

Université du Québec en Outaouais

L'influence des forces sociales sur la prise de décision des arbitres sportifs.

Thèse doctorale
Présentée au
Département de psychoéducation et de psychologie

Comme exigence partielle du doctorat en psychologie,
cheminement recherche (Ph. D.)

Par
© Joël GUÉRETTE

Août 2025

Composition du jury

L'influence des forces sociales sur la prise de décision des arbitres sportifs.

Par
Joël Guérette

Cette thèse doctorale a été évaluée par un jury composé des personnes suivantes :

Daniel Fiset, Ph. D., directeur de recherche, Département de psychoéducation et de psychologie, Université du Québec en Outaouais.

Caroline Blais, Ph. D., codirectrice de recherche, Département de psychoéducation et de psychologie, Université du Québec en Outaouais.

Stéphane Bouchard, Ph. D., examinateur interne, Département de psychoéducation et de psychologie, Université du Québec en Outaouais

Jérémie Verner-Filion, Ph. D., examinateur interne, Département des sciences de l'éducation, Université du Québec en Outaouais.

Patrick Gaudreau, examinateur externe, Faculté des sciences sociales, École de psychologie, Université d'Ottawa

Pour Natasha, Jack, et Charlie qui s'apprête à voir le jour.

REMERCIEMENTS

La réalisation de cette thèse a été rendue possible grâce au soutien inestimable de nombreuses personnes clés ayant contribué, chacune à leur manière, au fruit du travail ici présenté.

Je tiens à exprimer ma plus profonde gratitude à l'endroit de mes directeurs de thèse, Daniel et Caroline, qui ont su faire preuve du parfait équilibre entre le développement de mon autonomie et mon besoin d'encadrement. Vous m'avez permis de me dépasser et d'acquérir les compétences essentielles dans tous les aspects du travail de professeur, tout en me laissant le plein contrôle sur le choix de mes projets et de mes thématiques de recherche. Je ne serais jamais rendu où j'en suis sans vous et je vous en serai éternellement redevable.

Je voudrais exprimer ma sincère reconnaissance à mes parents et aux membres de ma famille pour leur amour, leur soutien, les valeurs qu'ils m'ont transmises et les leçons qu'ils m'ont enseignées, autant par la parole que par l'exemple. Merci de m'avoir appris qu'on n'est jamais mieux servi que par soi-même et que le travail acharné et le respect des autres paient toujours en bout de piste.

Je souhaite transmettre toute mon appréciation envers mes collègues, passés et présents, en provenance du domaine de la kinésiologie comme de la psychologie, du décanat des études, de l'École interdisciplinaire de la santé et du Laboratoire de Perception Visuelle et Sociale, pour leur partage de connaissances, les défis relevés ensemble et les amitiés créées qui vont bien au-delà des relations de travail.

Je tiens également à saluer le support de mes ami.es qui, avec admiration ou humour, me rappellent que de retourner sur les bancs d'école à mon « âge avancé » est courageux et digne de mention. Un merci spécial à Jason Lacelle pour la relecture des articles et les propositions ayant permis de les présenter dans un niveau d'anglais qui dépasse de loin le mien. Des remerciements

tout aussi spéciaux à Jonathan Laurin pour l'aide en programmation, sans laquelle je n'aurais jamais pu mener à terme cette thèse.

Je désire aussi souligner que je suis très touché par la contribution des personnes qui ont permis de faire rayonner mes travaux, que ce soit dans les événements scientifiques, dans les médias traditionnels ou dans leurs conversations.

Je souhaite également saluer l'implication des membres du jury de la présente thèse doctorale, pour le temps investi dans son évaluation et pour les commentaires permettant d'améliorer sa version finale.

Je ne saurais passer sous silence l'impact de toutes les autres personnes qui ont, de près ou de loin, contribué à leur manière à mon parcours académique et personnel.

Finalement, je ne saurais trop remercier monoureuse, Natasha, et mon fils, Jack, pour leur présence et leur support indéfectible. Merci de me rappeler quelles sont mes priorités chaque jour, tout en me permettant de me dépasser dans les autres sphères de ma vie. Je vous aime de tout mon cœur.

RÉSUMÉ

L'environnement sportif est un contexte écologique idéal pour observer les interactions humaines en situation de compétition. Afin d'assurer un climat de confrontation juste et équitable, les différentes ligues sportives font habituellement appel à des arbitres. Ces derniers sont toutefois aux prises avec une tâche extrêmement complexe, celle de rendre des décisions impartiales dans un environnement où l'information à traiter dépasse leurs capacités cognitives. Les arbitres sont ainsi à risque de prendre des décisions incorrectes ou biaisées. Une importante source de biais provient des forces sociales présentes dans l'environnement sportif. La présente thèse doctorale vise à étudier l'impact de certaines de ces forces sur la prise de décision des arbitres sportifs. En premier lieu, l'influence de la foule a été analysée dans le contexte du hockey professionnel. Prenant racine dans les événements inattendus liés à la pandémie de la COVID-19, il a été possible de comparer le nombre de pénalités attribuées par les arbitres de la Ligue Nationale de Hockey en séries éliminatoires en fonction de la présence ou non de spectateurs dans les gradins. Les résultats suggèrent que les arbitres accordent significativement moins de pénalités à l'équipe locale qu'à l'équipe adverse lorsque la foule est présente. En revanche, lorsqu'il n'y a pas de spectateurs, le nombre de pénalités accordées aux deux équipes ne diffère pas. Une généralisation de ces résultats a été réalisée dans d'autres contextes, permettant de constater que l'avantage accordé à l'équipe locale par les arbitres devant une foule s'observe aussi en saison régulière, autant dans la Ligue Nationale de Hockey que dans une ligue de niveau inférieur, la Ligue Canadienne de Hockey. Dans un second temps, l'influence de la critique excessive du travail des arbitres par les joueurs et les entraîneurs a été étudiée dans le domaine du baseball professionnel. Dans ce projet, l'influence de l'agression verbale, mesurée par une éjection de la partie, a été analysée pour déterminer son impact sur la prise de décision de l'arbitre au marbre. Les résultats suggèrent un avantage bidirectionnel du recours à la violence verbale. Après avoir été excessivement critiqués, les arbitres au marbre sont moins susceptibles d'accorder des prises aux frappeurs de l'équipe agressive et plus enclins à appeler des prises aux frappeurs de l'équipe adverse. Une série d'analyses supplémentaires permet de rejeter une hypothèse alternative, soit que les arbitres cherchent à compenser les conséquences négatives liées à la perte d'un coéquipier qui découle de leur décision préalable d'éjecter quelqu'un. Les résultats des deux projets de recherche permettent de mettre en lumière l'influence des forces sociales et de leur valence sur la prise de décision des arbitres sportifs, dans le but d'aider ces derniers à réduire les biais décisionnels dont ils sont victimes et à rendre le contexte sportif plus équitable et sécuritaire.

Mots-clés : Arbitres sportifs, Prise de décision, Sport professionnel, Biais décisionnels, Pression sociale, Forces sociales.

TABLE DES MATIÈRES

REMERCIEMENTS.....	IV
RÉSUMÉ	VI
TABLE DES MATIÈRES	VII
LISTE DES TABLEAUX	IX
LISTE DES FIGURES.....	X
LISTE DES ABRÉVIATIONS	XI
AVANT-PROPOS.....	XIII
CHAPITRE I	15
INTRODUCTION	15
<i>Les exigences du métier d'arbitre</i>	16
Arbitrer, une tâche cognitive.....	16
Le traitement de l'information	17
<i>Les sources de biais décisionnels</i>	21
Les joueurs.....	21
Les entraîneurs	25
La foule.....	26
<i>Les lacunes de la recherche sur le jugement et la prise de décision dans le sport</i>	29
La pression sociale.....	31
Étudier la pression sociale en contexte particulier	32
<i>Objectifs</i>	32
L'influence de la foule.....	33
L'influence de l'agression verbale des joueurs et entraîneurs	34
CHAPITRE II	35
ARTICLE 1: THE ABSENCE OF FANS REMOVES THE HOME ADVANTAGE ASSOCIATED WITH PENALTIES CALLED BY NATIONAL HOCKEY	
LEAGUE REFEREES.....	35
<i>Abstract</i>	35
<i>Introduction</i>	36
<i>Methods and results</i>	39
NHL playoffs data set	39
NHL playoffs model	42
NHL playoffs analysis and results.....	43
Regular season data sets	46
Regular seasons model	48
Regular seasons analysis and results	49
<i>Discussion</i>	53
<i>Conclusion</i>	56
<i>Acknowledgments</i>	57
CHAPITRE III	58
ARTICLE 2: VERBAL AGGRESSIONS AGAINST MAJOR LEAGUE BASEBALL UMPIRES AFFECT THEIR DECISION MAKING	58
<i>Abstract</i>	58
<i>Statement of Relevance</i>	59
<i>Introduction</i>	60

<i>Method</i>	65
Data acquisition.....	65
Analyses	68
<i>Results</i>	72
Primary analysis.....	74
Additional analysis	76
<i>Discussion</i>	85
CHAPITRE IV.....	91
DISCUSSION GÉNÉRALE	91
<i>L'influence des forces sociales sur la prise de décision des arbitres</i>	91
<i>Les mécanismes explicatifs potentiels</i>	92
La pression sociale.....	92
Le biais de négativité.....	93
Le biais de négativité pour expliquer le comportement des arbitres sportifs.....	94
<i>Implications potentielles des résultats</i>	97
Implications théoriques	97
Implications pratiques	100
<i>Contexte et limites des résultats</i>	103
<i>Directions futures</i>	105
Les projets en réalité virtuelle.....	105
L'influence des caractéristiques des participants	107
L'impact du biais de négativité.....	109
Les sous-catégories de pression sociale	110
<i>Synthèse</i>	111
CHAPITRE V	113
CONCLUSION	113
RÉFÉRENCES.....	115
ANNEXE A – RÉPONSES DU COMITÉ D'ÉTHIQUE À LA RECHERCHE DE L'UQO À LA SUITE DES DÉCLARATIONS DE PROJET DE RECHERCHE	139
ANNEXE B - SAGE CREATIVE COMMONS LICENSE: CONTRIBUTOR' S PUBLISHING AGREEMENT	142
ANNEXE C – REPORT FOR THE PSYCHOLOGICAL SCIENCE MASS REPRODUCTION INITIATIVE	146
ANNEXE D – MATÉRIEL SUPPLÉMENTAIRE DE L'ARTICLE 2.....	157

LISTE DES TABLEAUX

Tableau 1 : Penalty kills per year during the NHL playoffs.....	41
Tableau 2 : Estimated regression parameters for NHL playoffs Poisson GLMM (n=1094).....	43
Tableau 3 : Post-hoc results of the NHL model during playoffs (n=1094)	44
Tableau 4 : Penalty kills per year during the NHL and CHL regular seasons.....	48
Tableau 5 : Estimated regression parameters for NHL (n=3278) and CHL (n=3418) regular seasons Poisson	49
Tableau 6 : Post-hoc results of the NHL (n=3278) and CHL (n=3418) models during regular seasons.....	50
Tableau 7 : Summary of analysis from all seven models and their anticipated results according to the primary (excessive criticism) and alternative (compensate for ejection) hypotheses	63
Tableau 8 : Descriptive statistics and accuracy of MLB umpires by team, before and after ejection for excessive strike zone criticism	73
Tableau 9 : Impact of ejections for strike zone related criticisms on strike decisions (model 1).....	75
Tableau 10 : Summary of results from all seven models and their interpretation according to the primary and alternative hypotheses	87

LISTE DES FIGURES

Figure 1 : Séquence du traitement de l'information sociale (Bless et al., 2004)	18
Figure 2 : Cadre théorique de la cognition sociale du jugement humain. Modèle simplifié de Betsch et al. (2011), cité dans Plessner et al. (2023).....	20
Figure 3 : Mean Penalty kills per team with and without attendance during the NHL playoffs	45
Figure 4 : Mean Penalty kills per team with and without attendance during the NHL regular season.....	51
Figure 5 : Mean Penalty kills per team with and without attendance during the CHL regular season	52
Figure 6 : Called strike probability in games with ejections related to excessive strike zone criticism as a function of the baseline probability observed in games without ejections and the team ejected (n = 123,060).	76
Figure 7 : Called strike probability in games with ejections as a function of the baseline probability observed in games without ejections and the reason for the ejection (n = 153,255)	80
Figure 8 : Called strike probability in games with ejections as a function of the baseline probability observed in games without ejections and the team member ejected (n = 103,913).....	82

LISTE DES ABRÉVIATIONS

ABTO	Aggressive behavior towards opponent
APA	American Psychological Association
AS	All-Star
BCSprob	Baseline called-strike probability
C	Coaches
CER	Comité d'éthique de la recherche
CHL	Canadian Hockey League
CI	Confidence interval
FE	Fixed effects
GAM	Generalized additive model
GLMM	Generalized linear mixed model
HSD	Honestly significant difference
IRR	Incidence rate ratio
MLB	Major League Baseball
NAS	Not All-Star
NHL	National Hockey League
NSZ	Non-strike zone related criticism
OHL	Ontario Hockey League
OR	Odds ratio
P	Players
PK	Penalty kills
QMJHL	Quebec Major Junior Hockey League

RE	Random effects
ROC	Receiver operating characteristic
SD	Standard deviation
SZ	Strike zone related criticism
WAR	Wins above replacement
WHL	Western Hockey League

AVANT-PROPOS

La thèse doctorale qui suit propose un format par article. Ce sont deux articles distincts qui ont été insérés à titre de partie intégrante de la présente thèse. Pour les deux articles, l’auteur principal se veut le candidat au doctorat, Joël Guérette, alors que les coauteurs représentent son directeur de recherche, Daniel Fiset, ainsi que sa codirectrice, Caroline Blais. L’étudiant est responsable des étapes de la conceptualisation, de la recension des écrits, de l’élaboration de la méthodologie, de la collecte des données, des analyses, de la création du contenu visuel et de la rédaction des versions originales des articles. Pour leur part, le directeur et la codirectrice ont contribué à la conceptualisation, à la méthodologie, à la validation et à la supervision, en plus de la révision et de l’édition des manuscrits. Les deux articles qui découlent de ce processus ont été soumis à des journaux scientifiques, ont subi un processus de révision par les pairs et sont publiés.

