

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/102836/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Souchon, Nicolas, Livingstone, Andrew G., Bardin, Brigitte, Rasclé, Olivier, Cabagno, Geneviève and Maio, Gregory R. 2016. Influence of competition level on referees' decision-making in handball. *Social Influence* 11 (4) , pp. 246-258. 10.1080/15534510.2016.1259656

Publishers page: <http://dx.doi.org/10.1080/15534510.2016.1259656>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Running head: COMPETITION LEVEL AND REFEREEING

Influence of Competition Level on Referees' Decision Making in Handball

Nicolas Souchon

Centre de Recherche sur le Sport et le Mouvement (CeRSM), UFR STAPS,
Université de Paris Ouest, France
200 Avenue de la République, 92001 Nanterre Cedex.
nicolas.souchon@u-paris10.fr

Andrew G. Livingstone

School of Psychology,
University of Exeter, England, UK
A.Livingstone@exeter.ac.uk

Brigitte Bardin

Cognition, Langues, Langage, Ergonomie (CLLE), UFR de psychologie
Université de Toulouse
brigitte.bardin@univ-tlse2.fr

Olivier Rascle and Geneviève Cabagno

Violences, Identités Politiques & Sports (VIPS)
Université de Rennes 2, France
olivier.rascle@uhb.fr, genevieve.cabagno@uhb.fr

Gregory R. Maio

Values In Action (ViA), School of Psychology
Cardiff University, Wales, UK
maio@cardiff.ac.uk

Leading author : Nicolas Souchon

Word count : 5156

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper

Abstract

The influence of competition level on referees' decision making was investigated. Referees' decisions in 90 handball games (30 games X 3 competition levels) were observed in different situations related to the advantage rule, and 100 referees from two different level of expertise were subsequently asked to offer explanations for the competition-level effects from the first part of the study. Results revealed that at the highest level of competition referees intervened less frequently with sporting sanctions, but more frequently with disciplinary sanctions. These effects were apparent mainly in immediate intervention situations and unsuccessful advantage situations, but not in successful situations. Referees explained these effects of competition level in terms of a player competence stereotype, in addition to referees' different expertise across competition level. The implications of the findings for understanding how status-related stereotypes impact on intervention behavior are discussed.

Key words: Competition level, Gender, Judgmental heuristics, Refereeing, Sport

The Influence of Competition Level on Referees' Decision Making in Handball

Decision making in complex settings such as aviation, military, fire fighting or refereeing in team contact sports requires individuals to make vital or consequential decisions very quickly, and decisions have been found to be affected by numerous processes and heuristics (e.g., Betsch & Haberstroh, 2005). With regard to refereeing in team contact sport, several studies have tested the impact of gender stereotypes on decisions (e.g., Coulomb-Cabagno, Rascle, & Souchon, 2005; Souchon, Livingstone, & Maio, 2013). However, little is known about the role of other factors such as competition level and associated stereotypes in shaping referees' decisions (see Debanne & Fontayne, in press; Souchon, Cabagno, Traclet, Trouilloud, & Maio, 2009, for exceptions). Moreover, existing studies involved observations of only a limited number of matches (Souchon et al., 2009) or used an archival method (Debanne & Fonayne, in press). The primary aims of the present study were (a) to systematically observe referees' decisions at different competition levels, (b) to provide the first specific examination of how competition-level may shape referees' decisions, and (c) to measure referees' explicit stereotypes related to competition level.

Refereeing Decisions

According to the advantage rule in team contact sports, referees should not intervene with sporting sanctions (e.g., free kick in soccer) following a "foul" by the defending team if the attacking team would benefit from letting play continue. Alternatively, if the attacking team is disadvantaged, referees should return the ball to the victim (e.g., 9-meter throw in handball). During this process, referees have to make decisions quickly regarding whether to let play continue by evaluating the victim's ability to keep going *and* by assessing the severity of the transgression. Moreover, referees have to decide whether to punish aggressive and dangerous actions through disciplinary sanctions (e.g., temporary exclusion).

Referees thus tend to use different judgmental heuristics that can help guide judgments

or bias decision making (Plessner & Haar, 2006). Such cognitive shortcuts include the color of players' shirt (Frank & Gilovich, 1988), the passage of the game (e.g., Unkelbach & Memmert, 2008), the noise of the crowd (e.g., Nevill, Balmer, & Williams, 2002), or the players' aggressive reputation (Jones, Paull, & Erskine, 2002).

There is increasing evidence that referees also use stereotypes – defined here as beliefs and associations individuals develop toward members of social categories (e.g., Schneider, 2004) – to guide decisions. For example, referees tend to develop expectations that female players are less skillful and aggressive than male players, and when faced with very similar situations involving male or female players, referees at different levels of expertise sanction female players more than male players (Souchon et al., 2013). In the present research, we aimed to extend this literature by testing whether referees' decisions differ as a function of competition level, and how these relate to stereotypes of player competence at different levels of competition (Souchon et al., 2009).

