
RESPONSE & REBUTTAL

JOURNAL OF SPORT & EXERCISE PSYCHOLOGY, 2002, 24, 328-333

© 2002 Human Kinetics Publishers, Inc.

The Art of Reason Versus the Exactness of Science in Elite Refereeing: Comments on Plessner and Betsch (2001)

Duncan R.D. Mascarenhas, Dave Collins, and Patrick Mortimer
University of Edinburgh

Plessner and Betsch's (2001) investigation into officiating behavior may be representative of a shift from stress-oriented research (Anshel & Weinberg, 1995; Rainey & Winterich, 1995; Stewart & Ellery, 1996) to consideration of decision-making (Craven, 1998; Ford, Gallagher, Lacy, et al., 1999; Oudejans, Verheijen, Bakker, et al., 2000), the primary function of referees in any sport. Commendably, Plessner and Betsch have investigated the most important focus of referee performance, the application of the rules (Anshel, 1995). However, methodological weaknesses, together with a fundamental error in the attribution of causation to the findings, significantly dilute the paper's contribution to extending knowledge in this important area.

Key Words: judgment biases, decision-making, soccer

As sport becomes more pressured, referees are ever more accountable for their decisions and, reflecting this performance demand, governing bodies now sanction, or even demote, referees when they apply the laws inconsistently (*The Guardian*, 2001). The performance emphasis, both for the referees themselves and for the sports, makes this an area worthy of investigation by sport scientists. Therefore, Plessner and Betsch (2001) should be commended for their research into referee decision-making in soccer.

Plessner and Betsch studied the decisions made by 58 referees and 57 players on 20 videotaped scenes taken from a Spanish Primera Division soccer match. The use of videotaped scenarios to study this area is laudable, since video provides detailed presentations of events in real time, in which content can be manipulated (Cannon-Bowers & Bell, 1997). It allows different participants to independently consider an event that holds exactly the same information so that crucial comparisons can be made, and the paper represents an excellent step in examining this important domain. However, the paper also reflects a methodological flaw, a design weakness and a questionable interpretation which, singularly and in combination, serve to limit the paper's contribution to this important area.

The authors are with the Dept. of Physical Education, Sport and Leisure Studies, University of Edinburgh, St. Leonard's Land, Edinburgh, EH8 8AQ Scotland.

First, the methodological flaw. Plessner and Betsch (2001) apparently failed to match the testing perspective to the referee's usual view, employing a spectator's perspective as opposed to the "in the thick of it" angle that officials experience. Accordingly, the study has forfeited ecological validity, which may in turn have nullified the benefits of the referees' expertise (Abernethy, Thomas, & Thomas, 1993; Chamberlain & Coelho 1993; Williams & Davids, 1995). This contention is also well supported by the literature on naturalistic decision making, which suggests the need for a familiar perspective since an alternative angle may differ in important detail, distorting both the complexity and dynamics of a situation (Omodei, Wearing, & McLennan, 1997; Stokes, Kemper, & Kite, 1997).

A recent study of elite rugby-union referees and their coaches also highlighted the importance of the viewing angle. Here the referee coaches, a group who are used to seeing games from the sidelines, were up to 10% less accurate at making correct decisions when presented with the referees' perspective (Mascarenhas, Collins, & Mortimer, 1999). In short, different perspectives may generate different decisions, a situation that is exacerbated when the participants are asked to execute decision-making tasks from an unfamiliar perspective.

Interestingly, Plessner and Betsch's results seem to support this contention. They reveal that, in the game from which their test footage was culled, the original referee only gave a penalty in the second situation, although on video this offense was less clear than the first. It is perhaps not surprising that only 18% (13 out of 72) of the participants made the same decision as the original referee in both situations. The sideline perspective may well have distorted important information, and thus eradicated or limited the application of the referees' expertise. Unfortunately, since the reasoning that underpinned the participants' decisions was not tested, it is difficult to draw any firm conclusions.

Our second concern, the design flaw, also relates to this factor of expertise. By combining referees and players, Plessner and Betsch compare participants who surely must be unlike in their capacity to make *refereeing* decisions. In fact, the similarity in performance of these two "subsamples" on the tests implies that the design is flawed. It seems obvious that referees should be better at decision-making on rule application than players (whatever players might think!), since that is their primary task. This contention receives support from the work of Williams and Davids (1995), who showed increased expertise associated with actively doing rather than just passively watching soccer. Players don't regularly make rule decisions, whereas referees do. Future studies on referee decision-making need to use referees, specifically those with sufficient practice and expertise at their craft, if useful data are to be accrued concerning subsequent enhancement of performance.