Le premier article de la thèse, présenté au chapitre II et intitulé *The absence of fans removes the home advantage associated with penalties called by National Hockey League referees*, est publié par le journal *PloS one* depuis le 20 août 2021 (<https://doi.org/10.1371/journal.pone.0256568>). Pour sa part, le second article, qui fait l’objet du chapitre III, *Verbal Aggressions Against Major League Baseball Umpires Affect Their Decision Making*, est publié dans la revue *Psychological Science* et est disponible en ligne depuis le 20 février 2024 (<https://doi.org/10.1177/09567976241227411>).

Le texte inséré dans le présent document représente exactement celui proposé dans les articles à la suite du processus de révision par les pairs. Le format a été ajusté pour respecter les normes exigées par le Département de psychoéducation et de psychologie quant à la présentation d’un projet de thèse doctorale. À l’exception de ces ajustements de nature esthétique et des numéros et noms des tableaux et des figures, le contenu a été repris avec le plus de fidélité

possible. Pour assurer l'exactitude des termes, le texte présenté a d'ailleurs été conservé dans la langue de publication, soit l'anglais, en respect de l'article 2.13 du *Règlement des études de cycles supérieurs* de l'UQO.

Prenant appui sur des données libres d'accès, les deux projets complétés ne requerraient pas de certification éthique. Des déclarations de projets de recherche ont été soumises au Comité d'éthique de la recherche (CER) de l'UQO, lequel a confirmé que les travaux réalisés ne nécessitaient pas d'évaluation éthique (Annexe A).

De plus, les deux articles sont publiés sous la licence *Creative Commons Attribution (CC BY 4.0)*, laquelle permet à toute personne ou entité de réutiliser la contribution sans autorisation, à condition que la production scientifique soit correctement référencée. Cette licence permet également aux auteurs de conserver les droits d'auteurs, comme stipulé sur la page couverte de l'article publié dans *PloS one* et dans l'accord signé avec *Psychological Science* présenté à l'annexe B.

CHAPITRE I

Introduction

Le domaine sportif est un environnement écologique de choix pour observer les interactions sociales entre les êtres humains. Dans un contexte où l'esprit de compétition est valorisé et que les résultats peuvent apporter des gains émotionnels et financiers importants aux participants, les comportements les plus primaux vont souvent ressurgir au détriment de ceux considérés comme socialement adéquats (Maher et al., 2014). Même si les différentes disciplines sportives possèdent des règles bien définies, il est utopique de penser que les participants, influencés par leur soif de victoire, vont respecter les règles de manière impartiale. C'est pourquoi la presque totalité des compétitions sportives fait appel à des arbitres. Ces derniers ont le mandat de faire respecter les règles, dans l'optique de favoriser une compétition juste et équitable où les meilleurs pourront être déclarés vainqueurs.

Bien que ce mandat semble clair et bien défini, il n'en demeure pas moins que son application est beaucoup plus ardue que ce que suggère la théorie. Lors de leur prise de décision, les arbitres se retrouvent souvent à l'interface entre le désir d'être impartial et celui d'être équitable, tout en devant prendre des décisions sur des situations ambiguës dans un contexte où la vitesse de l'action ne leur laisse que quelques fractions de seconde pour réagir. Ils sont donc à haut risque de rendre des décisions biaisées et de recevoir des critiques vis-à-vis de ces décisions (Anderson & Pierce, 2009). Si la probabilité élevée que les arbitres expriment des décisions biaisées fait rager les joueurs, entraîneurs et spectateurs, elle ouvre toutefois la porte à l'observation des éléments qui peuvent influencer le processus décisionnel humain. C'est pour cette raison qu'en prenant racine dans l'environnement sportif, la thèse doctorale qui suit vise à

explorer et à comprendre comment certains éléments et comportements influencent la prise de décision des arbitres.

Les exigences du métier d'arbitre

Arbitrer, une tâche cognitive

Les exigences du travail d'arbitre varient selon la discipline sportive. Certains sports exigent que les officiels se déplacent pour suivre le jeu, comme au football, au basketball ou au soccer. Dans d'autres disciplines, comme le plongeon ou la boxe, les arbitres, aussi nommés juges, doivent demeurer assis pour observer l'action. Si la demande physiologique varie considérablement d'un sport à l'autre, la composante psychologique du métier d'arbitre se généralise quant à elle à toutes les disciplines. En effet, tous les arbitres doivent traiter une grande quantité d'information en provenance de leur environnement avant de pouvoir rendre des décisions. C'est pourquoi le métier d'officiel est considéré comme étant principalement une tâche cognitive (MacMahon & Mildenhall, 2012).

Ainsi, pour exceller dans l'accomplissement de son travail, un arbitre doit posséder des habiletés mentales spécifiques. MacMahon et al. (2007a) proposent qu'une des qualités les plus importantes soit la capacité à prendre des décisions rapides et précises. Les officiels doivent aussi être en mesure de maintenir leur attention et leur concentration sur les éléments pertinents, gérer la pression, faire face aux situations négatives et se redresser après une mauvaise décision (Guillén & Feltz, 2011). Le travail d'arbitre implique aussi la gestion du comportement des différents individus faisant partie de l'environnement de compétition. Que ce soit les entraîneurs qui tentent de faire pencher les décisions en leur faveur (Debanne, 2014), les plaintes de certains joueurs (Lex et al., 2015) ou les commentaires des spectateurs (Myers & Balmer, 2012), chaque intervention des participants en présence vient compliquer le travail d'un arbitre. Même certains

acteurs non impliqués directement dans l'action peuvent exercer une influence sur le travail d'un officiel. En effet, les directives des dirigeants de la ligue ou des superviseurs des arbitres peuvent contraindre l'officiel à traiter de l'information supplémentaire, selon les consignes qu'il a reçues avant le match (Price et al., 2012). Le métier d'officiel est donc complexe et les sources d'informations à manipuler sont abondantes. C'est pourquoi il est pertinent de se pencher sur la demande cognitive qu'engendre le traitement de toutes ces informations pour bien comprendre le travail psychologique nécessaire à l'arbitrage.

Le traitement de l'information

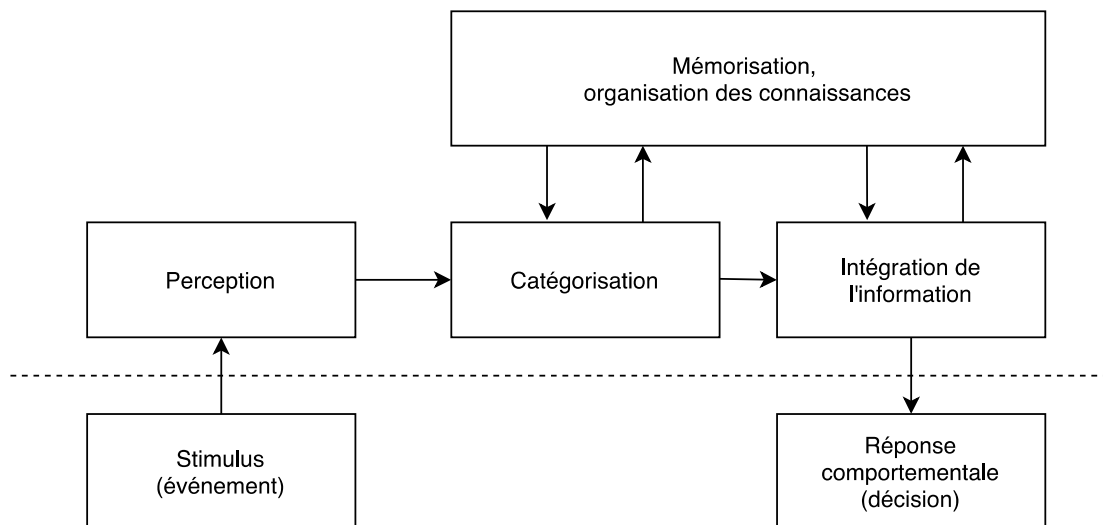
Si l'environnement sportif abonde en stimuli, cela ne veut pas dire que toutes les informations soient utiles pour un arbitre. Au contraire, il semblerait que la complexité de l'environnement nuise au travail des officiels. En effet, MacMahon et Mildenhall (2012) soulignent que l'information pertinente disponible est souvent faible pour l'arbitre, ce qui augmente la difficulté de sa tâche. L'information manquante devra être inférée par l'officiel, ce qui augmente la demande cognitive. S'ajoute au manque d'information utile la vitesse spectaculaire à laquelle les athlètes exécutent leurs actions et la nécessité pour l'arbitre de prendre une décision rapide et juste. Comme si la situation n'était pas déjà assez compliquée, une grande partie des situations que l'arbitre doit interpréter est ambiguë et un jugement parfaitement objectif de la situation est souvent impossible (Van Quaquebeke & Giessner, 2010). Ces éléments créent une demande perceptive et cognitive supérieure à la capacité de traitement de l'information d'un être humain (MacMahon et al., 2007b). L'arbitre devient ainsi vulnérable à plusieurs facteurs pouvant influencer ses décisions, en raison de sa capacité limitée à traiter l'information (MacMahon & Mildenhall, 2012). Juger les situations de match devient un travail

excessivement complexe et force les arbitres à se tourner vers des stratégies non optimales pour fournir la performance la plus adéquate possible.

Pour s'acquitter de leur tâche, certains auteurs suggèrent que les arbitres analysent l'information en suivant les principes de la cognition sociale (Plessner & Haar, 2006), qui s'intéresse à la connaissance sociale et aux processus cognitifs impliqués lorsqu'un individu tente de comprendre la réalité (Fiske & Taylor, 1991). Les processus cognitifs impliqués dans le traitement de l'information sociale incluent la perception, la catégorisation, la mémorisation et l'intégration de l'information (Bless et al., 2004; Plessner & Haar, 2006). La Figure 1 présente la séquence du traitement de l'information et l'interaction entre les processus cognitifs impliqués.

Figure 1

Séquence du traitement de l'information sociale (Bless et al., 2004)



La séquence commence lorsqu'un arbitre perçoit un stimulus, par exemple, lorsqu'un joueur de hockey donne un coup de bâton à son adversaire. L'arbitre doit ensuite catégoriser ce geste comme légal ou illégal. Pour y arriver, l'arbitre s'appuie sur sa connaissance des règlements et sur les décisions antérieures, conservées en mémoire, qu'il a rendues sur des gestes similaires depuis le début de la partie ou durant les parties précédentes. Le stimulus perçu sera également

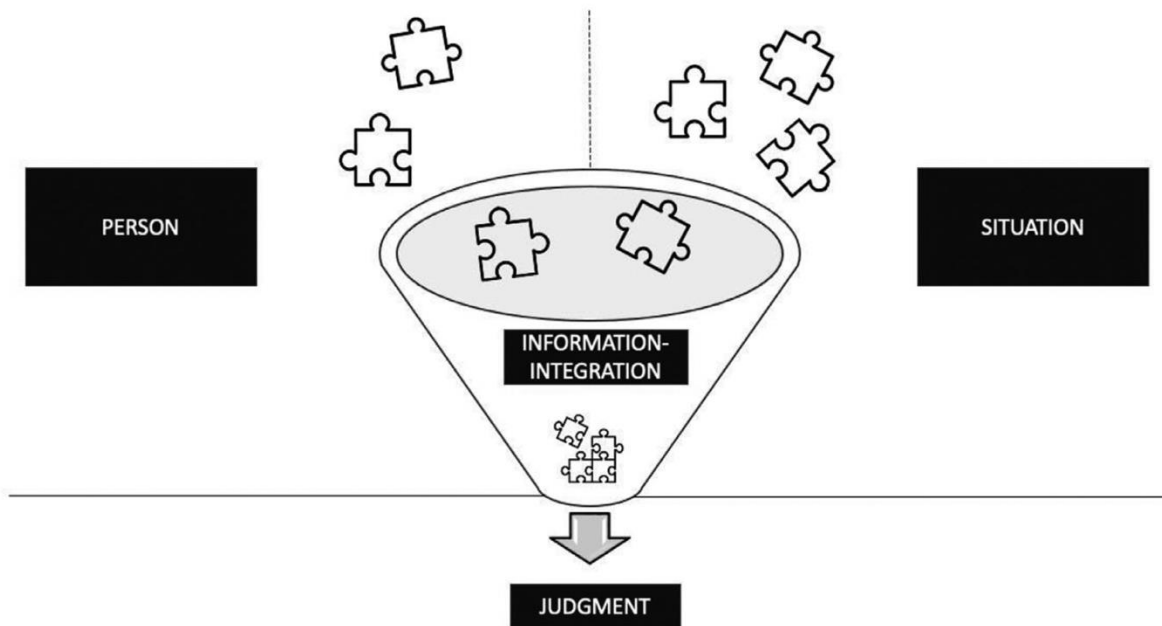
encodé en mémoire et pourra influencer les décisions futures de l'arbitre. Finalement, l'arbitre intègre les informations perçues, catégorisées et mémorisées, en plus d'inférer, si nécessaire, l'information manquante. Cette intégration de l'information lui permet de rendre une décision, comme décerner une pénalité au joueur qui a donné un coup de bâton.

