesliga, Champions League, and World Championship data sets. More precisely, we conducted a 2 (player height: perpetrator vs. victim) × 7 (season) ANOVA with repeated measures on the first factor on the Bundesliga and Champions League data sets. On the World Cups data set, we conducted a 2 (player height: perpetrator vs. victim) × 3 (tournament) ANOVA with repeated measures on the first factor.

The analysis of the German Bundesliga data set yielded a significant main effect of season, $F(6, 85{,}255) = 39.70$, $p < .001$, $\eta_p^2 = .003$, indicating that there are slight variations in the average players' heights involved in fouls between the different seasons. Actually, there seems to be a trend that the average player's height

increased over the years (as shown later in Table 3). More importantly, the analysis yielded a significant effect of player height, $F(1, 85{,}255) = 979.70$, $p < .001$, $\eta_p^2 = .01$. As hypothesized, foul perpetrators were on average taller ($M = 182.82$, $SD = 6.10$ cm) than foul victims ($M = 181.94$, $SD = 6.30$ cm). The analysis yielded no interaction effect, $F(6, 85{,}255) = 1.15$, $p = .33$, $\eta_p^2 < .001$. In addition, simple effect analyses within seasons indicated that the main effect of player height is significant within each season.

The analysis of the UEFA Champions League data set also yielded a significant main effect of season, $F(6, 31{,}970) = 17.43$, $p < .001$, $\eta_p^2 = .003$. Similar to the German Bundesliga data set, there is a trend that the average player's height increased over the years, with the exception of the last championship season (Table 3). More importantly, we again found a significant effect of player height, $F(1, 31{,}970) = 299.55$, $p < .001$, $\eta_p^2 = .01$. As hypothesized, foul perpetrators were on average taller ($M = 181.37$, $SD = 5.84$ cm) than foul victims ($M = 180.59$, $SD = 5.96$ cm). The analysis yielded no interaction effect, $F(6, 31{,}970) = 1.10$, $p = .36$, $\eta_p^2 < .001$. In addition, simple effect analyses within seasons indicated that the main effect of player height is significant within each season (Table 1).

Finally, the analysis of the FIFA World Cups data set also yielded a significant effect with respect to year of the world championship, $F(2, 6{,}068) = 3.00$, $p = .05$, $\eta_p^2 = .001$. Similar to the German Bundesliga and UEFA Champions League data sets, this might be a reflection of the fact that the average player's height competing in the World Cup seems to have increased over the years, with the exception of the last championship season (Table 1). More importantly for the objective of this article, we also found in the FIFA World Cup data set that foul perpetrators were on average taller ($M = 180.92$, $SD = 6.02$ cm) than foul victims ($M = 180.18$, $SD = 6.25$ cm), $F(1, 6{,}068) = 52.31$, $p < .001$, $\eta_p^2 = .009$. The analysis yielded no interaction effect, $F(2, 6{,}068) = 1.70$, $p = .18$, $\eta_p^2 < .001$. In addition, simple effect analyses within World Cups indicate that the main effect of player height is significant within all three World Cups (Table 1).

A possible concern regarding our field data results is that the relationship between height and fouls might actually result from the nature of players' respective tactical positions (i.e., goal keeper, defense, midfield, forward). Potentially taller defense players have to stop potentially smaller forward players, if necessary even by means of a foul. Table 2 provides some evidence for such a height difference between tactical positions; however, please note that because the present data set contains only information about players who were involved in a called foul play, we are unable to determine the average height of players on different tactical positions in general. To investigate the matter of tactical positions further, we thus plotted the amount of fouls by tactical position of perpetrators against the position of foul victims (Table 3). These results suggest that most of the fouls were actually committed by the relatively small midfield players fouling other midfield players. Moreover, we do find that defense players foul forward players the most, but also the reverse picture in that forward also foul defense players the most. Ultimately, to rule out that tactical positions might be the explaining factor responsible for the found effect of height, we created a match of position variable that indicates whether both players had the same tactical position (= same) or played different tactical positions (= different). We then ran additional analyses to check whether results are different in the same position versus different position condition.

Table 1 Foul Statistics in Association Football

Data Set	Player Height (in cm)				Inferential Statistics			
	Perpetrators		Victims					
	M	*SD*	*M*	*SD*	*N*	*F*	*p*	η_p^2
German Bundesliga								
2000/01	182.55	5.55	181.74	5.82	12,339	120.75	<.001	.001
2001/02	182.59	6.09	181.80	6.15	12,581	117.74	<.001	.001
2002/03	182.64	6.23	181.76	6.55	12,697	148.79	<.001	.002
2003/04	182.92	6.17	181.96	6.42	12,614	170.95	<.001	.002
2004/05	182.89	6.30	181.90	6.62	11,761	172.55	<.001	.002
2005/06	183.38	6.21	182.46	6.35	12,112	151.93	<.001	.002
2006/07	182.81	6.08	182.0	6.06	11,158	106.29	<.001	.001
UEFA Champions League								
2000/01	181.02	5.74	180.12	6.04	5292	68.28	<.001	.002
2001/02	181.41	5.80	180.46	5.82	5271	76.40	<.001	.002
2002/03	181.30	6.03	180.50	6.20	5087	50.95	<.001	.002
2003/04	181.25	5.77	180.59	5.68	4021	28.01	<.001	.001
2004/05	181.55	5.62	180.79	5.72	4223	39.07	<.001	.001
2005/06	181.78	5.81	181.05	5.99	4117	34.22	<.001	.001
2006/07	181.44	6.06	180.82	6.18	3966	24.00	<.001	.001
FIFA World Cup								
1998	180.84	6.33	179.84	6.30	1677	25.31	<.001	.004
2002	180.94	5.73	180.17	5.97	2205	19.71	<.001	.003
2006	180.96	6.08	180.45	6.46	2189	8.79	.003	.001