Stereotyping and competition levels

According to the stereotype content model (SCM; Fiske, Cuddy, Glick, & Xu, 2002), two basic dimensions – competence and warmth – underlie the content of stereotypes. The dimension of competence is dependent on perceived group status, with high-status groups being stereotyped as competent. In contrast, competitive groups are stereotyped as lacking warmth. The model predicts that individuals are more likely to help member of low-status groups stereotyped as warm, and be less helpful toward more competitive groups stereotyped as cold (Cuddy, Fiske, & Glick, 2008).

As status in sport is defined in large part through competition level, referees may perceive high-level players to be “competent” (e.g., skillful) and “cold” (e.g., aggressive, argumentative, dishonest), relative to lower-level players (see for example Conroy, Silva, Newcomer, Walker, and Johnson, 2001 on the relation between level and perception of

legitimacy of dangerous actions). Consequently, referees may be less helpful or benevolent toward high-level players than low-level players. Evidence consistent with this comes from analyses of referees' stereotypes based on player gender (Souchon et al., 2013).

Overall, the SMC would predict that referees are harsher toward highest-level players, such that physical contact between players would need to be severe before referees would feel the need to stop the game or return back the ball to the attacking players. Moreover, stereotypes guide cognitive interpretation of ambiguous information in order to confirm stereotype expectations (e.g., Darley & Gross, 1983). If referees perceive high-level players to be highly skillful, they may also anticipate that a player in possession of the ball could continue his or her action despite being victim of several fouls under the advantage rule (i.e., allowing the game to continue without intervention).

Accordingly, a recent study revealed that referees punished transgressions by highest-level players less severely with sporting and disciplinary sanctions than those made by lowest-level players (Souchon et al., 2009). Nevertheless, this study focused on a limited number of observations: 15 matches at the intermediate vs. 15 matches at the highest national level within the male championship. Also, those authors focused only on referees' reactions toward players' transgression and did not specifically analyze the conditions under which referees may use stereotypes of players at different levels of competition. However, Debanne and Fontayne (in press) found using archival data (i.e, without directly observing players' aggressive behaviors) that referees actually gave *more* disciplinary punishment to male handball players at the European level of competition than lower levels, in contrast to Souchon et al. (2009)'s results. The general aim of the present study was therefore to analyze in a more systematic way the influence of competition level on referees' decisions.

Conditions under which referees apply stereotypes of players

Handball players on the attacking team have to develop a collective strategy to create

free space within the defensive team. During these phases of organised attack, the player in possession of the ball may be victim of one or several illegal actions from the defensive team (e.g., being pushed) and be consequently be blocked, miss their pass or shot, or carry on to be successful in their pass or shot despite the transgression.

Under the advantage rule, referees in situations in which the attacking player could be blocked have to determine if they should let the play continue or intervene. As these situations are ambiguous, referees may consciously or unconsciously apply stereotypes. For example, Souchon et al. (2010) found that referees at an intermediate level intervened immediately more often for female players than male players, and suggested that this may have been because referees applied their gender stereotypes related to sporting skill (i.e., that female players are less able to successfully continue after a transgression).

Concerning the effect of competition level on decisions, we expected in the present study that referees would tend to intervene sooner with lower-level players than with higher-level players. This is because referees are likely to regard players at the highest level to be more able to continue after a transgression than at lower levels (Souchon et al., 2009).

Souchon et al. (2010) also suggest that two different scenarios can arise if the referee does not intervene immediately after a transgression: the attacking player is successful or unsuccessful in his or her pass or shot. In theory, referees should not intervene if the attacking team gains an advantage. Consequently, the first scenario presents no ambiguity and means that referees have less need to apply their stereotype. In these cases, there should be no effect of competition level or gender on referees' decisions. Consistent with this, Souchon et al. (2010) found no effect of player gender on male referees' decisions in such "successful situations", and intervention was very rare for both genders.

In contrast, referees' stereotypes may be more relevant when the attacking player is unsuccessful. Observations (Souchon et al., 2010) and experiments (Souchon et al., 2013)

revealed that referees in unsuccessful situations tend to be more benevolent toward female players, giving back the ball more frequently to female players than male players, and explaining this in terms of stereotypical expectations. Similarly, referees might be more benevolent toward low-level than higher-level players. We therefore expected that referees would intervene more frequently at the lowest than at the highest levels. Moreover, referees would explain this tendency in terms of competence stereotypes of the players.