Finally, to the questionable interpretation. One's consideration of reasons that underpin decisions provides a sensitive measure of knowledge (Abraham & Collins, 1998) and reveals the individual's level of understanding (Klinger, 2000; Langan-Fox, Code, & Langfield-Smith, 2000). Unfortunately, while Plessner and Betsch ably demonstrated that the probability of awarding a penalty increased if no penalty had been awarded to the same team before, there is no evidence to show the nature of the association between these two decisions since they didn't ask *why* these decisions were awarded. As such, the results *may* represent the game-

dysfunctional but perhaps understandable socio-psychological phenomenon of making contingencies to "balance things out," which is deemed illegal by FIFA rules. If so, as asserted by Plessner and Betsch, this presents a problem. However, without an analysis of the rationale behind decisions, it is equally plausible to draw other conclusions; in fact, these conclusions fit well with the functions that referees are *required* to fulfill.

Referees are asked to adjudicate games by exercising the rules and accounting for the spirit of the rules (Askins, 2001; Grunski, 2001). Crucially, Grunski (2001) also suggests that referees are increasingly encouraged to establish a "feel for the game," where the aim is to be realistic rather than literal in applying the rules. This is termed game management and represents the art rather than the science of refereeing which is essential in the real world. Thus, expert referees make decisions that are appropriate for the nuances of a particular game, allowing the game to flow, using the whistle only when the consequences for not doing so may adversely affect the tempo or temper of the game. Such "judicious use" of the whistle is a rational, game-management focused act, which is not only praiseworthy but also an essential skill for an elite referee (Bunting, 1999; Wilson, 2000; Colin High, Rugby Football Union elite referee development manager, personal communication, Oct. 2, 2001). Indeed in elite rugby, anecdotal evidence from rugby referees and their assessors, as well as RFU referee technical literature (Bunting, 1999), sees game management as the most important aspect of a referee's skill set.

Therefore, a more positive attribution to Plessner and Betsch's findings would be that the participants felt the game intensity had increased and the players were becoming more physical. So, in order to restrain this escalation of physical contact and prevent further discrepancies, they would logically, and within the letter and spirit of their prescribed duties, decide to award the penalty. Whatever the case, without a thorough analysis of the participants' *reasons*, it is unclear whether decisions were due to a contrast effect as suggested by Plessner and Betsch, or to some other more game-management focused rationale.

Discussion

Reinforcing the need to analyze the reasons underpinning decisions, and in light of our second comment presented earlier, it is worth noting that Plessner and Betsch asserted there were no differences between the referees and the players. We would suggest that investigating the rationale behind decisions would significantly delineate the referees from the players (cf. Rutt-Leas & Chi, 1993). In this regard, and further to the previous paragraph, research reveals that referees take into consideration things such as the previous levels of contact in the game and the potential consequences of their decisions (Mascarenhas, Collins, & Mortimer, 2002). This level of understanding seems to be a function of the referees' expertise at making these types of decisions, an attribute that the players only encounter as a third party.

Similarly, Klein (1993) would argue that by failing to match the time pressure that exists in the real-world environment, namely refereeing out on the field of play, you are not testing their real ability to make refereeing decisions. As such, it is perhaps predictable that the players are equally as skilled (or in this case,

perhaps unskilled) in their decision-making as the referees. Since both groups have essentially been tested as spectators, without time pressure and a familiar perspective, neither have been examined in their own specific areas of experience. In this regard, Cohen (1993) argues that examining decision biases with a naturalistic perspective opens to question certain basic assumptions of the decision-bias research. Cohen suggests that the naturalistic domain helps to answer "what is the actual impact of each bias in real-world domains?" and questions whether the dynamic quality of real tasks, as opposed to the Plessner and Betsch lab task, helps reduce the effects of biases.

Furthermore, Smith and Marshall (1997) suggest that decision-making, in the absence of complete knowledge of the complexity of situations, often requires some form of approximate reasoning. Thus decision-makers are forced to either guess or use higher order judgments such as their philosophy. In fact, Plessner and Betsch acknowledge that participants may "shift their criterion for awarding a penalty to the same team to a higher level" (p. 258). Thus they may accurately attribute these judgments to a "fair play" type of bias that is likely to form part of the participants' philosophy, but which could be very different from their action in a real-world setting.