Table 2 Average Size of Perpetrator and Victims (in Centimeters) for Each Tactical Position Across All Data Sets

	Goal		Defense		Midfield		Forward		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Perpetrator	189.65	4.80	184.11	5.71	180.56	5.65	183.23	6.41	182.35	6.07
Victim	189.01	4.59	183.44	5.95	179.68	5.69	182.48	6.49	181.50	6.25
Total	189.06	4.62	183.82	5.82	180.11	5.69	182.82	6.47	181.93	6.18

Table 3 Number of Fouls Separated by Perpetrators' and Victims' Tactical Positions Across All Data Sets

	Position of Victim				
Position of Perpetrator	**Goal**	**Defense**	**Midfield**	**Forward**	**Total**
Goal	1	9	34	91	135
Defense	368	3,010	14,470	23,423	41,271
Midfield	329	10,743	34,215	9,282	54,569
Forward	858	17,226	8,86	1,099	27,869
Total	1,556	30,988	57,405	33,895	123,844

On the German Bundesliga and Champions League data sets, we conducted a 2 (player height: aggressor vs. victim) × 2 (position: same vs. different) × 7 (season) ANOVA with repeated measures on the first factor. On the World Championship data set, we followed the same design, except that we exchanged season with tournament and thus conducted a 2 (player height) × 2 (position) × 3 (tournament) ANOVA.

The analysis on the Bundesliga data set yielded a main effect of player height, $F(1, 85{,}248) = 912.79$, $p < .001$, $\eta_p^2 = .011$; of season, $F(6, 85{,}248) = 43.30$, $p < .001$, $\eta_p^2 = .003$; and of match of position, $F(1, 85{,}248) = 5{,}225.61$, $p < .001$, $\eta_p^2 = .06$. The last main effect is explained by the fact that players in the same position condition were on average smaller than in the different position condition because the same position condition contains mainly midfield players, who are on average smaller than players from other positions (we find this in all three field data sets). These main effects were qualified by two-way interactions of season × fit of position, $F(6, 85{,}248) = 7.16$, $p < .001$, $\eta_p^2 = .001$; player height × match of position, $F(1, 85{,}248) = 11.54$, $p = .001$, $\eta_p^2 < .001$; and a three-way interaction, $F(6, 85{,}248) = 2.46$, $p = .02$, $\eta_p^2 < .001$. The three-way interaction is explained by a significant player height × match of position interaction within the season 2002/2003, $F(1, 12{,}695) = 5.90$, $p = .02$, $\eta_p^2 < .001$, whereas all other interactions within season were nonsignificant. Importantly, however, even within the same position condition of all seasons, we found a significant simple main effect of player's height in the expected direction (see the table in the Appendix).

The analysis on the Champions League yielded a main effect of player height, $F(1, 31{,}963) = 236.25$, $p < .001$, $\eta_p^2 = .007$; of season, $F(6, 31{,}963) = 22.25$, $p < .001$, $\eta_p^2 = .004$; and of match of position, $F(1, 31{,}963) = 1068.93$, $p < .001$, $\eta_p^2 = .03$. The analysis yielded also a two-way interaction between season and match of position, $F(6, 31{,}963) = 7.16$, $p < .001$, $\eta_p^2 = .002$. No other significant effect

was found. Thus, the main effect of player height was not qualified by match of position (see also the Appendix table).

Finally, the analysis on the World Championship yielded only a significant effect of player height, $F(1, 6{,}065) = 42.39$, $p < .001$, $\eta_p^2 = .007$, and of match of position, $F(1, 6{,}065) = 376.62$, $p < .001$, $\eta_p^2 = .06$ (see Appendix for simple effects). Thus, overall these analyses consistently indicate that different tactical positions cannot explain away the height difference effect in foul decisions.

As part of further post hoc analyses, we sought to investigate whether the probability that taller players are more often assumed to be foul perpetrators is categorical or increases with height difference. As can be seen in Table 1 and in the Appendix table, the average difference between the aggressor and the victim is approximately 1 cm in every data set. However, this is just an average score. Indeed, it seems unlikely that referees can even perceive such small height differences—especially at a distance. We thus assumed that the bias of deciding for a foul against a taller (vs. small) player is stronger the greater the height difference between the players actually is.