En proposant d'utiliser la séquence du traitement de l'information sociale pour étudier le jugement et la prise de décision en contexte sportif, l'une des intentions de Plessner et Haar (2006) était d'intégrer les études existantes dans un cadre théorique global afin d'offrir au domaine une structure conceptuelle cohérente. Cette structure allait pouvoir, toujours selon le souhait des auteurs, orienter les travaux de recherche futurs. En imposant un modèle qui requiert la présence des mécanismes cognitifs sous-jacents à la prise de décision, les chercheurs souhaitaient mettre de l'avant l'importance de porter attention à ces éléments lors des travaux. Malheureusement, la complexité des interactions entre les mécanismes cognitifs dans ce modèle semble avoir posé entrave à son utilisation par les scientifiques.

Cette particularité a mené à la révision du modèle dans un but de simplicité, encore une fois en appuyant la réflexion sur les théories de la cognition sociale (Betsch et al., 2011, cité dans Plessner et al., 2023). La nouvelle structure, décrite comme un cadre unifié de jugement social, élimine la nécessité d'isoler les nombreuses étapes sous-jacentes pouvant influencer la prise de décision et se concentre plutôt sur les deux sources d'influence potentielles, soit : 1- la personne qui émet le jugement et ses dispositions personnelles (l'individu) ; 2- la situation ou l'environnement à partir desquels les informations relatives au jugement ont été générées (le contexte).

Figure 2

Cadre théorique de la cognition sociale du jugement humain. Modèle simplifié de Betsch et al. (2011), cité dans Plessner et al. (2023)



En s'appuyant sur ce modèle et en l'appliquant à la réalité des arbitres sportifs, deux grandes catégories de projets peuvent maintenant émerger : les études concernant l'influence des différences individuelles sur la prise de décision et celles en lien avec l'influence du contexte sur le processus décisionnel. Dans le cadre de la présente thèse doctorale, les travaux seront orientés du côté de l'influence que peuvent avoir les éléments du contexte sur les processus décisionnels des arbitres sportifs. Puisque de nombreux éléments liés au contexte peuvent interférer avec ces processus, une attention particulière sera portée à certaines forces sociales faisant partie du contexte de l'arbitre. Plus spécifiquement, l'attention sera orientée vers la manière dont les autres individus qui font partie de l'environnement de l'arbitre peuvent l'influencer. Pour l'arbitre, ces individus sont principalement représentés par les joueurs, les entraîneurs et les spectateurs. Après avoir établi l'état des connaissances en lien avec l'influence de ces acteurs sur la prise

décisionnelle des arbitres, il sera possible de déterminer comment la présente thèse doctorale pourra contribuer à l'évolution des connaissances du domaine de la prise de décision en contexte sportif.

La prochaine section brosse le portrait des connaissances actuelles en matière de biais décisionnels chez les arbitres, en les regroupant selon les acteurs présents durant les compétitions sportives ayant le potentiel d'influencer la prise de décision, c'est-à-dire les joueurs, les entraîneurs et les spectateurs (Bar-Eli et al., 2011).

Les sources de biais décisionnels

La recherche antérieure permet de soulever de nombreux facteurs ayant le potentiel d'influencer la prise de décision des arbitres sportifs. Ces facteurs, qu'ils soient individuels en ce qui concerne les joueurs et les entraîneurs ou qu'ils soient propres au groupe dans le cas des spectateurs, sont décrits ci-après, dans les sections spécifiques à chacun des acteurs présents dans le domaine sportif.

Les joueurs

Ethnie et nationalité

Comme dans beaucoup d'autres domaines, l'ethnie peut influencer la prise de décision d'un officiel. Les biais ethniques ont d'ailleurs été observés dans de nombreuses disciplines sportives. Au soccer professionnel, on peut observer un avantage pour les joueurs en termes de pénalités accordées lorsque ces derniers proviennent du même pays natal que l'arbitre (Pope & Pope, 2015). Un biais de nationalité serait à l'origine des résultats, une distinction que les auteurs apportent par rapport au biais ethnique. Dans un ordre d'idées similaire, des chercheurs ont observé que le nombre de pénalités accordées aux joueurs de soccer de couleur plus foncée sous forme de fautes et de cartons est supérieur, même en contrôlant pour la nationalité (Magistro et

Wack, 2023). Toujours au soccer, une différence a été répertoriée dans la vitesse de jugement d'une pénalité lorsqu'un geste illégal est commis par un joueur blanc comparativement à un athlète noir, ces derniers étant réprimandés plus rapidement par les participants à l'étude (Wagner-Egger et al., 2012).

Des arbitres de volley-ball féminin sont également plus sévères à l'endroit des joueuses provenant de collèges et d'universités historiquement noirs (Dix, 2023). L'iniquité est observable par un nombre disproportionné de fautes de maniement de ballon, que l'auteur explique par la théorie de la violation des attentes. Au basketball professionnel, les arbitres ont tendance à accorder plus de fautes aux joueurs d'une ethnie différente comparativement à ceux de leur propre ethnie (Price & Wolfers, 2010). Les auteurs proposent que les biais ethniques implicites des arbitres puissent être responsables du phénomène.

Les athlètes parlant une langue différente de celle de l'arbitre peuvent aussi être victimes de biais décisionnels de la part de ce dernier. Mongeon et Longley (2015) ont montré que les officiels franco-canadiens appellent des pénalités plus fréquemment aux joueurs anglo-canadiens que le font les officiels anglo-canadiens dans la Ligue Nationale de Hockey. En Suisse, les arbitres de soccer ont tendance à désavantager les joueurs des équipes qui ne proviennent pas de la même région linguistique qu'eux en leur accordant plus de cartons jaunes (Faltings et al., 2023).

Genre

Le genre d'un athlète peut aussi mener un officiel à juger différemment son comportement. Souchon et al. (2009; 2013b) ont montré que les officiels de handball de différents niveaux sont plus sévères envers les femmes que les hommes. Leurs résultats, basés sur des études expérimentales où les arbitres devaient juger des situations de matchs présentées en

format vidéo, seraient principalement explicables par des stéréotypes de genre, lesquels affectent autant les officiels masculins que féminins. Un constat similaire est rapporté par Coulomb-Cabano et al. (2005) au soccer, dans un projet étudiant le comportement des arbitres de matchs de niveau national. Les participants, tous des arbitres masculins, pénalisent plus souvent les joueuses que les joueurs pendant les compétitions. Les arbitres féminines seraient également plus enclines à pénaliser les femmes qui pratiquent le basketball de niveau collégial et secondaire, comparativement aux hommes de même niveau (Graf et al., 2009). Ces résultats sont tirés d'une étude utilisant une méthodologie différente des auteurs précédents, soit l'utilisation de vignettes plutôt que de situations de matchs présentées sous forme de vidéo.

Taille

Lors de situations de fautes ambiguës au soccer, le joueur le plus grand a tendance à être catégorisé comme l'agresseur alors que le plus petit reçoit souvent le titre de victime, menant à des sanctions plus fréquentes chez les joueurs de grande taille (Van Quaquebeke & Giessner, 2010). Les résultats s'appuient autant sur des statistiques en situations réelles de matchs que des tâches en laboratoire. La différence de taille entre les joueurs et les arbitres pourrait aussi influencer la prise de décision de ces derniers, alors que les arbitres plus petits auraient tendance à accorder plus de fautes aux joueurs qui leur sont supérieurs en grandeur. C'est le cas au basketball (Gift & Rodenberg, 2014) et pour certaines sanctions au soccer (McCarrick et al., 2020), selon deux études corrélationnelles effectuées sur la base des statistiques de ligues professionnelles. Les préjugés implicites basés sur la taille, communément appelés le complexe de Napoléon, pourraient expliquer ce phénomène.

Statut

Les joueurs vedettes pourraient bénéficier de décisions avantageuses de la part des arbitres. Ce type de biais est appuyé par Kim et King (2014), qui soutiennent que les lanceurs de baseball possédant la meilleure réputation obtiennent des décisions favorables en regard de la zone de prises. Mills supporte aussi cette hypothèse (2014), en suggérant que les frappeurs vedettes et ceux possédant plus d'expérience bénéficient de la clémence de l'arbitre au marbre lorsque vient le temps de décerner des prises au baseball professionnel. Les joueurs étoiles de basketball seraient également avantagés par les décisions des arbitres de leur discipline sportive, recevant plus d'opportunité de lancers francs en fin de partie éliminatoire que les joueurs de statuts moins élevés (Caudill et al., 2014). Les arbitres auraient également une prédisposition à décerner plus de fautes de manière générale aux joueurs qui tentent de contrer les joueurs vedettes, offrant ici aussi un avantage lié au statut de joueur étoile (Barrett, 2021). Le statut des joueurs a été déterminé par le salaire dans cette étude réalisée auprès de la Ligue Nationale de Basketball. La réputation des athlètes offrirait également des avantages en patinage artistique, où les juges qui connaissent les patineuses auraient tendance à les évaluer plus favorablement que leurs collègues non familiers avec les athlètes (Findlay & Ste-Marie, 2004). Ce constat découle d'un projet effectué auprès de 12 juges et de 14 patineuses artistiques et propose comme mécanisme explicatif un biais de réputation.

Séances d'échauffement

Certains arbitres seraient également victimes de biais liés à des actions perpétrées avant la compétition qu'ils se doivent de juger en temps réel. Par exemple, les juges de gymnastique ont tendance à utiliser l'information acquise lors des séances d'échauffement et conservée en

mémoire pour prédire la performance lors de la compétition, et ce, même si la séance d'échauffement a été vue par le juge il y a plus d'une semaine (Ste-Marie & Valiquette, 1996).

Proximité des joueurs par rapport à l'arbitre

Les joueurs ayant une plus grande proximité avec l'officiel par la nature de leur positionnement, comme les receveurs au baseball, profitent de décisions biaisées (Mills, 2014). Selon l'auteur, la proximité physique du receveur, qui se positionne tout juste devant l'arbitre au marbre dans la Ligue majeure de Baseball, augmente la pression sociale générée envers ce dernier et serait à l'origine de l'influence dans la prise de décision.

Les entraîneurs

Ethnie

À l'instar des joueurs, l'ethnie des entraîneurs peut influencer la prise de décision des arbitres. C'est le cas au football collégial, où les équipes dirigées par un entraîneur noir sont pénalisées plus souvent que celles sous les commandes d'un entraîneur blanc (Davis et al., 2022). Cette conclusion prend appui sur deux indicateurs, soit le nombre moyen de pénalités par match et le nombre moyen de verges de pénalité par match. Comme pour l'impact de l'ethnie des joueurs présenté plus tôt, les stéréotypes ethniques pourraient expliquer les observations des chercheurs.

Expertise et charisme

Debanne (2014) suggère que certains entraîneurs de handball utilisent des techniques de persuasion basées sur leur expertise ou leur charisme pour influencer les décisions des arbitres. Selon l'auteur, lorsqu'un entraîneur critique les décisions d'un arbitre en mettant en lumière sa supériorité au niveau des connaissances techniques, il aurait plus de probabilité d'influencer les futures décisions de l'officiel de manière à obtenir des avantages pour son équipe. Le projet

exploratoire de Debanne nécessite, selon ses propres dires, d'autres études pour confirmer l'impact des divers types d'influence sociale sur le processus décisionnel des arbitres.

