Evidence from baseball that umpiring decisions are influenced by game context

Clare MacMahon

Traditional approaches in sport decision-making research test problems with a 'correct' and agreed upon answer. In many cases, however, the most critical decisions are often the least clear-cut. This is particularly the case for officiating decisions. The approach used in this study was to examine borderline or ambiguous decisions for the ball-strike judgment in baseball.

When an individual batter faces a pitcher and does not hit the ball, the plate umpire must judge the flight path of the ball relative to the strike zone. If a pitcher throws four balls, the batter is 'walked' to first base, with the opportunity eventually to score. If a pitcher throws three strikes, the batter is retired for an 'out'. Three outs end an inning, diminishing the batting team's opportunity for points. The pitch count, which tallies the number of balls and strikes during an individual batter's 'at bat', is thus an important piece of contextual information, showing whether or not the pitcher or batter is ahead, and how critical the situation is.

Baseball pitches were filmed from the umpire's perspective using a helmet-mounted camera. Pitch clips were viewed by three expert baseball umpires who classified them as 'balls' or 'strikes'. This resulted in a pool of definite 'ball' clips, and a pool of definite 'strike' clips where the three experts agreed. Disagreements between the experts were used to identify ambiguous or borderline pitches.

We recruited 31 baseball umpires, 16 baseball players and 27 control participants with little to no baseball experience to complete three tasks. The basic task was to watch a clip and classify it as a ball or a strike. Participants were then asked to rate their certainty on a scale from 1 to 5, where a 1 was definite and 5 was borderline. Post data collection, balls were designated -1 and strikes +1. When the two components of the decision were combined, responses ranged from 'definite ball' at -5 to 'definite strike' at +5.

In a basic judgment task, participants viewed 32 video clips of definite balls and definite strikes. In a 'direct information task', participants made a ball-strike decision for the same 16 borderline pitches presented in four different conditions: 1, following one definite strike; 2, following two definite strikes; 3, following one definite ball; and 4, following two definite balls. All borderline clips were seen in each condition.

In the 'summary information' or 'pitch count task', participants were shown 12 borderline pitches in four different pitch count conditions: 1, a neutral count of 0 balls and 0 strikes; 2, a 'batter-ahead' count of 3 balls and 0 strikes; 3, a neutral but critical count of 3 balls and 2 strikes (where a ball decision leads to a walk to first base and a strike decision leads to an out); and 4, a 'pitcher-ahead' count with 0 balls and 2 strikes.

Umpires and players were significantly better at calling pitches than controls in the basic judgment task, and thus it is a task that taps into experience-based expertise. In the 'direct information task', all participants called target clips closer to the strike end of the scale when viewed after definite balls than when they followed definite strikes. Similarly, in the 'pitch count task', participants called pitches more towards the strike end of the scale when there were three balls in the count (3-0, 3-2), then in the other two pitch counts (0-0, 0-2). A final interesting finding is that, compared to the players and control participants, umpires showed a greater overall tendency to call strikes as opposed to balls.

These findings provide evidence that the standard for evaluation of a baseball pitch changes based on the context. Moreover, previous pitches (ie, definite strikes, definite balls) and the pitch count appear to be strong contextual factors, given that their influence is shown not only in umpires and players but also in participants with little experience with the sport. The results of this study also provide evidence that umpires use a normative rule of 'hardening the game', or moving it along by calling strikes (Larsen & Rainey, 1991).

This work not only shows that context influences decisions but also provides a useful approach to understanding the complexity of critical decisions in a variety of sports. Specifically, these results allude to further work in cricket umpiring to study decisions and potentially design interventions to improve decision making. For example, this method can be used to examine critical difficult decisions such as wides, no balls, outs and LBWs. Contextual factors such as the score, stage in the match and status or ranking of a player and strength of an appeal may all influence decisions. Studying these factors is the next step in this area of research.

Acknowledgments

The author would like to acknowledge the invaluable input of Janet Starkes from McMaster University in Ontario, Canada, and Neil Charness from Florida State University, as well as post-doctoral funding from the Social Sciences and Humanities Research Council of Canada.


Clare MacMahon is at Victoria University, Melbourne, Victoria

Correspondence:
Clare.macmahon@vu.edu.au