To investigate this, we created a new variable. It represents whether the taller player was the aggressor (= 1) or not (= 0). Accordingly, we term the variable the *foul-against-the-tall-player probability*. We then regressed this new variable on the absolute height differences between the players involved in a foul. These analyses yielded the expected positive relationship across all data sets (overall: $r = .13$, $p < .001$; World Cup: $r = .14$, $p < .001$; European Champions League: $r = .12$, $p < .001$; German Bundesliga: $r = .12$, $p < .001$). Figure 1 illustrates this relationship further by depicting the probabilities along arbitrarily chosen categorical height

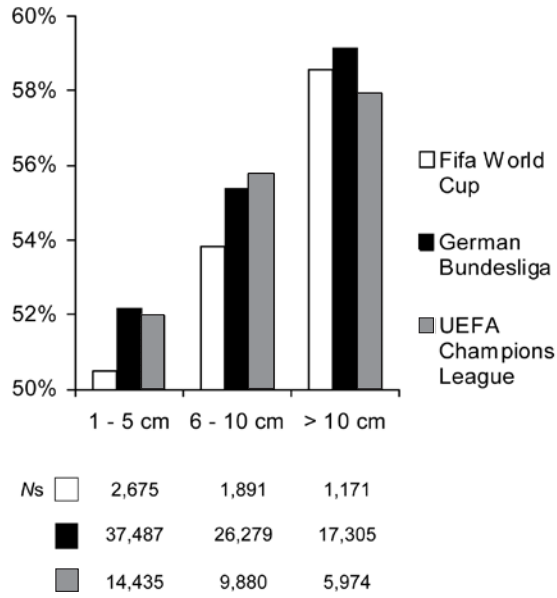


Figure 1 — Mean foul-against-the-taller-player probability depending on absolute height difference categories in Study 1.

differences. As predicted, the probability to call a foul against the taller player was on average lowest for absolute height differences of players between 1 cm and 5 cm, that is, 52.04%. This increased to 55.40% for height differences between 6 and 10 cm and even to 58.81% for absolute height differences above 10 cm.

Study 2

Although the field data of Study 1 provide initial important evidence corroborating our hypothesis, it was impossible to control for all possibly confounding factors. Especially, such field data are not able to rule out that taller players are actually committing more fouls. We thus designed an experiment in which no actual fouls are committed to investigate whether depicting height differences between players is already enough to inform a biased foul judgment.

Method

Participants. Participants for the experimental study were recruited via various German football-related Web sites (e.g., 11freunde.de, sport1.de, and sportal.de). They were informed that the study was about foul perceptions. One hundred twenty participants completed the experiment (87 male, 33 female; age $M = 30.04$, $SD = 8.72$ years). The majority of participants had no experience as referees ($n = 115$). Most of the participants, however, actively played soccer ($n = 95$), and most of these played in a leisure-time context ($n = 62$), whereas some of them (additionally) played on an amateur ($n = 56$) or professional level ($n = 2$). In addition, the majority of the participants indicated that they watch soccer on television ($n = 118$), including the German Bundesliga ($n = 84$), UEFA Champions League ($n = 84$), and FIFA World Cups ($n = 116$).

Procedures. The online experiment was set up as a between-subjects design. Participants saw a picture portraying two football players moving toward a ball in the middle of the picture. The ball was said to be lying still on the ground. The height of the two depicted players was randomly manipulated so that one player would always be bigger than the other (the pictures of both experimental conditions are depicted in Figure 2). We told participants that the player on the left (in some pictures being the bigger and in others being the smaller player) would fall to the ground after the portrayed scene. We then asked the participants to indicate the reasons why. Possible reasons for falling to the ground were (1) foul, (2) dive, (3) chance. The question was a forced-choice format for external validity reasons; that is, because referees, too, have to categorically decide if they are going to call a foul when someone fell or if they attribute the falling to another likely reason such as a dive or pure chance (Plessner, Schweizer, Brand, & O'Hare, 2009). Paralleling decisions in the field and in light of our theoretical rationale, we collapsed dive and chance decisions for our analyses into a "no foul" decision.

At the end of the experiment, we showed the picture that participants had seen in the beginning once more and asked participants to estimate the height of the players (in centimeters) as well as their muscularity (from 1 = *not muscular* to 5 = *very muscular*). Afterward, we asked participants for their demographic information (i.e., gender and age) as well as for their experience with refereeing, playing



Figure 2 — The two (potential) foul scenes presented to participants in Studies 2 and 3. (The height difference between players is 10% in two-dimensional space. Of the two pictures, only one was presented to each participant in a between-subjects design).

soccer at all, playing leisure-time soccer, playing amateur-level soccer, playing professional soccer, watching soccer on television, watching German Bundesliga, watching UEFA Champions league, and watching FIFA World Cups. All of these questions were to be answered with yes or no. At the end of the study, participants were debriefed about the hypothesis behind the study and could leave their e-mail addresses in a separate database if they wanted to be informed of the results.

Results

As intended by our manipulation, *t* test results indicate that participants perceived the tall player to be significantly taller compared with the small player, $t(119) = 17.70$, $p < .001$, $d = 1.84$ ($M = 184.34$, $SD = 4.90$ cm vs. $M = 174.76$, $SD = 5.50$ cm). Similarly, participants perceived the tall player to be significantly stronger compared with the small player, $t(119) = 7.40$, $p < .001$, $d = 1.14$ ($M = 3.56$, $SD = 0.83$ vs. $M = 2.59$, $SD = 0.87$).