Réactions

Les réactions des entraîneurs sont des stimuli ayant le potentiel d'influencer les décisions des arbitres. C'est du moins ce que suggèrent Souchon et al. (2013a), qui mentionnent que l'exclamation d'un entraîneur de judo modifie l'interprétation que fait l'arbitre de situations ambiguës et tend à offrir un avantage pour son athlète. Les auteurs admettent qu'il est cependant impossible de confirmer les processus motivationnels ou cognitifs impliqués.

La foule

Bruit

Le bruit de la foule a été identifié comme un stimulus susceptible d'influencer les décisions des arbitres à l'avantage des athlètes locaux. Un peu comme c'est le cas avec les réactions de l'entraîneur au judo, Myers et Balmer (2012) ont montré que le combattant local en Muay Thai peut bénéficier du bruit de la foule pour obtenir des jugements favorables de l'officiel. L'explication la plus intuitive, selon les auteurs, provient de la conformité sociale.

Dans un même ordre d'idées, différents auteurs ont étudié l'effet du bruit de la foule sur les fautes accordées par des arbitres de soccer. Lorsque le bruit de la foule est complètement retiré expérimentalement, il semble que les officiels soient plus sévères à l'endroit de l'équipe locale que lorsque le bruit de la foule est conservé (Nevill et al., 2002). Une explication plausible réside dans le stress induit par les représailles de la foule lors de décisions défavorables à l'endroit de l'équipe locale. Les arbitres pourraient chercher à éviter ces réactions négatives à leur égard en présence d'une foule ayant le potentiel d'exprimer son mécontentement. Lorsque l'intensité du bruit est prise en considération, il semble cette fois que les arbitres soient plus

sévères à l'endroit de l'équipe visiteuse (Unkelbach & Memmert, 2010). Les auteurs avouent humblement que leur étude n'a pas mesuré les processus cognitifs sous-jacents. Néanmoins, leurs résultats, combinés à ceux de Nevill et al. (2002), propose que les biais liés à la présence ou à l'intensité du bruit avantagent l'équipe locale, que ce soit par des décisions plus sévères envers l'équipe visiteuse ou par des décisions plus clémentes relativement à l'équipe locale.

Ces résultats doivent toutefois être considérés avec prudence avant d'être généralisés. En effet, selon Sors et al. (2019) qui se sont intéressés à la question du bruit de la foule au basketball, il est possible que seuls les arbitres plus anxieux se laissent influencer par les réactions des spectateurs. Leur étude, comparant une condition où la foule exerce beaucoup de pression à une autre où les spectateurs génèrent peu de pression, permet de tirer ces conclusions.

Nombre de spectateurs

Le volume de la foule pourrait aussi influencer la prise de décision chez des arbitres de soccer, offrant encore une fois un avantage à l'équipe locale (Downward & Jones, 2007; Page & Page, 2010; Picazo-Tadeo et al., 2017). Les arbitres auraient tendance à vouloir apaiser la foule, surtout lorsque les spectateurs augmentent en nombre. La réponse à la pression sociale qu'exerce la foule pourrait toutefois mener à des réponses différentes selon les arbitres, ceux-ci possédant des caractéristiques individuelles distinctes permettant de faire face de manière plus ou moins efficace à la pression grandissante.

Ces résultats semblent néanmoins difficiles à généraliser à d'autres disciplines que le soccer. En effet, selon Szabó (2022), qui a utilisé les statistiques de fréquentation des stades de la dernière décennie, un plus grand nombre de spectateurs pourrait avoir des effets divergents en fonction du sport étudié. L'auteur suggère qu'une plus grande foule pourrait générer des décisions favorables pour l'équipe locale dans la Ligue Nationale de Basketball, des décisions

avantageuses pour l'équipe visiteuse dans la Ligue Nationale de Football et n'aurait aucun impact dans la Ligue Nationale de Hockey. L'auteur propose que les résultats soient explicables par la pression sociale, laquelle a des effets différents selon le contexte.

Proximité

La proximité de la foule par rapport au terrain vient aussi moduler la prise de décision des officiels au soccer. Lors de parties disputées dans des stades où les spectateurs sont séparés du terrain par une piste d'athlétisme, la tendance des arbitres à avantager l'équipe locale est plus faible que lorsque la foule est collée sur la surface de jeu (Buraimo et al., 2010; Buraimo et al., 2012; Dohmen, 2008; Scoppa, 2008). Encore une fois, l'explication primaire résiderait dans la pression sociale, laquelle augmente lorsque les spectateurs se retrouvent plus près de l'action. L'effet semble généralisable à plusieurs ligues de soccer professionnelles établies dans différents pays d'Europe.

Densité

La densité de la foule, mesurée par le nombre de spectateurs divisé par la capacité du stade, semble aussi avoir un rôle à jouer dans la prise de décision des arbitres. Boyko et al. (2007), tout comme Goumas (2014), proposent que ce soit le cas au soccer. Dans le cadre du premier article, les auteurs suggèrent que des mécanismes psychologiques individuels permettent aux arbitres de faire face à la pression différemment, soulignant la nécessité de pousser l'investigation plus loin pour comprendre ces mécanismes. En ce qui concerne l'étude de Goumas (2014), l'auteur se contente de décrire la densité de la foule comme étant la cause probable de biais décisionnels chez les arbitres.

Composition

Dohmen (2008) ajoute que la composition de la foule, soit la proportion de partisans de chaque équipe dans les amphithéâtres, peut aussi influencer l'intégration de l'information chez un arbitre. Ici aussi, l'auteur propose l'hypothèse de la pression sociale pour expliquer ses résultats. Cette proposition est appuyée par Ponzo et Scoppa (2018), eux qui ont, comme leur prédécesseur, étudié le phénomène dans les stades de soccer européens. Ces derniers parlent toutefois de support de la foule plutôt que de pression sociale, laquelle exercerait une influence positive sur les joueurs locaux en plus d'influencer négativement le travail des arbitres.

Les lacunes de la recherche sur le jugement et la prise de décision dans le sport

L'ensemble des sources de biais décisionnels identifiés dans les écrits scientifiques suggère que l'information sociale en présence dans l'environnement sportif, notamment les différents acteurs impliqués, peut avoir un impact important sur la prise de décision des arbitres. Toutefois, une majorité des études présentées à la section précédente se soldent par des propositions prudentes lorsque vient le temps de suggérer les mécanismes explicatifs potentiels. L'identification des biais est présentée, mais repose souvent sur des interprétations générales, c'est-à-dire les stéréotypes et la pression sociale.

C'est d'ailleurs une des principales critiques de la recherche sur le jugement et la prise de décision dans le sport des vingt dernières années (Raab et al., 2019). En effet, les experts du domaine en provenance d'approches diversifiées considèrent insatisfaisant le fait que la recherche sur le jugement et la prise de décision en contexte sportif tente rarement de mettre en lumière les processus cognitifs impliqués ou ne suggère pas de théories alternatives à celle de Plessner et Harr (2006) pour appuyer les résultats. Les auteurs renchérissent en affirmant que c'est en s'intéressant aux processus sous-jacents du jugement social que la recherche pourra aider à

améliorer la prise de décision en contexte sportif et que se contenter de rapporter des effets originaux n'offre que peu d'avancées pour le domaine. Les lacunes observées dans les écrits scientifiques sont potentiellement explicables par : 1- la quantité et l'interrelation des processus cognitifs impliqués dans le traitement de l'information sociale ; et 2- la complexité de l'information contenue dans l'environnement sportif, qui rend laborieux le processus visant à isoler complètement les variables étudiées.

Si cette complexité semble à l'origine des principales faiblesses des études en jugement et prise de décision en contexte sportif, elle est également à la source de la simplification du modèle visant à encadrer la recherche dans ce domaine. En retirant la nécessité d'isoler les étapes du traitement de l'information sociale, comme le suggère le modèle de Betsch et al. (2011), il est possible que les avancées en termes de connaissances dans le champ d'études visé progressent plus rapidement. Néanmoins, la simplicité du nouveau modèle pourrait aussi provoquer l'effet contraire, en exacerbant le problème d'identification unique des sources de biais, sans que les chercheurs poussent leur réflexion en termes de processus impliqués. Le risque principal de l'utilisation du modèle simplifié réside dans la possibilité que les scientifiques expliquent grossièrement leurs résultats par le processus global d'intégration de l'information, représenté par les morceaux de casse-têtes dans la Figure 2.

L'utilisation du modèle simplifié pourrait donc mener à des lacunes similaires à celles observées dans la recension des écrits présentée plus tôt en regard du travail des arbitres sportifs, soit que dans bon nombre de projets, l'explication proposée postule tout simplement que les arbitres sont victimes de stéréotypes ou de pression sociale. Pour éviter de tomber dans ce piège, les travaux élaborés dans la présente thèse doctorale prendront appui sur des situations qui pourraient permettre de préciser les mécanismes explicatifs, allant au-delà des concepts généraux préalablement suggérés. Comme mentionné précédemment, les éléments de contexte, plus

précisément les forces sociales générées par les individus en présence dans l'environnement sportif, seront priorisés comme sujet d'étude. C'est donc en prenant racine dans des situations ayant la capacité de générer de la pression sociale que les contextes d'études seront identifiés, laissant de côté les éléments relatifs à l'individu, comme les stéréotypes.

La pression sociale

Si la pression sociale peut expliquer un grand nombre de biais cognitif, c'est principalement parce qu'elle est un concept qui englobe pratiquement toutes les interactions sociales possibles. En effet, selon l'*American Psychological Association* (APA), la pression sociale se définit comme « l'exercice d'une influence sur une personne ou un groupe par une autre personne ou un autre groupe » [traduction libre] (American Psychological Association, 2018, p. 1). Ce type d'influence se subdivise en trois sous-catégories, nommées l'influence informationnelle, l'influence normative et l'influence interpersonnelle. Toujours selon l'APA, la première sous-catégorie, l'influence informationnelle, comprendrait les stratégies de persuasion et les arguments rationnels. La seconde forme d'influence, la normative, serait pour sa part un appel à la conformité. Finalement, l'influence interpersonnelle serait constituée des formes directes d'influence, comme les menaces, les attaques personnelles, les promesses de récompenses et l'approbation sociale.

En regard des définitions présentées, il n'est pas surprenant de constater la tendance des auteurs du domaine du jugement et de la prise de décision en contexte sportif à expliquer les biais décisionnels des arbitres par la pression sociale : cette dernière inclut pratiquement tous les types de comportements sociaux pouvant exercer une influence sur l'arbitre. Si la stratégie généraliste d'expliquer les observations par la pression sociale évite de mener à de faux constats, elle demeure néanmoins imprécise et peu utile pour approfondir la compréhension des biais

décisionnels des arbitres. Ce manque de profondeur limite également les contributions potentielles au développement de formations visant à diminuer les biais décisionnels chez les officiels.

Étudier la pression sociale en contexte particulier

Comme mentionné précédemment, expliquer les biais décisionnels des arbitres par la simple pression sociale demeure insatisfaisant. Pour tenter de pallier cette lacune, la présente thèse doctorale utilisera des situations extrêmes de pression sociale pour tenter de mieux comprendre comment les forces sociales présentes dans l'environnement et induites par les joueurs, entraîneurs et spectateurs peuvent influencer la prise de décision des arbitres. En s'inspirant du cadre théorique de la cognition sociale du jugement humain qui suggère que la situation influence la prise de décision en milieu sportif, des situations particulières permettant de mettre à l'épreuve des théories sous-jacentes à la pression sociale seront utilisées.

Afin d'éviter de succomber à la tentation de simplement identifier des phénomènes originaux, les situations particulières seront analysées en regard de théories permettant de préciser de quelle manière la pression sociale contribue au processus décisionnel. L'approche proposée vise à répondre à l'une des lacunes identifiées par Raab et al. (2019), soit que la recherche sur le jugement et la prise de décision dans le sport ne suggère pas de théories alternatives à celle de Plessner et Haar (2006) pour expliquer les observations rapportées.

Objectifs

Deux situations extrêmes de pression sociale impliquant les différents acteurs présents dans les compétitions sportives ont été identifiées pour jeter les bases de deux projets distincts. L'objectif de la présente thèse doctorale est de mesurer l'impact de ces situations sur la prise de décisions des arbitres sportifs. D'abord, l'impact de la présence et de l'absence des spectateurs

sur la prise de décision des arbitres professionnels sera étudié, dans un contexte rendu possible par les consignes sanitaires mises en place lors de la pandémie de la COVID-19. Ensuite, la critique verbale excessive des décisions des arbitres par les acteurs principaux des joutes sportives, c'est-à-dire les joueurs et les entraîneurs, sera explorée.

L'influence de la foule

Comme la pression sociale exercée sur un arbitre provient de la présence d'autres individus dans son environnement, le premier projet aura comme objectif d'étudier le comportement des officiels lorsque les spectateurs sont complètement retirés des événements sportifs. Pour y arriver, le domaine du hockey professionnel servira de laboratoire. L'impact de la foule sur la prise de décision des arbitres de hockey a déjà été mis de l'avant par certains auteurs pour expliquer, en partie, le phénomène de l'avantage de la glace, mais jamais dans un contexte d'étude où la foule a pu être physiquement retirée des arénas. Avec la venue de la pandémie de la COVID-19, cette opportunité d'étudier le comportement des arbitres dans des arénas vides s'est présentée pour la première fois. Le premier article sera inspiré de cette opportunité et visera à comparer comment les forces sociales générées par la foule influencent la prise de décision des arbitres de hockey.