Effects on Disciplinary Sanctions

Concerning disciplinary sanctions, Souchon et al. (2010) found a player gender effect in situations that involved a “failure” (i.e., immediate intervention or unsuccessful situations), but not a “success”. These former situations may be perceived to be more dangerous by referees, who are subsequently more likely to apply their stereotypes when making disciplinary decisions. For example, Souchon et al. (2013) found experimentally that referees tended to punish female player more severely than male players in failure situations (see also Souchon, Coulomb-Cabagno, Traclet, & Rascle, 2004). Concerning the influence of competition level, observational data have so far been mixed, indicating either that referees punished aggressive players with disciplinary sanctions less (Souchon et al., 2009) or more (Debanne & Fontayne, in press) at the highest competition level than at the intermediate competition level.

Summary, aims and predictions

The aims of the present research were to extend the approach of Souchon et al. (2009) and Debanne and Fonatyne (in press) by examining the effect of competition level on referees’ decisions, and the extent to which referees explain such decisions in terms of competition-level stereotypes. Based on the SCM, we predicted that for similar situations, referees would intervene less frequently with sporting sanctions with higher-level players than with lower-level players, based on the former’s perceived ability to make use of the advantage

rule (i.e., the high competence component of the stereotype). However, for *disciplinary* sanctions, it may be that referees are actually more severe with higher-level players, due to the concurrent stereotype of these players as being aggressive and competitive (i.e., the low warmth component of the stereotype).

Method

Participants

The main experimental study examined 90 matches from the period 2001-2003 in the French Handball Championships with the agreement of the French Handball Federation. Thirty matches at the highest local (i.e., ‘excellence departmental’), 30 matches at the intermediate (i.e., ‘pre-national’) and 30 matches at the highest national level of competition (i.e., ‘première division’) were videotaped in both the male and the female championships (15 matches X 2 player gender at each competition level). Matches included as many different referees (30 individuals referees within the local championships; 27 different pairs of referees within the intermediate championships; 27 different pairs of referees within the national championships) and teams (92 in total: 28 local teams, 32 intermediate teams, and 32 national teams) as possible. Referees were all men.

The second part of the study involved a fresh sample of 100 referees (age $M = 35.52$, $SD = 11.1$; experience $M = 10.81$, $SD = 6.84$, 97 men and 3 women) from two level of expertise (50 intermediate level referees: $M = 36.21$, $SD = 10.91$, experience $M = 11.5$, $SD = 6.51$; and 50 national referees: age $M = 34.89$, $SD = 10.78$; experience $M = 10.12$, $SD = 4.52$).

Procedure

The “attacking team” is defined as the team in possession of the ball. Handball referees can return the ball to the attacking team through a sporting sanction when at least one defensive opponent displays a transgression (International Handball Federation, 2005). Nonetheless, referees must not intervene according to the advantage rule before the player has

lost possession of the ball or cannot pursue their actions because of the transgression (IHF, 2005, Rule 13.2). We therefore focused our observations on transgressions committed against players in possession of the ball.

Our observations also centred on what we define as ‘organized attack’ situations, in which a player in possession finds himself or herself behind a line of at least four opposing defenders. This is based on a pilot sample and previous research (Souchon et al., 2004) which indicated that handball referees perceive such attacks to comprise the main part of handball games, and to be more physically and technically demanding than counter attacks. They are thus the best situations to observe the application of the advantage rule. Attackers’ fouls were not measured due to their shortage.

Players’ transgressions. Observation criteria for identifying transgressions strictly followed the rules of handball. These include any ‘pushing’, ‘bumping into’, ‘pushing away’ ‘holding back’, ‘catching and holding’, or ‘seizing the player with possession around the waist’ (rule 8.2, IHF, 2005). Multiple transgressions were recorded when a player in possession was victim of two or more transgressions before passing the ball or shooting.

Handball refereeing decisions. For each observed transgression, we recorded referees’ decisions: i.e., sporting (9-m throw vs. 7-m throw: direct shot at the goal) and/or disciplinary sanctions (yellow card vs. 2-min suspension vs. red card). Advantage rule application was inferred from a referee’s decision not to sanction an observed transgression.

Type of situations. Immediate intervention situations, unsuccessful situations, and successful situations were observed as in Souchon et al. (2010). In immediate intervention situations, the referee intervened instantaneously following the defensive transgression. In unsuccessful situations, the player missed a pass or shot after the defensive transgression. In successful situations, the player accomplished a pass or shot, despite the defensive transgression.

Coding

Before undertaking our final observations, one game for each level in the male championship and in the female championship were observed by three people, including two handball experts ($M_{\text{age}} = 40$ years, $M_{\text{Experience}} = 12$ years officiating at the highest national level), and the main author. Two other games were observed two weeks later by the main author. The Kappa coefficients (i.e., agreed inter-observer and intra-observer coefficient) between 0.85 and 0.95 were satisfactory. Each game was then observed by the main author and one of the two handball experts, with agreement rates of between 0.87 and 0.92. As in Souchon et al. (2010), both observers recorded players' transgressions, referees' decisions and type of situations. One single measure for each variable was averaged.