Table 4 reveals the obtained distribution. Unfortunately, the automated randomization algorithm within the employed software produced a very unlikely uneven distribution. Thus, we cannot continue assuming that the expected marginal distributions are equal, but instead have to turn to a variant of the χ^2 test that bases the expected marginal distributions upon the number of participants subjected to

Table 4 Observed "Foul" / "No Foul" Decisions in Study 2

	Decision		
	Foul	**No foul**	**Total**
Player on the ground			
Tall player	8 (13.42)	27 (21.58)	35
Small player	38 (32.58)	47 (52.42)	85
Total	46	74	120

Note. Expected frequencies are in parentheses.

each manipulation. This specific χ^2 test uses a contingency table to test whether two (or more) variables are independent of each other. In other words, finding an insignificant result for the χ^2 test indicates that the two variables are independent, and conversely, finding a significant result indicates that the one variable in part depends on the other. The present unfortunate disparity in columns and rows thus does not represent any threat to this kind of χ^2 test statistics (Camilli & Hopkins, 1978; Overall, 1980). We consequently accepted the uncommon distribution and continued with our analyses.

Central to our hypothesis, the Pearson χ^2 test results indicate that observed foul attributions significantly differ from the expected frequencies, $\chi^2(1) = 5.00$, $p = .03$, $\varphi = .20$ (Table 4). In line with our reasoning, participants called fouls more often when the player said to be lying on the ground after the portrayed scene was the smaller one.

Study 3

The second experimental study was designed to address some potential shortcomings of the first. Whereas the scenario in Study 2 always stated that the player on left side of the picture would fall to the ground after the portrayed scene, we now left it to the participants to decide which player they anticipate to fall to the ground. By not dictating which player will fall to the ground, we not only eliminated any effect the left side might have had on foul judgments, but are also able to additionally assess participants' a priori evaluation of the situation, that is, whether participants already anticipate that one player would lie more often on the ground than the other.

Furthermore, one possible criticism of Study 2 might be that the manipulation of height combined with the abstract triangular background image affects participants' perceptions of the players' distance to the ball. Thus, if smaller players appear to be closer to the ball, it might drive the effect, that is, that the taller (= further away) player needs to foul to get to the ball. To control for this alterative explanation, we measured the perceived distance of the players from the ball.

Method

Participants. Participants for the second study were recruited via a professional online consumer panel. Again, they were informed that the study was about foul perceptions. A total of 196 participants completed the experiment (176 male, 20 female; age $M = 25.83$, $SD = 6.62$ years). The majority of participants did not have any experience at being referees ($n = 172$). Most of the participants, however, actively played soccer ($n = 177$). Most of these did so in a leisure time context ($n = 109$), while some of them (additionally) played on an amateur ($n = 88$) or professional level ($n = 2$). In addition, the majority of the participants indicated that they watch soccer on television ($n = 188$), including the German Bundesliga ($n = 153$), UEFA Champions League ($n = 118$), and FIFA World Cups ($n = 183$).

Procedures. In the second experiment, we showed the same pictures as in Study 2 (Figure 2) in a similar between-subjects design. We again randomly manipulated the height of the players. Unlike Study 2, however, we gave no indication of who would fall to the ground this time, but instead asked participants which one of the

two players they assumed to be lying on the ground after the portrayed scene and what they assumed to be the reason for that. Possible reasons for falling to the ground were again: (1) foul, (2) dive, (3) chance. Similar to Study 2, the question was again a forced-choice format and we collapsed dive and chance decisions into "no foul."

At the end of the experiment, we showed the picture participants saw at the beginning once more and asked participants to estimate the height of the players (in centimeters) as well as their muscularity (from 1 = *not muscular* to 5 = *very muscular*). In addition, we measured the perceived distance of the players from the ball, in that we asked participants to estimate the distance of each player from the ball (in centimeters). Afterward, similar to Study 2, we asked participants for their demographic information (i.e., gender and age) as well as for their experience with refereeing, playing soccer at all, playing leisure-time soccer, playing amateur soccer, playing professional soccer, watching soccer on television, watching German Bundesliga, watching UEFA Champions league, and watching FIFA World Cups. All of these questions were again to be answered with yes or no. At the end of the study, participants were debriefed about the hypothesis behind the study and could leave their e-mail addresses in a separate database if they wanted to be informed of the results.

Results

Study 3 confirmed the results of Study 2. Again, our manipulation check indicated that participants perceived the tall player as significantly taller, $t(195) = 18.95$, $p < .001$, $d = 1.76$, and stronger, $t(195) = 5.58$, $p < .001$, $d = .65$, than the small player (height $M = 183.92$, $SD = 4.69$ cm vs. $M = 175.39$, $SD = 5.03$ cm; strength $M = 3.43$, $SD = 0.92$ vs. $M = 2.81$, $SD = 0.99$). In addition, regarding a potential ball-distance effect, the analysis indicates that there was no significant difference in the perceived distance of both players to the ball, $t(195) = 1.70$, $p = .09$, $d = .10$ (distance, small player $M = 52.08$, $SD = 23.16$ cm; distance, tall player $M = 54.27$, $SD = 22.12$ cm).

Central for our hypothesis, participants anticipated more often that the small player would be lying on the ground after the portrayed scene ($n = 114$) compared with the tall player ($n = 82$), $\chi^2(1) = 5.25$, $p = .02$, $\varphi = .16$. Moreover, in line with Study 2, participants again attributed more often a foul when they anticipated the small player to fall to the ground compared with when they anticipated that the tall player would fall down, Pearson $\chi^2(1) = 15.71$, $p < .001$, $\varphi = .28$ (Table 5). Thus, both experimental studies provide consistent evidence that individuals use players' heights as a decision cue when having to make a decision in an ambiguous foul situation.