Plutôt que de se contenter d'expliquer les résultats par la pression sociale de manière générale, une approche visant à déterminer si les biais liés à l'avantage de la glace sont dus à l'évitement des potentielles sanctions sociales ou plutôt à la recherche d'approbation sociale sera présentée. Cette dualité sera discutée en regard des biais d'omission et de commission. Pour un arbitre de hockey, omettre de pénaliser l'équipe locale permet d'éviter les sanctions sociales, lesquelles sont présentes sous forme des représailles de la foule. L'action de se commettre,

notamment en pénalisant l'équipe visiteuse, peut pour sa part engendrer des récompenses sociales, lesquelles sont observables par les acclamations de la foule.

L'influence de l'agression verbale des joueurs et entraîneurs

Le second projet visant à évaluer l'impact d'une situation extrême sur la prise de décision des arbitres se concentre sur la pression sociale engendrée par l'un des comportements négatifs les plus courants dans le domaine sportif, soit la critique excessive du travail des arbitres. Pour ce faire, les épisodes d'agressions verbales des joueurs et entraîneurs de la Ligue majeure de Baseball à l'endroit des arbitres seront étudiés. Après avoir mesuré l'influence de la présence d'un grand nombre d'individus, c'est-à-dire l'ensemble des partisans présents aux matchs, il est pertinent de vérifier l'impact potentiel des comportements propres à un seul ou à quelques individus sur la prise de décision des arbitres. L'objectif de cette étude sera de vérifier l'impact unique des comportements négatifs extrêmes dans un contexte où les réactions excessives sont omniprésentes, soit le baseball professionnel. Spécifiquement, ce projet visera à déterminer l'impact des critiques excessives sur les décisions des arbitres liées à la zone de prises.

Cette fois, la pression sociale exercée par la personne qui critique le travail d'un arbitre sera évaluée en opposant des théories contradictoires en psychologie sociale, soit : 1- que l'agression verbale peut altérer la relation avec la victime et réduire la possibilité qu'elle adopte le point de vue de l'agresseur ou ; 2- qu'elle peut mener la victime à se comporter pour plaire à l'agresseur dans l'optique d'éviter de futures agressions. Comme dans le cas de l'étude réalisée dans la Ligue Nationale de Hockey, l'approche permettra d'analyser si les arbitres sont influencés par un désir d'éviter les sanctions sociales.

Les hypothèses proposées dans chacun des contextes extrêmes étudiés sont décrites dans leurs articles respectifs aux chapitres II et III.

CHAPITRE II

Article 1: The absence of fans removes the home advantage associated with penalties called by National Hockey League referees

Joël Guérette¹, Caroline Blais¹, and Daniel Fiset¹

¹Département de psychoéducation et de psychologie, Université du Québec en Outaouais

Abstract

The COVID-19 pandemic has had a major impact on professional sports, notably, forcing the National Hockey League to hold its 2020 playoffs in empty arenas. This provided an unprecedented opportunity to study how crowds may influence penalties awarded by referees in an ecological context. Using data from playoff games played during the COVID-19 pandemic and the previous 5 years ($n=547$), we estimate the number of penalties called by referees depending on whether or not spectators were present. The results show an interaction between a team's status (home; away) and the presence or absence of crowds. Post-hoc analyses reveal that referees awarded significantly more penalties to the away team compared to the home team when there is a crowd present. However, when there are no spectators, the number of penalties awarded to the away and home teams are not significantly different. In order to generalize these results, we took advantage of the extension of the pandemic and the unusual game setting it provided to observe the behavior of referees during the 2020-2021 regular season. Again, using data from the National Hockey League ($n=1639$), but also expanding our sample to include Canadian Hockey League games ($n=1709$), we also find that the advantage given to the home team by referees when in front of a crowd fades in the absence of spectators. These findings provide new evidence

suggesting that social pressure does have an impact on referees' decision-making, thus contributing to explain the phenomenon of home advantage in professional ice hockey.

Introduction

During the 2019 National Hockey League (NHL) playoffs, controversies surrounding the work of referees received a lot of scrutiny (Allen, 2019; Khurshudyan, 2019). While blaming referees for a team's loss is a ritual as old as the game of hockey itself, several experts agree that the 2019 playoffs have been riddled with decisions that changed the outcome of many games (Blackburn, 2019; Boylen, 2019). With those controversies in mind, referees may have expected their work to be scrutinized even more closely than usual during the 2020 playoffs. Fortunately for them, the COVID-19 pandemic took some firsthand eyewitnesses out of the equation. Indeed, the 2020 playoffs were held in empty arenas, with no crowds allowed. While this may give referees a break from real-time criticism of their decisions by fans, it provides an unprecedented opportunity to examine the crowd's contribution to one of the most notorious phenomena in sport, the home advantage.

Home advantage is a well-documented concept in sports (Courneya & Carron, 1992; Jamieson, 2010; Jones, 2018; Ribeiro et al., 2016). In hockey, playing at home provides advantages related to the rules. For example, players from the home team have the right to position themselves last when taking a face-off, a situation where the players attempt to gain control of the puck after it is dropped on the ice by the referee, which provides a technical upper-hand. In addition, home team coaches are allowed to make the last change of players after a stoppage of play, meaning they can choose to put on their desired players to match the opposing teams' choice of players. Essentially, these two rules allow the home team to observe the opposing team's strategy and adapt to it. In addition to rules-related benefits, playing in front of a

local crowd is acknowledged to increase the probability of winning (Boudreau et al., 2017; Inan, 2020; Nevill & Holder, 1999). While the presence of fans can positively impact home team players effort-level (Leard & Doyle, 2011), supporters can also have an impact on referee decisions (Brimberg & Hurley, 2009; Dohmen, 2008; Myers, 2014; Nevill et al., 2002). The crowd's behaviour is processed by referees as social information, which in turn can influence their decision-making (Plessner & Haar, 2006).

Since crowds are predominantly composed of local fans, their reactions are mainly in response to events that affect outcomes for the home team. Supporters conventionally cheer decisions benefitting the home team and boo decisions against them. Among the decisions that provoke the strongest reactions from crowds, are penalty calls. In hockey, penalties generate an important disadvantage for the offending team by offering a powerplay to the opposite team. In this situation, the penalized player must serve his penalty in the penalty box and his team cannot replace his presence on the ice. This means his team must perform with one less player for a period of time that is determined by the seriousness of the reprehensible act (2, 4 or 5 minutes). Crowd reactions to what fans consider illegal maneuvers may act as salient cues for referees, having the potential to influence their decisions on penalty calls (Nevill et al., 2002).

Referees' decisions leading to penalty calls, which then result in the home team benefitting from more powerplay opportunities than the visiting team, are considered a contributing factor to home advantage in the NHL (Beaudoin & Swartz, 2010; Brimberg & Hurley, 2009; Leard & Doyle, 2011). A plausible justification for that observation is the social pressure exerted by the home crowd. The presence of supporters in sporting events increase social pressure on referees, a phenomenon reputed to influence decisions in favor of the home team (Dawson & Dobson, 2010; Di Corrado et al., 2011; Dohmen, 2005; Dohmen & Sauermann, 2016; Scoppa, 2008). This pressure could lead referees to internalize crowd preferences in their

decision-making (Garicano et al., 2005). Since calling penalties is a decision that relies entirely on the judgment of a referee, analyzing this type of decision offers an opportunity to assess how the crowd may influence NHL referees' decision-making processes.

Although the effect of social pressure exerted by crowds on NHL referees' behavior has already been investigated (Beaudoin & Swartz, 2010; Brimberg & Hurley, 2009), completely isolating the impact of supporters is methodologically challenging. While controlling for different characteristics of the crowd is possible, totally removing supporters from NHL arenas is almost unimaginable. That is until the COVID-19 pandemic became problematic, and the implementation of social distancing measures forced the NHL to rethink the way their 2020 playoffs games were to be played. The NHL designated the cities of Edmonton and Toronto in Canada as hub cities. These hub cities allowed the NHL to control social distancing by holding the playoffs in empty arenas, creating an ecological environment where the impact of the home crowd was completely removed. This context offered a unique opportunity to analyze how referees make decisions in a context without direct social pressure from fans.

The purpose of this study is to investigate whether the presence of a crowd can influence the decision-making of NHL referees. This is done by taking advantage of the 2020 playoffs being played in front of empty seats, and comparing the number of penalties called by NHL referees against the home and away teams respectively in 2020 to the previous five years (2015 to 2019). Based on previous studies, we hypothesize that playing in crowded arenas leads to a reduction in penalties called against the home team and that this trend will disappear in a context where no local fans are present to exert social pressure on the referees. This would validate to a certain extent that the presence of crowds at professional ice hockey games do have an impact on the referees' decision-making, thus reinforcing the concept of home advantage in professional sports. To ensure the generalizability of the results, analysis of referee behavior in various

contexts is also presented. This was done by comparing 2020-2021 regular season games from the NHL, as well as from the Canadian Hockey League (CHL), the top junior ice hockey league in Canada, to games from their respective previous regular seasons. Those additional data sets offer the possibility to observe penalties awarded by referees in different contexts, when compared to the initial sample. Notably, the new data sets looked at games in which the stakes are minimized compared to the playoffs, at games played in the teams' local arenas rather than in neutral hub cities, and at games where the playing level and the degree of expertise of the referees are lower.

Methods and results

NHL playoffs data set

Our database includes all NHL playoff games since 2015, giving five years of data prior to the COVID-19 pandemic. As suggested by Lakens and Evers (2014), the 547 games played during this 6-year period are sufficient to achieve 80% power to observe a small effect size with an alpha of .05. This relatively short period also allows for consistency in the referees assigned to playoff games. We accessed NHL official web site (nhl.com) to obtain game reports from 2015 to 2019 games, played in front of a live audience, and from 2020 games, played without crowds. Due to the shortened season, the 2020 playoffs were held under an adjusted format. First, home and away status was granted as if the games were played in the normal playoffs format. This means that even though teams were not playing in their home cities, the higher ranked team was given home status for games 1, 2, 5 and 7 of the best-of-7 series, while the lower ranked team was given home status for games 3, 4 and 6.

Second, the number of eligible teams was increased, and the number of games was slightly modified. Eight teams were added to the 16 teams normally qualified for the playoffs. Of

the 24 teams that qualified in 2020, the top 4 of the eastern conference and the top 4 of the western conference got a free pass to the first round. However, those teams had to play a round-robin, where each team plays the other top three teams in their conference, to determine their final ranking for the rest of the playoffs. Those twelve round-robin games, 6 per conference, have not been considered in our data set since these games did not have the same stakes as a standard series where the losing team is eliminated. As the importance of these games is diminished, the behavior of the referees could be different compared to series where the stakes are higher. In addition, these games were played under regular season rules when it came to overtime. More specifically, instead of playing 5-on-5 until a team scores a goal like in other playoff games, round-robin overtime periods were played 3-on-3 for a maximum of 5 minutes, followed by a shootout if the tie persists. This rule restricts the number of minutes played in the event of a tie compared to normal playoffs, thus limiting the opportunity for referees to award penalties. Nevertheless, we conducted additional statistical analyses, which included the twelve round-robin games, confirming that their exclusion did not affect the pattern of findings.

The remaining 16 teams, representing the bottom eight teams in each conference, competed in a qualifying round consisting of best-of-five series, with the winners continuing on to the first round. Because best-of-five games lead to a team's elimination and have the same importance and rules as normal playoff games, we included these games in the data set. We excluded games involving the Edmonton Oilers and Toronto Maple Leafs from the data set, as those teams were the only ones playing in their home arenas. As with round-robin matches, excluding this data from statistical analyses did not change the results.

All penalties generating powerplays were collected from NHL official game summaries. We converted each game's powerplay opportunities for one team into the number of penalties taken by the opposing team. This ensured that only penalties creating a numerical disadvantage

for the offending team were retained. These are also known as penalty kills (PK). We rejected all penalties where players on both teams were being punished and neither team had a power play, specifically minor double penalties and fights. We also rejected 10-minute penalties for misconduct since these types of penalties do not generate powerplay opportunities for the opposing team. Descriptive statistics per year are reported in Tableau 1.

Tableau 1

Penalty kills per year during the NHL playoffs

Year	Games	Crowd	Away PK	Home PK	Mean Away PK	Mean Home PK
2015	89	Yes	276	244	3.10 (1.438)	2.74 (1.434)
2016	91	Yes	313	283	3.44 (1.551)	3.11 (1.567)
2017	87	Yes	296	250	3.40 (1.521)	2.87 (1.310)
2018	84	Yes	291	252	3.46 (1.766)	3.00 (1.575)
2019	87	Yes	306	243	3.52 (1.704)	2.79 (1.511)
2020	109	No	380	389	3.49 (1.772)	3.57 (1.612)

Note. Standard deviations are presented in parentheses.