Analytic strategy and statistical analysis

Number of situations was analyzed by way of a 3 (competition level) X 3 (situation type: immediate intervention, unsuccessful advantage, successful advantage) factorial ANOVA, as was the hypothesis that players' competition level would influence the prevalence of immediate intervention situations (i.e., number of immediate intervention situations divided by total number of situations). The hypothesis that competition level would influence the application of sporting sanctions in unsuccessful and successful situations was tested using a 3 (competition level) X 2 (successful vs. unsuccessful advantage situations) X 3 (sanction type: 9-meter throw vs. 7-meter throw vs. no punishment) factorial ANOVA.

Concerning disciplinary decisions, "failure situations" were defined following Souchon et al. (2010) as immediate intervention situations pooled with unsuccessful advantage situations. The hypothesis that referees' disciplinary sanctions in successful and failure situations would be influenced by competition level was tested using a 2 (success vs. failure) X 3 (no sanction vs. yellow card vs. 2-minute suspension) X 3 (competition level) factorial ANOVA. No red cards were observed.

Content analysis of referees' explanations. The fresh sample of 100 referees were given four questions based on the findings of the analysis of decisions. Because the questions were developed in view of these findings, they are described in the Results section below. Testing was carried out by groups during four meetings between referees in 2004. After data collection, which took approximately 25 minutes, participants were asked to place the questionnaire in a box to guarantee anonymity.

Results

Main experiment

Number of situations. Results indicated that main effects of competition level, $F(2, 261) = 52.92, p < .001, \eta_p^2 = .29$, type of situation, $F(2, 261) = 150.5, p < .001, \eta_p^2 = .54$, and the 2-way interaction between competition level and type of situation, $F(4, 261) = 25.29, p < .001, \eta_p^2 = .28$, were significant.

Table 1 describes the 2-way interaction. Players, whatever their competition level, were as frequently involved in immediate intervention situations (p 's $> .05$) and unsuccessful situations (p 's $> .05$). Nevertheless, highest-level players were more frequently involved in successful situations than mid-level players, who were in turn more frequently involved in these situations than lowest-level players (p 's $< .001$).

Proportion of Different Situations. A significant main effect of the type of situation, $F(2, 261) = 239.83, p < .001, \eta_p^2 = .64$, and a two-way interaction between competition level and type of situation, $F(2, 261) = 25.34, p < .001, \eta_p^2 = .28$, were significant.

The 2-way interaction indicated that players were marginally more frequently involved in immediate intervention situations in the lowest ($M = .30, SD = .08$) than in the intermediate level ($M = .27, SD = .05$), $p = .09$, who were in turn more frequently involved in these type of situations than highest-level players ($M = .22, SD = .04$), $p < .01$. Also, lowest-level players ($M = .30, SD = .06$) were more involved in unsuccessful situations than highest-level players

($M = .26$, $SD = .05$), $p < .01$, but players at the intermediate level ($M = .30$, $SD = .06$) were as frequently involved in these types of situations as lowest and highest level players ($ps > .05$). Furthermore, players were more frequently involved in successful situations at the highest ($M = .51$, $SD = .06$) than at the intermediate level ($M = .44$, $SD = .06$), who were in turn more frequently involved in this type of situation than lowest-levels players ($M = .38$, $SD = .07$), p 's $< .01$.

Sporting Decisions. The main effect of severity of sanction, $F(2, 522) = 3,056.19$, $p < .001$, $\eta_p^2 = .92$, the 2-way interaction between type of situation and severity of sanction, $F(2, 522) = 2,226.62$, $p < .001$, $\eta_p^2 = .89$, and the 2-way interaction between competition level and severity of sanction, $F(4, 522) = 34.56$, $p < .001$, $\eta_p^2 = .21$, were all qualified by a 3-way interaction between competition level, type of situation, and severity of sanction, $F(4, 522) = 21.23$, $p < .001$, $\eta_p^2 = .14$.

The two way interaction between competition level and severity of sanction indicates that players were less frequently sanctioned with a 9-meter throw at the highest level ($M = .23$, $SD = .19$) than at the intermediate level ($M = .30$, $SD = .28$, $p < .01$), who were in turn less frequently sanctioned with a 9-meter throw than lowest-level players ($M = .34$, $SD = .29$, $p < .01$). Also, referees decided not to intervene more frequently at the highest level ($M = .71$, $SD = .25$) than at the intermediate level ($M = .65$, $SD = .31$, $p < .01$) and more frequently at the intermediate level than at the lowest level ($M = .61$, $SD = .33$, $p < .01$). Nevertheless, referees intervened with a 7 meter throw with similar frequency across all competition levels.