Just to complicate matters, however, research into the judgments of air defense personnel (Adelman, Bresnick, Black, Marvin, & Sak, 1996; Adelman, Tolcott, & Bresnick, 1993) predicted and showed a more pronounced order bias in naturalistic tests when compared to pencil and paper tests. Thus, given an analysis of the reasons that underpin decisions, to test the extent of such bias in this study, it may actually prove to be even more prevalent than Plessner and Betsch propounded. Clearly, however, without examining the reasons that underpin decisions, their results remain speculative.

Conclusion

Currently there is no benchmark in refereeing research since it is such a new area and, as such, Plessner and Betsch should be applauded for their scientifically rigorous and "tight" investigation. However, this study of human judgment bias presents a lab-based test that unfortunately may be too sterile and too contrived to generalize to the real-world refereeing environment. If the results are to be applied with any confidence, realistic conditions, as offered under the banner of naturalistic decision making, provides the most appropriate paradigm (Cannon-Bowers, Salas, & Pruitt, 1996). In the future, research on referee performance should consider a naturalistic environment that demands an ecologically valid perspective, accounts for the extreme time pressures in which referees must make judgments, and looks at how experts make decisions (Klein, 1997; Orasanu & Connolly, 1993; Zsombok, 1997) through consideration of their reasoning.

References

- Abernethy, B., Thomas, K.T., & Thomas, J.T. (1993). Strategies for improving understanding of motor expertise (or mistakes we have made and things we have learned). In J.J. Starkes & F. Allard (Eds.), *Cognitive issues in motor expertise* (pp. 317-356). Amsterdam: Elsevier Science Publ. B.V.

- Abraham, A., & Collins, D.J. (1998). Examining and extending research in coach development. *Quest*, **50**, 59-79.
- Adelman, L., Bresnick, T., Black, P.K., Marvin, F.F., & Sak, S.G. (1996). Research with Patriot Air Defense officers: Examining information order effects. *Human Factors*, **38**, 250-261.
- Adelman, L., Tolcott, M.A., & Bresnick, T.A. (1993). Examining the effect of information order on expert judgment. *Organizational Behavior and Human Decision Processes*, **56**, 348-369.
- Anshel, M.H. (1995). Development of a rating scale for determining competence in basketball referees: Implications for sport psychology. *The Sport Psychologist*, **9**, 4-28.
- Anshel, M.H., & Weinberg, R.S. (1995). Sources of acute stress in American and Australian basketball referees. *Journal of Applied Sport Psychology*, **7**, 11-22.
- Askins, R. (2001). Common myths about officiating. *Referee*, **26**(10), 44-47.
- Bunting, N.J. (1999, January). *Allied Dunbar premiership game management instructions for referees, players and coaches*. Castlecroft, Wolverhampton: RFU Referees Technical Report from conference on the game.
- Cannon-Bowers, J.A., & Bell, H.H. (1997). Training decision-makers for complex environments: Implications of the naturalistic decision-making perspective. In C.E. Zsombok & G.A. Klein (Eds.), *Naturalistic decision making* (pp. 99-110). Mahwah, NJ: Erlbaum.
- Cannon-Bowers, J.A., Salas, E., & Pruitt, J.S. (1996). Establishing the boundaries of a paradigm for decision making research. *Human Factors*, **38**, 193-205.
- Chamberlain, C.J., & Coelho, A.J. (1993). The perceptual side of action: Decision making in sport. In J.J. Starkes & F. Allard, (Eds.), *Cognitive issues in motor expertise* (pp. 135-157). Amsterdam: Elsevier Science Publ. B.V.
- Cohen, M.S. (1993). Three paradigms for viewing decision biases. In G.A. Klein, J. Orasanu, R. Calderwood, & C.E. Zsombok, (Eds.), *Decision making in action: Models and methods* (pp. 36-50). Norwood, NJ: Ablex.
- Craven, B.J. (1998). A psychophysical study of leg-before-wicket judgments in cricket. *British Journal of Psychology*, **89**, 555-578.
- Ford, G.G., Gallagher, S.H., Lacy, B.A., Bridwell, A.M., & Goodwin, F. (1999). Repositioning the home plate umpire to provide enhanced perceptual cues and more accurate ball-strike judgments. *Journal of Sport Behavior*, **22**, 28-44.
- Grunska, J. (2001). Yeah, you're good...but you can be better. *Referee*, **26**(6), 36-40.
- Klein, G. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G.A. Klein, J. Orasanu, R. Calderwood, & C.E. Zsombok (Eds.), *Decision making in action: Models and methods* (pp. 138-147). Norwood, NJ: Ablex.
- Klein, G. (1997). An overview of naturalistic decision making applications. In C.E. Zsombok & G.A. Klein (Eds.), *Naturalistic decision making* (pp. 49-59). Mahwah, NJ: Erlbaum.
- Klinger, D. (2000). *Observing teams in natural settings: A helpful tool to aid in team observation and feedback*. Paper presented at 5th Annual Conference on Naturalistic Decision-Making, Stockholm.
- Langan-Fox, J., Code, S., & Langfield-Smith, K. (2000). Team mental models: Techniques, methods and analytic approaches. *Human Factors*, **42**, 242-271.
- Mascarenhas, D.R.D., Collins, D., & Mortimer, P. (1999). *Shared mental models: A naturalistic approach to training coherent decision making in rugby union referees*. Paper presented at Scottish Sport Science and Medicine Conference, Edinburgh.
- Mascarenhas, D.R.D., Collins, D., & Mortimer, P. (2002). The 'four cornerstones' model of refereeing. *Referee*, **27**(4), 38.