Our data also includes dummy variables for each team in each year, to control the characteristics of the teams that may contribute to penalty decisions by NHL referees. For example, previous results suggest that the relative strength of teams may contribute to differences in playing style, resulting in more penalties being awarded to the weaker team (Beaudoin et al., 2016). Style of play not related to a team's relative strength, but rather related to the offensive system of play favored by the team, has also been identified as a key factor in home advantage in the National Basketball Association (Harris & Roebber, 2019), suggesting that the same effect could be observed in the NHL. Finally, dummy variables for game referees have been included to serve as fixed effects to control for individual propensity to call penalties. Individual differences exist in the tendency of referees to call more penalties to home or away teams (Beaudoin et al., 2016).

NHL playoffs model

To estimate the number of penalties called by referees, we used a Poisson generalized linear mixed model (GLMM) with a log link function. The model provides positive fitted values through the logarithmic link, and the Poisson distribution best fits the count data in a fixed period of time, in this case the number of penalties awarded in an NHL game (Coxe et al., 2009; Gardner et al., 1995). Standard Poisson regression was selected, since the model meets assumption of equidispersion and data were not zero inflated or zero truncated [28,29]. Our model is represented by:

$$\ln(\mu_{i,j}) = \text{Home}_{i,j} + \text{Crowd}_j + \text{Home}_{i,j} * \text{Crowd}_j + \text{Team}_k + \text{Ref1}_j + \text{Ref2}_j + \epsilon_{i,j,k} \quad (1)$$

where $\mu_{i,j}$ indicates the penalties awarded that resulted in penalty kills. Independent categorical variables are $\text{Home}_{i,j}$, representing the status of the team i (home = 0; away = 1), and Crowd_j indicating the presence or absence of an audience in match j (spectators = 0; no spectator = 1), with each match being counted twice, once for the away team and once for the home team (Endrich & Gesche, 2020; Garicano et al., 2005). The interaction between these variables is also included in the model to capture the effect on referee decisions of the status of teams, with and without crowd. Team_k fixed effects for each team during year k are included to control for style of play. Fixed effects for both referees were also included to control for differences in propensity to award penalties, with Ref1_j and Ref2_j . We use the package *stats* included in the software R to fit the GLMM (R Core Team, 2022), the package *rsq* to estimate the adjusted R^2 (Zhang, 2020), and the package *emmeans* (Lenth, 2020) to run post-hoc analyses.

NHL playoffs analysis and results

The model revealed a significant effect of the *Home*Crowd* interaction ($b = .17, p=.024$) for NHL playoff games. The results suggest that the likelihood of penalties being awarded by NHL referees differs depending on the status of the team and the presence or absence of spectators. Results of the NHL playoffs Poisson GLMM are presented in Tableau 2.

Tableau 2

Estimated regression parameters for NHL playoffs Poisson GLMM (n=1094)

	Estimate (b)	Std. error	z value	P-value	IRR Exp(b)
Intercept	1.388	0.196	7.089	<0.001	4.006
Home	-0.155	0.038	-4.022	<0.001	0.857
Crowd	0.077	0.269	0.285	0.775	1.080
Home : Crowd	0.186	0.083	2.254	0.024	1.205
Referees FE			Yes		
Team FE			Yes		
Adjusted R ²			0.113		

Note. Estimates are in log-odds from a Poisson regression. FE: Fixed effect. IRR: Incidence rate ratio.

Simultaneous pairwise comparisons using Tukey's HSD test indicated that in front of a crowd, penalties awarded to the away team were significantly higher than those awarded to the home team ($Z = 4.022, p<.001$). In contrast, no significant difference is observed when teams played in absence of the crowd ($Z = -.434, p=.973$). Referees also award significantly fewer penalties to the home team when they are supported by the crowd, compared to the home team playing in front of empty seats ($Z = -2.833, p=.024$). A similar trend can be observed when comparing the home team supported by a crowd with the opposing team playing in an empty arena ($Z = 2.462, p=.066$). However, there is no difference in the penalties awarded to the visiting team when playing in front of a crowd compared to an empty arena ($Z = -.475, p=.965$), nor in comparison to the home team playing without spectators ($Z = -.877, p=.817$). The model explains

11.3% of the variance in penalties called by NHL referees, which is considered a small effect size [34]. A small effect size was to be expected, as an unbiased referee should only consider the actions of the players and eliminate all external factors from his environment, i.e., the variables in the model. The main explanation for penalty calls by a referee remains the actual presence of illegal actions by the players, a variable that cannot be included in the model, as no objective data exists to that effect.

Tableau 3

Post-hoc results of the NHL model during playoffs (n=1094)

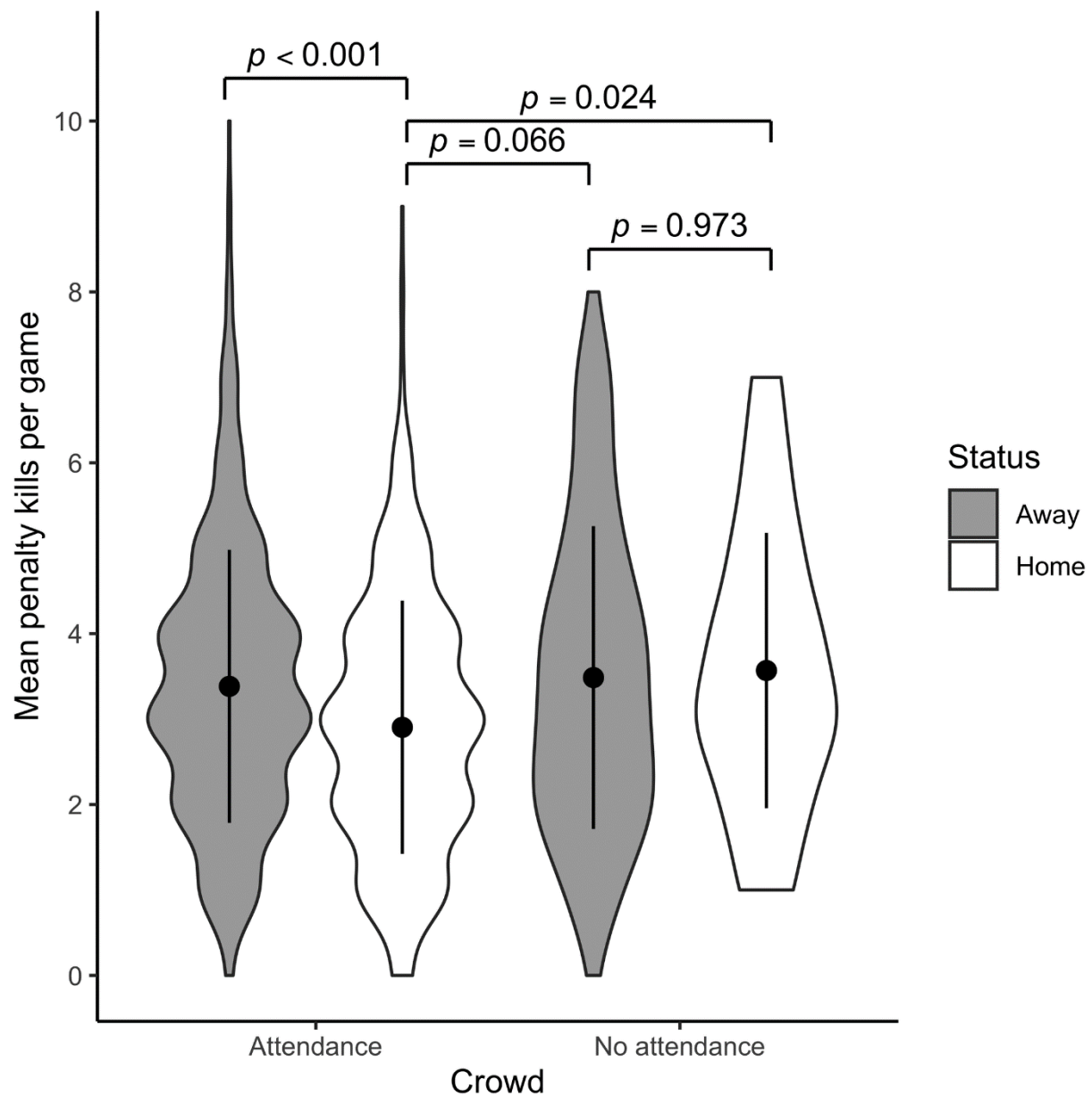
	Estimate (b)	Std. error	z ratio	P-value	IRR Exp(b)
Away Crowd – Away Empty	-0.037	0.077	-0.475	0.965	0.964
Away Crowd – Home Crowd	0.155	0.039	4.022	<0.001	1.168
Away Crowd – Home Empty	-0.068	0.078	-0.877	0.817	0.934
Away Empty – Home Crowd	0.191	0.078	2.462	0.066	1.210
Away Empty – Home Empty	-0.032	0.073	-0.434	0.973	0.969
Home Crowd – Home Empty	-0.222	0.079	-2.833	0.024	0.801

Note. Results are averaged over the levels of: *Ref1*, *Ref2*, *Team*. Estimates are in log-odds.

Figure 3 presents a violin plot with mean penalty kills (± 1 SD error bars) for away and home teams, in front of crowds before the pandemic and in empty arenas during the 2020 playoffs.

Figure 3

Mean Penalty kills per team with and without attendance during the NHL playoffs



Note. For the sake of simplicity, the raw data have been presented in the figure, even though the model results in Tableau 2 and Tableau 3 are in log-odds.

Our model adds evidence that referee decision-making plays a part in explaining the concept of home advantage by showing that in front of local supporters during the playoffs, NHL

referees tend to award less penalties to the home team compared to the away team, as well as both teams when the seats are empty. The results also support our hypothesis, which suggests that the home team is no more privileged when it comes to penalty calls when there is no home crowd in the building. Our model offers a good way to define the role of referees by controlling for the style of play of the teams, and the individual propensity of referees to award penalties.

Regular season data sets

To investigate the generalizability of our results, we used the same strategy and variables as for the NHL playoff data set to construct additional data sets for the NHL and CHL regular seasons. First, we aimed to see if crowd pressure had an impact on referee behavior in games where the stakes are lower and where games are played in the local teams' arenas rather than in neutral hub cities. Secondly, we wanted to test whether the influence of spectators was observable on referee decision-making in a league below the NHL level. The collection of data from games played in empty arenas was made possible by the extension of restrictive health measures related to the COVID-19 pandemic that forced some hockey leagues to begin the 2020-2021 season without crowds. This is the case for the NHL and CHL. In addition to having to play some games in front of empty seats, these leagues have also been compelled to shorten their regular schedule to avoid cancelling their seasons altogether. For the NHL, we once again used the official website nhl.com to collect the data. As for the Canadian Hockey League, we used the subdivisions of the official website (chl.ca) to acquire game summaries from the individual leagues that make up the CHL (whl.ca, ohl.ca, theqmjhl.ca).

We started by collecting data for the 2020-2021 NHL season, in which teams were schedule to play 56 games instead of the usual 82 games that are normally played in non-pandemic regular season. Unlike the playoffs, no hub city was designated, and teams played in

their respective hometown. This is closer to the usual format of NHL games, providing an opportunity to verify that the 2020 playoff results are not due to hub cities being neutral ground. The NHL's four divisions were reformatted with the goal of having teams exclusively play against other teams in their division. Given that government restrictions prohibited border crossings between the Canada and the United States, an exclusively Canadian division was created, along with three American divisions. All the games in the Canadian division were played in empty arenas. For the American division games, state-specific rules dictated whether or not spectators were present, with some games played in front of limited crowds ($n=310$). These games were removed from the sample because the very limited number of fans might not generate the same pressure on the referees as in an arena full of spectators. Nevertheless, additional statistical analyses that include these games were performed to confirm that removing them does not change the results.

Next, we compiled the data for the 2020-2021 CHL season. The CHL is made up of three leagues spread across Canada and the United States. The Western Hockey League (WHL), the Ontario Hockey League (OHL) and the Quebec Major Junior Hockey League (QMJHL) includes teams from nine provinces of Canada and four American states. As public health measures vary between provinces and states, the structure of the regular schedule also varied between the 3 leagues that make up the CHL. For starters, health measures caused the OHL to completely cancel its 2020-2021 season. For their part, the WHL and QMJHL opted for a shorter schedule. Since the teams that make up these two leagues are spread across several Canadian provinces and some American states, the format of the games varied according to the public health measures of each province and state. As a result, some of the games in the QMJHL were played in front of fans ($n=102$). These games were removed from the sample. Once again, additional statistical analysis including the data from these games confirms that their removal does not alter the

significance level of the results. Public health restrictions also led both leagues to hold some games in hub cities, similar to the NHL during its 2020 playoffs, while the rest of the games were held in the respective teams' hometowns.