Table 2 describes the three-way interaction between competition level, type of situation, and severity of sanction. This table indicates that the interaction between competition level and severity of sanction described above in turn emerged only for unsuccessful situations: no competition level differences appear for any sanction in successful situations. For unsuccessful situations, referees were also more likely to let the game continue

without intervention with highest-level than with intermediate-level players ($p < .001$), and with intermediate players more than with lowest-level players ($p < .001$).

Disciplinary Decisions. The main effect of severity of sanction, $F(2, 522) = 139,324.6$, $p < .001$, $\eta_p^2 = .99$, the 2-way interaction between type of situation and severity of sanction, $F(2, 522) = 143.9$, $p < .001$, $\eta_p^2 = .35$, and the 2-way interaction between competition level and severity of sanction, $F(4, 522) = 43.1$, $p < .001$, $\eta_p^2 = .25$, were all qualified by a 3-way interaction between competition level, type of situation, and severity of sanction, $F(4, 522) = 11.4$, $p < .001$, $\eta_p^2 = .08$.

The 2-way interaction between competition level and severity of sanction indicates that referees gave more yellow cards to highest-level players ($M = .027$, $SD = .01$) than to intermediate-level players ($M = .019$, $SD = .01$, $p < .04$), who in turn received more yellow cards than lowest-level players ($M = .01$, $SD = .01$, $p < .01$). In terms of 2-min suspensions, while referees punished highest-level players ($M = .026$, $SD = .01$) and intermediate-level players ($M = .025$, $SD = .01$) more severely than lowest-level players ($M = .01$, $SD = .01$), p 's $< .01$, referees punished highest-level players and intermediate-level players equally severely, $p > .05$. Overall, referees decided not to intervene with disciplinary punishment more frequently with lowest-level players ($M = .98$, $SD = .02$) than with intermediate-level players ($M = .956$, $SD = .02$, $p < .01$), who were in turn less frequently punished than highest-level players ($M = .947$, $SD = .02$, $p < .01$).

Table 3 describes the 3-way interaction between competition level, type of situation, and severity of sanction. This table indicates that the interaction between competition level and severity of punishment described above in turn only emerged in failure/unsuccessful situations. More generally, no competition level effects emerged in successful situations.

Content analysis of referees' explanations for competition-level effects on decisions.

The findings described above were then used as a basis for questions posed to the 100

referees who participated in the second part of the study. Specifically, they were asked (1) why players display more transgressions toward attacking players in possession when the competition level rises, while referees intervene less frequently with sporting decisions; (2) why referees punish players more severely with disciplinary punishment at the highest than at the lowest levels; (3) why the proportion of immediate intervention situations is greater in the lowest than in the highest level in reference to the number of defensive transgression committed; and (4) why when a player misses their pass or their shot after being victim of a defensive transgression, referees tend to return back the ball more frequently to lowest-level than to intermediate-level players, and to intermediate-level players more than to highest-level players.

Responses were content analyzed. The first and the fifth author of the paper identified individual units in different inductively-created categories (Krippendorff, 2013). After agreement and corrections with another researcher (the second author of the paper), these individual units were organized into 27 different categories. The rate of agreement between the two main observers was 92.4% and the rate of agreement within the first coder was 96.1%. Participants broadly used 6 different types of explanations: (a) performance (9 sub-categories), (b) aggressiveness (5 sub-categories), (c) subjective explanations (2 sub-categories), (d) refereeing characteristic and skills (4 sub-categories), (e) contextual explanation (3 sub-categories), (f) others (4 sub-categories). Table 4 depicts the percentage of referees who stated at least one of the most common explanations (i.e., performance, aggressiveness, subjective explanation and refereeing characteristic).

Results revealed that referees frequently expressed the stereotype that highest-level players would be more skillful and would perform better than lowest-level players. Also, referees tended to expect that highest-level players would be more aggressive than lowest-level players. Interestingly, more than half of referees suggested that they would be more

benevolent (i.e., more readily return the ball to the victim of a foul after a missed pass or shot) with low- than high-level players. Thirty-nine percent of referees also mentioned the influence of refereeing skill in order to understand the competition-level effects. Notably, this was evoked as an explanation much less than were explanations based on player competence

Discussion

The aims of the present research were to test the effect of competition level on referees' decisions in handball, and to examine the extent to which referees invoke stereotypic beliefs about player competence in order to explain these effects. We expected that for both sporting and disciplinary sanctions, referees would intervene less with higher-level players than with lower-level players. Results were consistent with this prediction for sporting sanctions, but not for disciplinary sanctions. Referees tended to apply fewer sporting sanctions to highest-level players than lowest-level players, but punished highest-level players *more* severely with disciplinary sanctions than lowest-level players. In turn, competition level effects as expected occurred only in immediate intervention situations and unsuccessful situations, but not in successful situations. Overall, these findings greatly extend prior evidence that competition level may influence referees' decisions (Souchon et al., 2009) and previous studies on judgmental heuristics in refereeing (e.g., Unkelbach & Memmert, 2010).