Table 5 Observed "Foul" / "No Foul" Decisions in Study 3

	Decision		
	Foul	**No Foul**	**Total**
Player on the ground			
Tall player	32 (45.6)	50 (36.4)	82
Small player	77 (63.4)	37 (50.6)	114
Total	109	87	196

Note. Expected frequencies are in parentheses.

General Discussion

Consistently, in the last seven years of German Bundesliga and UEFA Champions League as well as in the last three FIFA World Cups, we find that assumed foul perpetrators were on average taller than their assumed victims. These field findings are corroborated by more controlled experimental evidence. In this respect, Study 2 shows that people, when presented with an ambiguous foul situation, are more likely to attribute fouls to the taller player. Further, Study 3 shows that people, when presented with an ambiguous foul situation, anticipate that smaller players are more likely to go to the ground than their taller counterparts. In line with the above, Study 3 also shows that people tend to attribute an anticipated fall by a shorter player to a foul by the taller opponent. Conversely, they are more likely to attribute an anticipated fall by a taller player to a non-foul context (dive or chance). Moreover, we were able to rule out that this effect is either due to position (left or right side of the picture) or the players' perceived distance to the ball. In summary, the present results confirm a height-related attribution bias regarding foul calls in football. As such, these findings squarely extend previous studies that found that other decision cues—for instance, the color of a player's uniform (Frank & Gilovich, 1988; Tiryaki, 2005), preceding foul judgments (Plessner & Betsch, 2001), a team's reputation (Jones et al., 2002), and crowd noise (Nevill et al., 2002)—also affect foul judgments.

Apart from complementing previous research on foul decision cues, the present research also significantly extends previous research on embodied cognition. While previous research consistently showed a cognitive association between vertical space and power (e.g., Giessner & Schubert, 2007; Schubert, 2005), our results are among the first to provide direct evidence that such an association also impacts on actual decision-making processes (Baumeister et al., 2007). Moreover, as the embodiment perspective proved to be fruitful in the present case, it now appears valuable to investigate embodiment effects even more widely in sport decision making (cf. Plessner et al., 2009). In the future, studies could, for instance, investigate whether a player's facial features (cf. Fink, Neave, & Seydel, 2007), a players' agility in movements (cf. Hugill, Fink, Neave, & Seydel, 2009), or perceiving players running from left to right versus vice versa (cf. Maass & Russo, 2003) have an effect on perceived player strength and thus also on foul decisions. Moreover, future studies could potentially even extend their theoretical reasoning to the area of embodied emotions (Niedenthal, 2007) and their mediating effects on players' and referees' performance alike. In this respect, one could, for instance, investigate the effects of players' body postures (cf. Stepper & Strack, 1993), a referee's holding of a whistle in the mouth (cf. Strack, Martin, & Stepper, 1988), or the effect of twitched eyebrows, for instance, due to playing against the sun (cf. Larsen, Kasimatis, & Frey, 1992).

Despite the exciting avenues the present research might open for future studies, the present results, however, also require further discussion, in particular, regarding (a) how the relatively small effect sizes in the field data are to be interpreted, (b) if the results support the argument of a misjudgment or are in fact a reflection of an adaptive strategy (i.e., taller people also commit more fouls), and (c) what the implications for practice are.

Let us commence with the first issue. One should bear in mind that height is only one of the many decision cues that referees process when perceiving a potential foul (see Plessner et al., 2009, for multicue information processing). The most informative of those cues would often be the actual foul behavior. This means that height assumingly only becomes informative as an additional decision cue when the primary decision cue, namely, the concrete foul behavior, yields no clear decision result (cf. search and stop rules in heuristics, Gigerenzer & Goldstein, 1996). Since ambiguous fouls represent only a fraction of actual fouls in football (see Andersen et al., 2004; Fuller et al., 2004), the present results thus imply that the effect size would be much higher when clear fouls, such as hitting, spitting, holding, or dangerous play, could be partialled out.

Likewise, we would argue that the effects in our field data might be additionally underestimated because many of the fouls smaller players commit are likely to be perceived as a taller player's dive or as mere chance that a taller player has tripped (as our experimental data suggest). These fouls are thus not called and thus also missing from the obtained field data sets.

A third and final reason to expect the field effect sizes to be underestimated is due to the fact that only visible height differences are likely to have an impact on the foul decision-making process. The present data support that notion. As can be gathered from Figure 2, almost half of the fouls are committed between players who did not differ more than 5 centimeters in height and are thus likely not to be detected as such (cf. the Weber–Fechner law on just noticeable difference). Hence, in this category, height difference unsurprisingly has the smallest effect on foul attributions (and thus weakens the overall effect size). Looking further into psychophysics research on visual acuity suggests that perceptions of height are moreover affected by a plethora of other factors, including perception angle, light intensity, and self and other movements (Bruce, Green, & Georgeson, 2003; Sedgwick, 2001). Because the present data are field data, it is likely that one or more of these factors were involved and thus reduce the chance to accurately detect a height difference—which would additionally weaken the overall effect size.

Given the above reasons that lead to an underestimation of effect size, it is the consistency of the effect across all 17 subsamples (and its accordance with results in both experiments) that is impressive and can be taken as firm evidence that height and dominance are not only associated on a cognitive level, but also that this association affects people's individual decision making (i.e., a behavioral outcome) in the "real world" (Baumeister et al., 2007).