In order to compare decisions taken by referees in this unusual context with decisions taken by referees during typical games played in front of crowds, we used the entirety of the previous season's games (2019-2020 regular season) for both the NHL and CHL. For the CHL, we excluded OHL games since it did not play games in 2020-2021 and meant it could not be used for comparison purposes. Although the 2019-2020 seasons were somewhat shortened due to the outbreak of the pandemic, the number of games in each league is far greater than the number of games suggested by Lakens and Evers (2014) to observe a small effect size. Descriptive statistics for the regular seasons are presented in Tableau 4.

Tableau 4

Penalty kills per year during the NHL and CHL regular seasons

Year	Games	Crowd	Away PK	Home PK	Mean Away PK	Mean Home PK
NHL						
2019-2020	1082	Yes	3368	3064	3.11 (1.484)	2.83 (1.353)
2020-2021	557	No	1634	1643	2.93 (1.484)	2.95 (1.452)
CHL						
2019-2020	1264	Yes	4979	4651	3.94 (1.791)	3.68 (1.648)
2020-2021	445	No	1660	1713	3.73 (1.727)	3.85 (1.830)

Note. Standard deviations are presented in parentheses.

Regular seasons model

The same Poisson GLMM, represented by Eq (1), was used to estimate the number of penalties called by NHL and CHL referees during regular seasons. The model for both data sets still meets assumption of equidispersion and data were not zero inflated or zero truncated (Coxe

et al., 2009; Gardner et al., 1995). The R packages *rsq* (Zhang, 2020) and *emmeans* (Lenth, 2020) were again used to estimate the adjusted R^2 and to run post-hoc analyses, respectively.

Regular seasons analysis and results

The model illustrates a significant effect of the *Home*Crowd* interaction for NHL regular season games ($b = .10, p = .010$) and for CHL regular season games ($b = .10, p = .018$). The results suggest that the probability of penalties being called by referees in these two leagues is modified by team status and the presence or absence of fans. The Poisson GLMM results for both leagues in regular season are presented in Tableau 5.

Tableau 5

Estimated regression parameters for NHL (n=3278) and CHL (n=3418) regular seasons Poisson GLMM

	Estimate (b)	Std. error	z value	P-value	IRR Exp(b)
NHL					
Intercept	1.033	0.111	9.437	<0.001	2.851
Home	-0.093	0.022	-4.179	<0.001	0.911
Crowd	-0.068	0.040	-0.114	0.909	0.995
Home : Crowd	0.101	0.043	2.571	0.010	1.116
Referees FE			Yes		
Team FE			Yes		
Adjusted R^2			0.081		
CHL					
Intercept	0.320	0.580	0.552	0.581	1.377
Home	-0.066	0.020	-3.251	0.001	0.936
Crowd	-0.035	0.034	-1.030	0.303	0.966
Home : Crowd	0.095	0.040	2.362	0.018	1.099
Referees FE			Yes		
Team FE			Yes		
Adjusted R^2			0.077		

Note. Estimates are in log-odds from a Poisson regression. FE: Fixed effect. IRR: Incidence rate ratio.

Simultaneous pairwise comparisons using Tukey's HSD test indicated that in front of a crowd, penalties awarded to the away team were significantly higher than those awarded to the home team in both the NHL ($Z = 3.729, p=.001$) and CHL ($Z = 3.251, p=.006$). Similarly to the 2020 NHL playoffs, these differences fade in regular season games where arenas are empty in both the NHL ($Z = -.230, p=.996$) and the CHL ($Z = -.820, p=.845$). In contrast to the NHL playoffs, no other significant differences are observed in the regular season in either the NHL or the CHL. The model for NHL explains 8.1% of the variance, while the model for CHL measures 7.7% of the variance. Again, a small effect size is observable, reinforcing the idea that external factors generate only a small influence on referee penalty calls.

Tableau 6

Post-hoc results of the NHL (n=3278) and CHL (n=3418) models during regular seasons

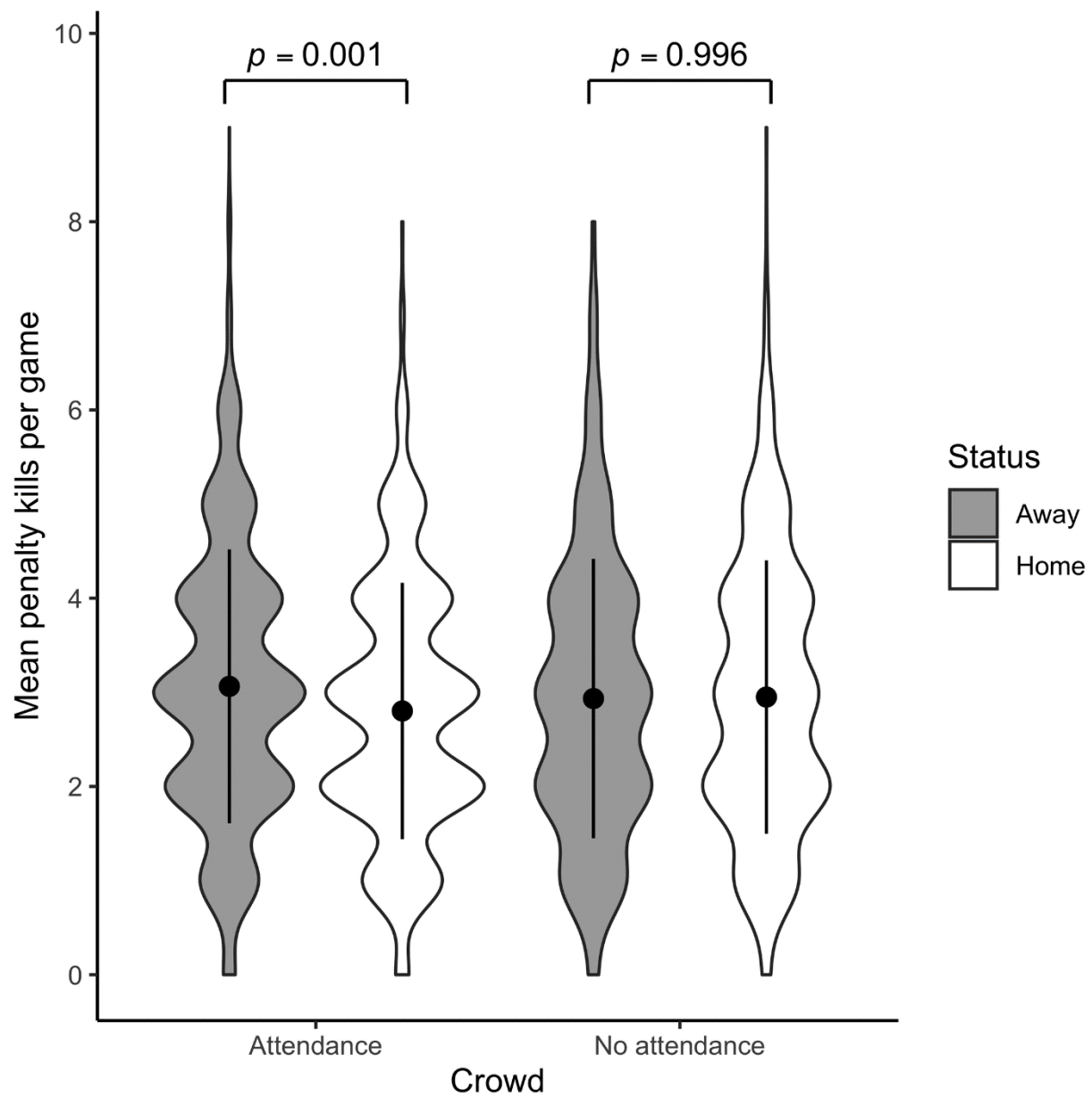
	Estimate (b)	Std. error	z ratio	P-value	IRR Exp(b)
NHL					
Away Crowd – Away Empty	0.068	0.031	2.181	0.129	1.070
Away Crowd – Home Crowd	0.093	0.025	3.729	0.001	1.098
Away Crowd – Home Empty	0.060	0.032	1.899	0.228	1.062
Away Empty – Home Crowd	0.025	0.032	0.790	0.859	1.025
Away Empty – Home Empty	-0.008	0.035	-0.230	0.996	0.992
Home Crowd – Home Empty	-0.033	0.032	-1.037	0.728	0.967
CHL					
Away Crowd – Away Empty	0.035	0.034	1.030	0.732	1.036
Away Crowd – Home Crowd	0.066	0.020	3.251	0.006	1.069
Away Crowd – Home Empty	0.007	0.034	0.199	0.997	1.007
Away Empty – Home Crowd	0.031	0.034	0.918	0.795	1.032
Away Empty – Home Empty	-0.028	0.035	-0.820	0.845	0.972
Home Crowd – Home Empty	-0.060	0.034	-1.755	0.236	0.942

Note. Results are averaged over the levels of: *Ref1*, *Ref2*, *Team*. Estimates are in log-odds.

Figures 4 and 5 present violin plots with mean penalty kills (± 1 SD error bars) for away and home teams, in front of crowds before the pandemic and in empty arenas during the 2020-2021 regular seasons.

Figure 4

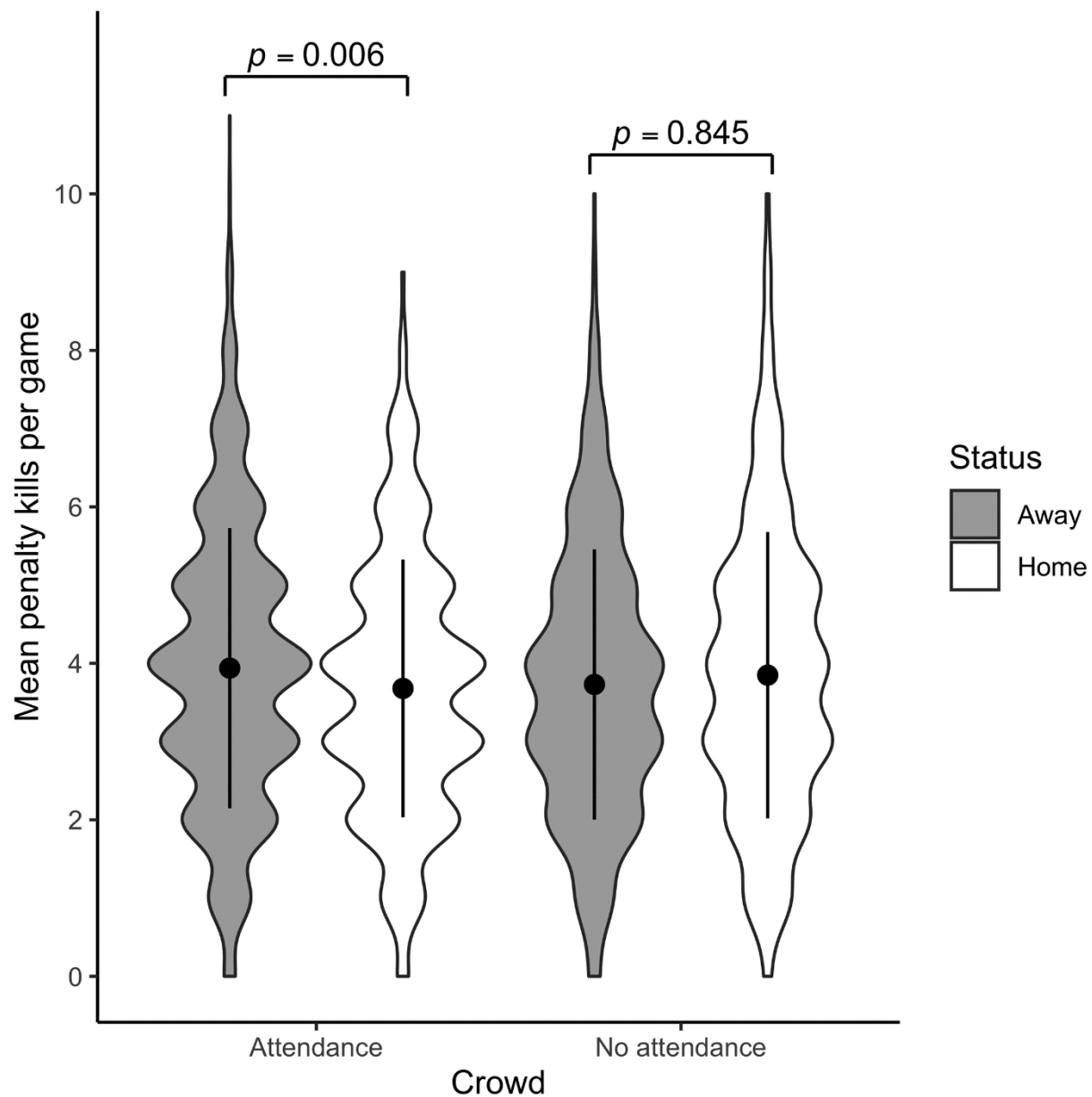
Mean Penalty kills per team with and without attendance during the NHL regular season



Note. For the sake of simplicity, the raw data have been presented in the figure, even though the model results in Tableau 5 and Tableau 6 are in log-odds.