Concerning sporting sanctions, referees as predicted intervened immediately more frequently with lowest- than with highest-level players, and were more likely in unsuccessful situations to award a 9-meter throw to lowest-level players than to intermediate-level players, and returned the ball more frequently to intermediate-level players than to highest-level players (see Souchon et al., 2010 for similar effects related to player gender). Different processes may explain these effects. For example, the high speed of play at a high level of competition may make a higher proportion of fouls more ambiguous (e.g., MacMahon, Starkes, & Deakin, 2007), while higher-player aggressiveness overall may make it more

difficult to notice each transgression (Mascarenhas, O'Hare, & Plessner, 2006). Referees may also consciously or unconsciously adjust their decisions in order to only sanction the defensive transgressions that surpass a certain level of intensity (Unkelbach & Memmert, 2008). In addition, the relatively high level of stress due to pressure from players (e.g., Kaissidis-Rodafinos, Anshel, & Sideridis, 1998) and from supporters (e.g., Nevill et al., 2002) at the highest competition level may impact upon referees' decisions in a manner that maintains the 'flow' of the game, and its value as a spectacle (e.g., Mascarenhas et al., 2006).

Nevertheless, the fewer sporting sanctions at the highest levels of play are consistent both with the idea that referees may wait longer before intervening at a high level of competition because they could expect that players can prolong their actions, despite the gravity of the fouls (Souchon et al., 2009), and with predictions derived from the SCM (Fiske et al., 2002). Specifically, analysis of referees' explanations for the effects of competition level described above revealed that the most common forms of explanation invoked were related to beliefs about player competence (e.g., the inability of lower-level players to gain from playing an advantage; the ability of higher-level players to resist fouls), the greater aggression of higher-level players (i.e., low warmth), and the need for benevolent intervention at lower levels. The SCM would predict more benevolence towards members of low-status groups (low-level players in this case), on the basis that they are less competent but warmer (e.g., less aggressive), than their high-status (high playing level) counterparts (Cuddy et al., 2008). These were all invoked more frequently than beliefs about referees' own ability at different competition levels. As characterizations of player competence, they thus logically serve to *justify* as well as explain the competition-level effects (Hoffman & Hurst, 1990), in keeping with theories of the social functions of stereotypic beliefs (e.g., Tajfel, 1981).

Concerning disciplinary sanctions, the results contradicted our predictions and the findings of Souchon et al. (2009), but were consistent with results obtained by Debanne and

Fontayne (in press). Referees in “failure situations” tended to punish higher-level players more severely than lower-level players. Specifically, the greater intensity of contact between attacking and defensive players at this level could justify more severe punishment of the defensive player with a disciplinary sanction. Analysis of referees’ explanations revealed that this may be subjectively explained in terms of the greater risk posed to attacking players at this level. Nevertheless, referees may react more leniently to lower-levels players as they would perceive them to be less competent and warmer than higher-levels players (Fiske et al., 2002).

In terms of the wider applicability of the present findings, models such as the SCM have predictions relating directly to behavior, but these have generally not been directly tested, with research tending to focus on perceptions, attitudes and emotions as outcomes. In contrast, the present research directly assesses the naturally-occurring behavior of a powerful group who are directly adjudicating upon, and intervening in, the activities of others. In turn, the study reveals important nuances in terms of how perceived competence and warmth may shape such behavior. Specifically, the finding that referees apply fewer sporting sanctions, but more disciplinary sanctions to relatively high-level players suggests that rather than being a simple matter of more or less intervention/punishment *per se*, actions aimed at regulating the behavior of outgroups may be taken more frequently, but with less severity for low competence/high warmth groups, whereas actions aimed at regulating the behavior of outgroups may be taken less frequently, but with more severity for high competence/low warmth groups.

Limitations and future research

Limitations of the present research include the fact that different referees officiate at different competition levels, while officiating alone or as pairs depending on competition levels. Also, because of the naturalistic observations, we were unable to test directly the effect

of stereotyping (or other processes) on referees' decision making. For example, an alternative explanation of these results may be referees' level of expertise, as suggested by referees in the second part of the study. Ste-Marie (2003) suggests, for example, that expert referees show better eye movement patterns to identify essential sources of information, and best predict the outcome of visually-presented information. Greater refereeing expertise at the highest level could thus facilitate the perception of dangerous transgression situations and if the advantage rule could be applied (MacMahon, Starkes, & Deakin, 2007). Addressing these limitations may require experimental research in different team contact sports, potentially using virtual reality technology. A tentative prediction based on previous work on referees' gender stereotypes (Souchon et al., 2013) would be that referees at all levels level of expertise – even the most experienced – would be influenced by their stereotype related to competition level.