Regarding the second issue, of course, it can be argued that bigger people might actually be more aggressive and thus objectively commit more fouls (Felson, 1996; Olweus, 1978). Indeed, because embodied cognitions are highly abstracted cognitive derivations of "real-life" patterns, we would not argue much differently. If taller people are more aggressive, it is only adaptive to associate height with aggression because it unconsciously enables people to behave differently toward taller people and thus, for instance, avoid potential harm (cf. "ecological rationality" in Goldstein & Gigerenzer, 2002). However, given the present data, we are unable to ultimately determine whether the height-related foul attribution bias is an adaptive strategy or an unjustified misjudgment (Kahneman, Slovic, & Tversky, 1982). Moreover, it goes beyond the scope and the objective of the current study, which was to confirm that a height-related attribution bias in football foul calls exists.

To determine whether the identified bias is adaptive or not should therefore be the subject of future studies in which, for instance, several football matches including the refereeing could be filmed from different angles. Subsequent expert coding, taking all camera perspectives into account, should then be able to reveal not only whether tall players objectively commit more fouls than smaller players, but also reveal the extent to which a referee's decision strategy is a reflection of that. In doing so, one should, however, bear in mind that even if the expected correlation between height and foul behavior were to be found, it likely does not explain the complete variance. In other words, the association of height and foul attributions might prove to be a good rule of thumb for a referee (i.e., show high cue validity), but, as rules of thumb go, there might be instances in which they are not right and thus unfair to employ.

This leads to the third issue: What can or should be done about the identified bias in refereeing? On the one hand, victims of uncalled or mistakenly called fouls legitimately want any foul to be called as accurately as possible. After all, an unwarranted or uncalled foul can in some cases determine the result of a game—for instance, when a penalty shot is (not) given or a player has to leave the field—and it can ultimately determine whether a team is relegated or whether it qualifies for an upcoming championship, and both of these are important for a club's financial performance (at least in professional clubs) (Plessner & Betsch, 2001; Sutter & Kocher, 2004). Thus, there is a legitimate argument to be made that each and every foul should be called as objectively as possible—despite the obtained effect size, which suggests that height perceptions are only occasionally responsible for potential foul misjudgments in the field.

On the other hand, some might argue that not all fouls need to be called 100% correctly. With regard to the height-related bias, one might even argue that, since all teams are likely to have some taller and some smaller players, there is probably no systematic disadvantage for any one team in a league. Thus, what is to one's team advantage in one moment might be to their opponent's advantage in the other. Even in cases where there might be systematic height differences such as in World Cups, one might point out that this systemic disadvantage against taller players regarding foul attributions might be leveled out by systematic height advantages in other instances, such as when height is advantageous for being able to head the ball or when a players' height is good for shielding the goal or intercepting corner shots. Given this potential balance between advantages and disadvantages, one might argue that the role of the referee is not only to ensure a fair game, but also to help preserve the natural flow of the game by minimizing interference. To quote a FIFA training manual, "Constant whistling for minor and dubious infringements can cause bad feeling and loss of temper from players and spectators." (FIFA, 2004).

Ultimately, both sides thus have a valid point. The question is how to resolve it. Whereas one could argue that referee assistance technology, such as multiple camera angles in slow motion, could increase the chance to determine the "true" perpetrator in ambiguous foul situations, FIFA president Joseph S. Blatter is very specific concerning the harm such practices would do to the game when stating, ". . . we must never stop the match with videos or monitors to look at what has happened. It will take away the spontaneity and fascination of our game; we must keep football with a human face." (Warshaw, 2006). Other routes to increase foul decision accuracy should thus be taken. One viable alternative could lie in the training of

referees. Plessner and colleagues (2009), for example, describe such a training in which they present ambiguous foul situations to football referees. These situations have been previously rated according to what the adequate referee decisions would be. Ultimately, it is the objective of their training to immunize referees against the undesired influence of irrelevant cues. Given the present evidence, it would be interesting to integrate ambiguous foul situations with players of different height in such a training to explore whether teaching about the height-based attribution bias and giving feedback on foul decisions right after each call could immunize referees against height as a decision cue—at least to the extent that it is a misjudgment.

With the current study, we have shown that refereeing in football has, to quote Joseph S. Blatter, a very "human face." Indeed, referees are not objective and perfect information processors, but human and thus are subject to socially learned and evolutionarily formed cognitive associations that sometimes bias their judgment. In the end, whereas the specific effect of a cognitive height-dominance association on actual behavior is in itself noteworthy and extends the scientific literature, it is, however, not our call if or how these insights will be used in practice.