- Omodei, M., Wearing, A., & McLennan, J. (1997). Head-mounted video recording: A methodology for studying naturalistic decision making. In R. Flin, E. Salas, M. Strub, & L. Marting (Eds.), *Decision making under stress: Emerging themes and applications* (pp. 137-146). Burlington, VT: Ashgate.
- Orasanu, J., & Connolly, T. (1993). The reinvention of decision making. In G.A. Klein, J. Orasanu, R. Calderwood, & C.E. Zsombok (Eds.), *Decision making in action: Models and methods* (pp. 3-20). Norwood, NJ: Ablex.
- Oudejans, R.R.D., Verheijen, R., Bakker, F.C., Gerrits, J.C., Steinbruckner, M., & Beek, P.J. (2000). Errors in judging 'offside' in football. *Nature*, **404**, 33.
- 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.
- Rainey, D.W., & Winterich, D. (1995). Magnitude of stress reported by basketball referees. *Perceptual and Motor Skills*, **81**, 1241-1242.
- The Guardian*. (2001, Oct. 29). Referee Gallagher relegated after Keane blunder. p. 15.
- Rutt-Leas, R., & Chi, M.T.H. (1993). Analyzing diagnostic expertise of competitive swimming coaches. In J.J. Starkes & F. Allard (Eds.), *Cognitive issues in motor expertise* (pp. 75-94). Amsterdam: Elsevier Science Publ. B.V.
- Smith, D.E., & Marshall, S. (1997). Applying hybrid models of cognition in decision aids. In C.E. Zsombok & G.A. Klein, (Eds.), *Naturalistic decision making* (pp. 331-341). Mahwah, NJ: Erlbaum.
- Stewart, M.J., & Ellery, P.J. (1996). Amount of psychological stress reported by high school volleyball officials. *Perceptual and Motor Skills*, **83**, 337-338.
- Stokes, A.F., Kemper, K., & Kite, K. (1997). Aeronautical decision making, cue recognition, and expertise under time pressure. In C.E. Zsombok & G.A. Klein (Eds.), *Naturalistic decision making* (pp. 183-196). Mahwah, NJ: Erlbaum.
- Williams, M., & Davids, K. (1995) Declarative knowledge in sport: A by-product of experience or a characteristic of expertise? *Journal of Sport & Exercise Psychology*, **17**, 259-273.
- Wilson, R. (Producer and Director). (2000, July 25). *Man in black*. Scotland: British Broadcasting Corp.
- Zsombok, C.E. (1997). Naturalistic decision making: Where are we now? In C.E. Zsombok & G.A. Klein (Eds.), *Naturalistic decision making* (pp. 3-16). Mahwah, NJ: Erlbaum.

Manuscript submitted: December 5, 2001

Copyright of Journal of Sport & Exercise Psychology is the property of Human Kinetics Publishers, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.