Future research could also examine how referees' decisions may be related to referees' implicit cognition, for example by using an implicit association task (Greenwald, McGhee, & Schwartz, 1998) to explore the role of implicit and explicit competition level stereotypes and the role of implicit and explicit attitudes toward higher vs. lower-level players. Implicit and explicit measures are better predictors of behaviors together than in isolation, especially when the correlation between implicit and explicit measures is high (see Greenwald, Poehlman, Uhlmann, & Banaji, 2009), as we would expect to be the case here.

References

- Betsch, T., Haberstroh, S. (2005). *The routines of decision making*. Hillsdale, NJ: Lawrence Erlbaum.
- Coulomb-Cabagno, G., Rascle, O., & Souchon, N. (2005). Players gender and male referees' decisions about aggression in french soccer: a preliminary study. *Sex Roles*, 52(7-8), 547-553.
- Conroy, D. E., Silva, J. M., Newcomer, R. R., Walker, B. W., & Johnson, M. S. (2001). Personal and participatory socializers of the perceived legitimacy of aggressive behavior in sport. *Aggressive Behavior*, 27, 405-418.
- Cuddy, A.J.C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, 40, 61-149.
- Debanne, T. & Fontayne, P. (in press). Les prises de décisions des arbitres d'un point de vue des théories de la domination. Approche psycho-sociale dans l'activité handball. *Psychologie Française*.
- Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labelling effects. *Journal of Personality and Social Psychology*, 44, 20-33
- Fiske, T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status competition. *Journal of Personality and Social Psychology*, 82, 879-902.
- 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.

- Greenwald, A. G., McGhee, D. E., & Schwartz, J. K., L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74, 1464-1480.
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97, 17-41.
- Hoffman, C., & Hurst, N. (1990). Gender stereotypes: Perception or rationalization? *Journal of Personality and Social Psychology*, 58, 197–208.
- International Handball Federation. (2005). Rules of the game. Båle: International Handball Federation.
- 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.
- Kaissidis-Rodafinos, A., Anshel, M., & Sideridis, G. (1998). Sources, intensity, and responses to stress in Greek and Australian basketball referees. *International Journal of Sport Psychology*, 29, 1-22.
- Krippendorff, K. (2013). Content analysis: An introduction to its methodology (3rd ed.). London: Sage.
- MacMahon, C., Starkes, J. & Deakin, J. (2007). Referee decision making in a video-based infraction detection task: application and training considerations. *International Journal of Sports Science and Coaching*, 2, 257-265.
- Mascarenhas, D., O'Hare, D., & Plessner, H. (2006). The psychological and performance demands of Association Football refereeing. *International Journal of Sport Psychology*, 37, 99-120.

- Nevill, A. M., Balmer, N. J., & Williams, A. M. (2002). The influence of crowd noise and experience upon refereeing decisions in football. *Psychology of Sport and Exercise*, 3, 261-272.
- Plessner, H., & Haar, T. (2006). Sports performance judgments from a social cognitive perspective. *Psychology of Sport and Exercise*, 7, 555-575.
- Schneider, D. J. (2004). The psychology of stereotyping. New York : The Guilford Press.
- Souchon, N., Cabagno, G., Traclet, A., Dosseville, F., Livingstone, A. G., Jones, M., & Maio, G.R. (2010). Referees' decision making and player gender: The moderating role of the type of situation. *Journal of Applied Sport Psychology*, 22, 1, 1-16.
- Souchon, N., Cabagno, G., Traclet, A., Trouilloud, D., & Maio, G. R. (2009). Referees' use of heuristics: The moderating impact of level of competition. *Journal of Sport Sciences*, 27, 695-700.
- Souchon, N., Coulomb-Cabagno, G., Traclet, A. & Rascle, O. (2004). Referees' decision making in handball and transgressive behaviors: influence of stereotypes about gender of players ? *Sex Roles*, 51 (7/8), 143-150.
- Souchon, N., Livingstone, A. G., & Maio, G. R. (2013). The influence of referees' expertise, gender, motivation, and time constraints on decisional bias against women. *Journal of Sport and Exercise Psychology*, 35(6), 585-599.
- Ste-Marie, D.M. (2003). Expertise in sport judges and referees: Circumventing information processing limitations. In J.L. Starkes & A. Ericsson (Eds.), *Expert performance in sport: Advances in research on sport expertise* (pp. 169–190). Champaign, IL: Human Kinetics.
- Tajfel, H. (1981). Human groups and social categories. New York: Cambridge University Press.
- Unkelbach, C., & Memmert, D. (2008). Game management, context effects, and calibration: the case of yellow cards in soccer. *Journal of Sport & Exercise Psychology*, 30, 95-109.

Unkelbach, C., & Memmert, D. (2010). Crowd noise as a cue in referee decisions contributes to the home advantage. *Journal of Sport & Exercise Psychology*, 32, 483-498.