Appendix

Extended Foul Statistics in Association Football Split by Same Versus Different Tactical Positions

		Player Height (in cm)							
		Perpetrators		Victims		Inferential Statistics			
Data Set	Tactical Position	*M*	*SD*	*M*	*SD*	*N*	*F*	*p*	η_p^2
German Bundesliga									
2000/01	Same	181.03	5.45	180.28	5.75	4267	36.65	<.001	>.001
	Different	183.35	5.44	182.51	5.71	8072	84.39	<.001	.001
2001/02	Same	180.81	5.63	179.89	5.66	3842	48.16	<.001	.001
	Different	183.37	6.12	182.64	6.17	8739	70.89	<.001	.001
2002/03	Same	180.89	5.80	179.72	6.06	3707	75.14	<.001	.001
	Different	183.37	6.25	182.60	6.57	8990	79.77	<.001	.001
2003/04	Same	181.40	5.94	180.23	6.06	3801	77.10	<.001	.001
	Different	183.58	6.16	182.71	6.43	8812	97.56	<.001	.001
2004/05	Same	180.99	5.84	180.13	6.03	3495	38.58	<.001	>.001
	Different	183.70	6.31	182.65	6.71	8266	135.29	<.001	.002
2005/06	Same	182.07	6.02	180.93	6.21	3635	70.34	<.001	.001
	Different	183.94	6.21	183.12	6.29	8477	85.43	<.001	.001
2006/07	Same	181.32	5.84	180.16	5.67	3427	68.61	<.001	.001
	Different	183.48	6.06	182.84	6.19	7731	47.23	<.001	.001
UEFA Champions League									
2000/01	Same	179.53	5.14	178.75	5.38	1629	15.70	<.001	>.001
	Different	181.68	5.86	180.72	6.22	3663	53.13	<.001	.002
2001/02	Same	179.98	5.59	179.14	5.41	1639	18.00	<.001	.001
	Different	182.06	5.78	181.05	5.90	3632	58.97	<.001	.002

(continued)

Appendix *(continued)*

Data Set	Tactical Position	Player Height (in cm)				Inferential Statistics			
		Perpetrators		Victims					
		M	*SD*	*M*	*SD*	*N*	*F*	*p*	η_p^2
2002/03	Same	179.97	5.88	179.06	5.67	1652	21.71	<.001	.001
	Different	181.93	6.00	181.19	6.34	3435	29.76	<.001	.001
2003/04	Same	179.91	5.52	179.38	5.29	1339	5.81	.016	>.001
	Different	181.92	5.77	181.18	5.77	2682	22.82	<.001	.001
2004/05	Same	180.74	5.31	180.09	5.43	1282	8.49	.004	>.001
	Different	181.91	5.72	181.10	5.82	2941	30.99	<.001	.001
2005/06	Same	180.65	5.58	180.32	5.50	1197	2.02	.16	>.001
	Different	182.24	5.58	181.35	6.16	2920	36.43	<.001	.001
2006/07	Same	180.41	5.95	179.81	6.19	1273	7.35	.007	>.001
	Different	181.92	6.05	181.30	6.12	5693	16.65	<.001	.001
FIFA World Cup									
1998	Same	179.23	6.08	178.23	5.98	571	8.71	.003	.001
	Different	181.67	6.30	180.67	6.31	1106	16.59	<.001	.003
2002	Same	179.27	5.69	178.81	6.14	704	2.31	.13	>.001
	Different	181.72	5.58	180.81	5.77	1501	18.84	<.001	.003
2006	Same	179.26	5.31	178.80	5.67	708	2.22	.14	>.001
	Different	181.78	6.26	181.24	6.67	1481	6.62	.01	.001

References

Andersen, T.E., Engebretsen, L., & Bahr, R. (2004). Rule violations as a cause of injuries in male Norwegian professional football: Are the referees doing their job. *American Journal of Sports Medicine, 32,* 625–685.

Archer, J. (1988). *The behavioural biology of aggression*. Cambridge, UK: Cambridge University Press.

Barsalou, L.W. (1999). Perceptual symbol systems. *The Behavioral and Brain Sciences, 22,* 577–609.

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology, 59,* 617–645.

Baumeister, R.F., Vohs, K.D., & Funder, D.C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior. *Perspectives on Psychological Science, 2,* 396–403.

Bruce, V., Green, P. R., & Georgeson, M. A. (2003). *Visual perception: Physiology, psychology and ecology*. Sussex: Psychology Press.

Camilli, G., & Hopkins, K.D. (1978). Applicability of chi-square to 2 X 2 contingency tables with small expected cell frequencies. *Psychological Bulletin, 85,* 163–167.

Cawley, J., Joyner, K., & Sobal, J. (2006). Size matters: The influence of adolescents' weight and height on dating and sex. *Rationality and Society, 18,* 67–94.

Chaiken, S. (1986). Physical appearance and social influence. In C.P. Herman, M. Zanna, & E.T. Higgins (Eds.), *Physical appearance, stigma, and social behavior* (Vol. 3, pp. 143–177). Hillsdale, NJ: Lawrence Erlbaum.

Felson, R.B. (1996). Big people hit little people: Sex differences in physical power and interpersonal violence. *Criminology, 34,* 433–452.

FIFA. (2004). Law 5: The referee. Retrieved September 27, 2007, from http://www.fifa.com/mm/document/afdeveloping/refereeing/2.%20law%205_540.pdf

FIFA. (2007). Platini rules out video replays. *Press releases from the 122nd Annual General Meeting of the International Football Association Board.* Retrieved September 27, 2007, from http://www.fifa.com/worldfootball/news/newsid=606659.html

Fink, B., Neave, N., & Seydel, H. (2007). Male facial appearance signals physical strength to women. *American Journal of Human Biology, 19,* 82–87.

Fiske, A.P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review, 99,* 689–723.

Frank, M.G., & Gilovich, T. (1988). The dark side of self- and social perception: Black uniforms and aggression in professional sports. *Journal of Personality and Social Psychology, 54,* 74–85.

Fuller, C.W., Junge, A., & Dvorak, J. (2004). An assessment of football referees' decisions in incidents leading to player injuries. *American Journal of Sports Medicine, 32,* 175–225.

Giessner, S.R., & Schubert, T.W. (2007). High in the hierarchy: How vertical location and judgments of leaders' power are interrelated. *Organizational Behavior and Human Decision Processes, 104,* 30–44.

Gigerenzer, G., & Goldstein, D.G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103,* 650–669.

Glenberg, A. (1997). What memory is for. *Brain and Behavioral Sciences, 20,* 1–55.

Goldstein, D.G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109,* 75–90.

Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge, MA: Harvard University Press.

Hugill, N., Fink, B., Neave, N., & Seydel, H. (2009). Men's physical strength is associated with women's perceptions of their dancing ability. *Personality and Individual Differences, 47,* 527–530.

Jones, M.V., Paull, G.C., & Erskine, J. (2002). The impact of a team's aggressive reputation on the decisions of association football referees. *Journal of Sports Sciences, 20,* 991–1000.

Judge, T.A., & Cable, D.M. (2004). The effect of physical height on workplace success and income: Preliminary Test of a theoretical model. *The Journal of Applied Psychology, 89,* 428–441.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, London: The University of Chicago Press.

Larsen, R.J., Kasimatis, M., & Frey, K. (1992). Facilitating the furrowed brow: An unobtrusive test of the facial feedback hypothesis applied to unpleasant affect. *Cognition and Emotion, 6,* 321–338.

Maass, A., & Russo, A. (2003). Directional bias in the mental representation of spatial events. *Psychological Science, 14,* 296–301.

Mascarenhas, D.R.D., O'Hare, D., & Plessner, H. (2006). The psychological and performance demands of association football refereeing. *International Journal of Sport Psychology, 37,* 99–120.

Meier, B.P., Hauser, D.J., Robinson, M.D., Friesen, C.K., & Schjeldahl, K. (2007). What's "up" with god? Vertical space as a representation of the divine. *Journal of Personality and Social Psychology, 93,* 699–710.

Montepare, J. (1995). The impact of variations in height on young children's impressions of men and women. *Journal of Nonverbal Behavior, 19,* 31–47.

Nevill, A.M., Balmer, N.J., & Williams, M. (2002). The influence of crowd noise and experience upon refereeing decisions in football. *Psychology of Sport and Exercise, 3,* 261–272.

Niedenthal, PM. (2007). Embodying emotion. *Science, 316,* 1002–1005.

Niedenthal, P.M., Barsalou, L.W., Winkielman, P., Kraut-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review, 9,* 184–211.

Olweus, D. (1978). *Aggression in the schools: Bullies and whipping boys*. Washington, D.C.: Hemisphere.

Overall, J.E. (1980). Power of chi-square tests for 2 X 2 contingency tables with small expected frequencies. *Psychological Bulletin, 87,* 132–135.

Parker, G.A. (1974). Assessment strategy and the evolution of fighting behaviour. *Journal of Theoretical Biology, 47,* 223–243.

Plessner, H., & Betsch, T. (2001). Sequential effects in important referee decisions: The case of penalties in soccer. *Journal of Sport & Exercise Psychology, 23,* 254–259.

Plessner, H., Schweizer, G., Brand, R., & O'Hare, D. (2009). A multiple-cue learning approach as the basis for understanding and improving soccer referees' decision making. *Progress in Brain Research, 174,* 151–158.

Schubert, T.W. (2005). Your highness: Vertical positions as perceptual symbols of power. *Journal of Personality and Social Psychology, 89,* 1–21.

Schubert, T.W., Waldzus, S., & Giessner, S.R. (2009). Control over the association of power and size. *Social Cognition, 27,* 1–19.

Schuett, G.W. (1997). Body size and agonistic experience affect dominance and mating success in male copperheads. *Animal Behaviour, 54,* 213–224.

Schwartz, B., Tesser, A., & Powell, E. (1982). Dominance cues in non-verbal behavior. *Social Psychology Quarterly, 45,* 114–120.

Sedgwick, H.A. (2001). Visual space perception. In E.B. Goldstein (Ed.), *Blackwell handbook of perception*. Malden, MA: Wiley-Blackwell.

Stepper, S., & Strack, F. (1993). Proprioceptive determinants of emotional and nonemotional feelings. *Journal of Personality and Social Psychology, 64,* 211–220.

Strack, F., Martin, L.L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology, 54,* 768–777.

Sutter, M., & Kocher, M.G. (2004). Favoritism of agents - The case of referees' home bias. *Journal of Economic Psychology, 25,* 461–469.

Tiryaki, M.S. (2005). Assessing whether black uniforms affect the decisions of Turkish soccer referees: Is the finding of Frank and Gilovich's study valid for Turkish culture? *Perceptual and Motor Skills, 100,* 51–57.

Warshaw, A. (2006, November 28). Blatter vows to block use of video evidence in the game. *The Independent*.

Weinraub, M., & Putney, E. (1978). The effects of height on infants' social responses to unfamiliar persons. *Child Development, 49,* 598–603.

Zebrowitz, L.A. (1996). Physical appearance as a basis of stereotyping. In C.N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 79–120). Hillsdale, NJ: Lawrence Erlbaum.

Manuscript received: April 6, 2009
Revision accepted: September 3, 2009