Table 1

Means (and standard deviations) concerning the number of transgressions (TR), multiple transgressions (X2) and situations (SIT) in the game, depending on competition level and type of situations (Immediate, Successful, Unsuccessful)

	Local			Intermediate			National		
	I	S	U	I	S	U	I	S	U
SIT	32.87 (10.98)	41.76 (13.04)	31.8 (7.30)	36.7 (12.01)	59.37 (15.63)	36.8 (8.37)	35.15 (8.05)	83.06 (16.38)	41.06 (6.98)
X2	.53 (.73)	.06 (.25)	.30 (.65)	3.1 (2.83)	.53 (.86)	.60 (.77)	4.15 (2.29)	1.7 (1.46)	2.66 (2.05)
TR	33.40 (11.11)	41.82 (13.05)	32.1 (7.22)	39.80 (13.31)	59.9 (15.95)	37.4 (8.59)	39.3 (9.65)	84.76 (16.87)	43.72 (7.01)

Table 2

Means (and standard deviations) concerning sporting sanctions for unsuccessful situations and successful situations according to competition level and the severity of sanction

	Unsuccessful advantage			Successful advantage		
	Local	Inter	Nat	Local	Inter	Nat
9-meter throw	0.63** (0.13)	0.57* (0.12)	0.41* (0.09)	0.07 (0.05)	0.04 (0.03)	0.04 (0.02)
7-meter throw	0.07 (0.08)	0.08 (0.05)	0.11 (0.05)	0 (0.00)	0 (0.00)	0 (0.00)
No sporting intervention	0.30** (0.14)	0.35* (0.11)	0.48* (0.08)	0.93 (0.05)	0.96 (0.03)	0.96 (0.02)

*Inter-category R/D or N/R difference is significant $p < .05$

** Inter-category D/N difference is significant $p < .05$

Table 3

Means (and standard deviations) concerning disciplinary punishment according to competition level for failure situations and successful situations

	Immediate and Unsuccessful			Successful situations		
	Local	Inter	Nat	Local	Inter	Nat
Yellow Card	0.01** (0.01)	0.03* (0.01)	0.04* (0.01)	0.005 (0.01)	0.01 (0.01)	0.01 (0.01)
2-min suspension	0.01** (0.01)	0.04* (0.01)	0.04 (0.2)	0.005 (0.01)	0.01 (0.01)	0.01 (0.01)
No punishment	0.98** (0.02)	0.93* (0.01)	0.92 (0.02)	0.99 (0.01)	0.98 (0.02)	0.98 (0.02)

*Inter-category R/D or N/R difference is significant $p < .05$

** Inter-category D/N difference is significant $p < .05$

Table 4

Percentage (number of units) of referees who stated at least one stereotypical explanation related to competition level

	Regional	National	Together
Performance			
High level players are more competent			
Attacking player performance			
<i>Resistance and ball liberation</i>	54% (32)	64%(50)	59% (82)
<i>Physical abilities</i>	30% (17)	52% (44)	41% (61)
<i>Technical abilities</i>	52% (49)	60% (62)	56% (111)
<i>Clear chances of scoring creation</i>	22% (11)	40% (21)	31% (32)
<i>Speed of the game/ creation free space</i>	12% (6)	12% (8)	12% (14)
<i>Combination (at least one of the above)</i>	92% (115)	94% (185)	93% (300)
Collective performance			
<i>Better collective game</i>	14% (8)	34% (24)	24% (32)
Other			
<i>High level players are better</i>	26% (15)	20% (11)	23% (26)
<i>Training</i>	26% (15)	28% (24)	27% (39)
<i>Rules of the game understanding</i>	10% (5)	6% (3)	6.5% (8)
Combination	98% (158)	98% (247)	98% (405)
Aggressiveness			
High level players are more aggressive			
Attacking player aggressiveness			
<i>Better involvement and risk taken</i>	30% (17)	28% (15)	29% (32)
<i>Continue their attack despite the defensive foul</i>	4% (3)	4% (2)	4% (5)
Defensive player aggressiveness			
<i>High level players are more aggressive</i>	18% (10)	16% (8)	17% (18)
In general			
<i>Game is harsher at high level</i>	24% (12)	16% (9)	20% (21)
Combination	54% (42)	58% (34)	56% (76)
Low level player are more aggressive	10% (5)	10% (5)	10% (10)
Subjective explanation			
More benevolence with low level players	52% (31)	52% (32)	52% (63)
Contacts are less dangerous at high level	10% (6)	6% (3)	8% (9)
Combination	56% (37)	56% (35)	56% (72)
Refereeing characteristic and skills			
More skilled at high level	34% (26)	34% (20)	34% (46)
More pedagogy at low level	4% (3)	4% (2)	4% (5)
Orders from institution	4% (2)	0% (0)	2% (2)
<i>Favor the show at high level</i>	4% (3)	12% (6)	8% (9)
Combination	38% (34)	40% (28)	39% (62)