Figure 5

Mean Penalty kills per team with and without attendance during the CHL regular season



Note. For the sake of simplicity, the raw data have been presented in the figure, even though the model results in Tableau 5 and Tableau 6 are in log-odds.

Consistent with the results stemming from the 2020 NHL playoffs, NHL and CHL referees, during regular season games, award the home team significantly fewer penalties on

average compared to the away team. The results again corroborate our hypothesis that the home team is no longer favored when it comes to penalty calls in situations where there are no local fans in the arena.

Discussion

Home advantage is a well-established concept in professional sport and the influence of the home crowd on referee decision-making is a factor that is assumed to contribute to this phenomenon. Using a unique ecological context generated by the COVID-19 pandemic, we found evidence of the effect of crowds on NHL referees' decisions during the playoffs. Our observations can also be generalized to different contexts, i.e., regular season and lower-level league. These findings add to existing scientific evidence supporting the involvement of sports referees in home advantage (Bryson et al., 2021; Cueva, 2020; Endrich & Gesche, 2020; Pettersson-Lidbom & Priks, 2010; Scoppa, 2021; Sors et al., 2020) and reinforce the existence of this phenomenon in elite ice hockey. The results also allow us to better understand the decision-making processes of ice hockey referees, who have the potential to change the outcome of games.

Our findings provide new evidence linking referees' decision-making with the concept of home advantage in elite ice hockey. More precisely, our model suggests that the home team receives less penalties when they play in front of their home crowd compared to the opposing team. This advantage disappears when the games are played without spectators. These results suggest that referees are more lenient towards home players when they are playing in front of their local fans.

The social pressure exerted by fans on referees is a plausible explanation for their behavior (Dawson & Dobson, 2010; Di Corrado et al., 2011; Dohmen, 2005; Dohmen & Sauermann, 2016; Scoppa, 2008). Crowd pressure seems to have an impact on referees in

different ice hockey leagues, as well as in competitive settings with varying stakes. However, the behavior of referees seems to be affected differently depending on the context. In the regular season, despite the difference in penalties called between the home and visiting teams when the crowd is present, it is difficult to determine whether the referees are calling more penalties on the visiting team, fewer penalties on the home team, or a combination of both. During the playoffs, a more specific direction of bias towards the home team is identifiable. Consistent with our results, crowd pressure during the playoffs can be observed by a decrease in the number of penalties awarded by the referees to the home team, a phenomenon that can be identified as an omission bias (Spranca et al., 1991). This bias would be more prevalent in critical moments, such as playoffs, where referees are known to try to limit their impact on the flow of the game (Moskowitz & Wertheim, 2011). Moreover, referees who display omission biases might be recognized as being fairer than referees who exhibit other types of biases. Indeed, the omission of penalties imposed on the home team can be considered less intrusive or damaging to the visiting team than an act of commission, such as an increase in the penalties imposed on the latter (Moskowitz & Wertheim, 2011). An excellent example of the perceived immorality of a commission call is the 2021 dismissal of Tim Peel, an NHL referee, who was caught admitting that he was trying to balance out the number of penalties awarded during the game by giving a questionable penalty to one of the teams (NHL Public Relations, 2021). To this day, there has never been a dismissal of an NHL referee who willingly omitted to call a penalty, or as they say in sports jargon, “swallowed his whistle” (Moskowitz & Wertheim, 2011). However, these examples raise the need to be cautious in interpreting an omission bias among NHL referees during the playoffs. While, as mentioned above, the greater stakes of playoff games may be sufficient to explain the omission bias, it is also important to consider that the referees selected to officiate these games are those whom the league considers their best. These referees can be

perceived as being the best precisely because they tend to omit penalties to the home team, rather than awarding more penalties to the opposing team, which is considered more morally acceptable. In this case, the presence of the omission bias during the playoffs could be due to the selection of referees by NHL executives, rather than the higher stakes of the games being officiated. It would be relevant to assess this in future studies.

Given that the results are based on a unique ecological environment, the number of games played without spectators remains low compared to the data available for usual games played in front of a crowd. Also, since playoff games played without an audience are limited to the NHL in 2020, our findings in playoffs are not generalizable beyond a reasonable doubt to other ice hockey leagues. As far as we know, the NHL is the only league to have played playoffs in 2020. In 2021, the professional leagues that have held playoffs have done so with a small number of games or in front of a limited number of fans. In the CHL, only the QMJHL has held its playoffs in an adjusted format, limiting the number of games in empty arenas to 44, which does not provide enough statistical power for analysis. Also, our sample is limited to NHL and CHL referees, meaning that other referees in minor and major hockey leagues may behave differently. Furthermore, given that these leagues are played in North America, referees may be influenced differently than those in other continents. Since the style of refereeing may be influenced by cultural factors, it could be conceivable to expect different results from European referees, for example. Our project is limited to the NHL and CHL, and more research needs to be conducted in different leagues and countries before we can generalize our results to the broad population of ice hockey referees.

However, our study has the obvious strength of presenting an ecological design made possible by the COVID-19 pandemic. For the very first time in studies relating to elite referee behavior, the local crowd and its possible impact on decision-making was physically eliminated

from the equation. This offers a significant advantage over designs where audience influence had to be statistically controlled. The statistical model used also adds strength to our results, whereas previous studies supporting the influence of the public on referees' decision-making, mainly those made in soccer, have relied on ordinary linear regression models that does not fit the count data as well as Poisson regression (Coxe et al., 2009). Moreover, our study examined the behavior of referees working for the world's best professional ice hockey league, as well as for one of the top junior hockey leagues. These referees are considered to be amongst the best in their discipline and the least likely to be influenced by factors outside of the game itself. The fact that our results present an impact of the crowd on elite referees' decision-making suggests that this impact should also be present, and possibly greater, in less experienced or capable referees. Finally, our study exposes the effect of crowds on referee decisions in a variety of contexts. Social pressure from the crowd influences referees in more than one league, in leagues of different levels, as well as in games with different stakes. Our research project illustrates that this impact disappears with the absence of crowds, both in neutral hub cities and in hometown arenas.

Conclusion

Home advantage can be explained in part by the social pressure exerted by home crowds on referees. A more exhaustive knowledge of which decisions are more predisposed to be influenced by the presence of crowds could help to better train referees, enabling them to deal with the existence of decision biases. With the advancement of technology, the work of referees is now highly scrutinized and improving their ability to make impartial decisions on every call can have a positive impact on both the length of their careers and the reputation of the league itself.

Acknowledgments

We would like to thank our collaborators, Jason Lacelle and Francis Gingras, for reviewing the text and providing constructive comments.

CHAPITRE III

Article 2: Verbal Aggressions Against Major League Baseball Umpires Affect Their Decision Making

Joël Guérette^{1,2}, Caroline Blais¹, and Daniel Fiset¹

¹Département de psychoéducation et de psychologie, Université du Québec en Outaouais; and

²École interdisciplinaire de la santé, Université du Québec en Outaouais

Abstract

Excessively criticizing a perceived unfair decision is considered to be common behavior among people seeking to restore fairness. However, the effectiveness of this strategy remains unclear. Using an ecological environment where excessive criticism is rampant—Major League Baseball—we assess the impact of verbal aggression on subsequent home-plate umpire decision making during the 2010 to 2019 seasons ($N = 153,255$ pitches). Results suggest a two-sided benefit of resorting to verbal abuse. After being excessively criticized, home-plate umpires ($N = 110$ adults, employed in the United States) were less likely to call strikes to batters from the complaining team and more prone to call strikes to batters on the opposing team. A series of additional analyses lead us to reject an alternative hypothesis, namely that umpires, after ejecting the aggressor, seek to compensate for the negative consequences brought on by the loss of a teammate. Rather, our findings support the hypothesis that, under certain conditions, verbal aggression may offer an advantage to complainants.

Statement of Relevance

The use of aggressive behavior to resolve conflicts is often considered socially unacceptable. Nevertheless, when people feel that they are victims of injustice, they tend to react aggressively in an attempt to restore justice. In this research, we used a context in which perceived injustices are numerous and real—professional baseball games in the United States—to test the effectiveness of verbal aggression toward officials. We found that players and coaches who yell at officials influence decisions of those officials to their team’s advantage in a bidirectional way. After being verbally abused, the officials call fewer strikes to that team’s batters and more to those of the opposing team. These findings must be considered when formulating strategies to mitigate human-to-human aggression. Recognizing the potential impact of aggression on decision makers raises questions about the ethical considerations surrounding its use, particularly in situations where individuals perceive themselves as victims of injustice.

Introduction

Witnessing a decision that we perceive as unfair often generates strong negative emotions, which can impair one's own decision-making and generate suboptimal responses (Blanchette & Richards, 2010; Fehr & Gächter, 2002; Hopfensitz & Reuben, 2009; Kenning & Plassmann, 2005; Lerner et al., 2015; Loewenstein & Lerner, 2003; Naqvi et al., 2006; Pillutla & Murnighan, 1996). Punishing the unjust person is a common way of responding, even if this behavior is likely to result in costly consequences (Henrich et al., 2006; Larrick et al., 2011; Nelissen, 2008; Scip et al., 2014; Seymour et al., 2007; Turillo et al., 2002; Xiao & Houser, 2005). In human relationships, punishment may take the form of verbal aggression (Baron & Richardson, 1994). This behavior is a way of expressing the magnitude of the perceived injustice, aiming to convince people to modulate their subsequent decisions to restore fairness (Canary et al., 1996; Crockett et al., 2014; Folger & Cropanzano, 2001; Greenberg & Cohen, 2014).

At first glance, aggressively criticizing people to change their future decisions to one's advantage may seem ineffective. Indeed, there is strong evidence that verbal aggression causes psychological pain, embarrassment, and suffering that damage the victim's self-esteem (Hamilton, 2012; Infante & Rancer, 1996; Perloff, 2020; Rancer & Avtgis, 2006). These negative emotions alter the relationship with the aggressor and reduce the likelihood that the victim will comply with the aggressor's viewpoint (Gass & Seiter, 2018; Guerin, 2003; Perloff, 2020; Rancer & Avtgis, 2006). On the other hand, research on social-information processing suggests that people may seek to avoid social punishment and may, after being aggressively criticized, seek to conform to their abuser's opinion to avoid further sanctions (Germar et al., 2016; Pryor et al., 2019; Rilling et al., 2008; Rilling & Sanfey, 2011; Sanfey, 2007; Schiebener & Brand, 2015; Starcke & Brand, 2012; Theriault et al., 2021; Wasylyshyn et al., 2018). This theoretical

incongruity is problematic for predicting the impact that excessive criticism of individuals may have on subsequent decisions. In an attempt to shed light on this contradiction, we propose to use an ecological environment where verbal aggression following perceived injustice is omnipresent—that is, baseball games.

Verbally abusing in-game officials, more commonly referred to as *umpires*, is almost a tradition in baseball (DeLong, 2017; Smith, 1983). Although pervasive, these aggressive behaviors are normally subject to an automatic punishment: The offender is ejected from the game. Although it is challenging to measure all verbal interactions that occur in a baseball game, those involving excessive verbal abuse are objectively measurable in Major League Baseball (MLB). They are associated with an ejection, which is recorded in the game statistics. (See Supplemental Material S1 in the Supplemental Material available online for an explanation of the use of ejection as an indicator of verbal aggression). Because many unacceptable behaviors can lead to an ejection, the game statistics provide the reason for the sanction, confirming whether or not the offense was verbal abuse of an umpire.

Much of the excessive criticism MLB umpires receive is related to strike-zone decisions (DeLong, 2017). In baseball, the strike zone is defined as an intangible area delimited by the width of the home plate and the height of the batter. Specifically, the zone height ranges from the midpoint between the batter's shoulders and the top of his pants to a point below his kneecap (see Fig. S1 in the Supplemental Material). One of the tasks of the home-plate umpire is to determine whether pitches directed at a batter travel through the strike zone when the batter does not swing. This decision should be objective; any pitch that touches the strike zone should be called a strike, whereas any other pitch should be called a ball. However, the task is extremely challenging because the speed of the pitch, which can exceed 100 mph, requires a high level of information processing (MacMahon & Starkes, 2008). The difficulty of the task is confirmed by

the camera system installed in all MLB stadiums. This system accurately measures the location of each pitch, allowing the umpires' performance to be evaluated (Mills, 2017). According to the pitch-tracking system, MLB umpires make the wrong call over 10% of the time. When umpires make mistakes, they give one team an advantage over the other. Each pitch classified as a ball gives a distinct advantage to a batter, because accumulating four balls results in the batter being walked (i.e., he may go to first base automatically). Conversely, each strike represents a disadvantage to batters, because accumulating three strikes results in the batter being called out. It is not surprising, then, that baseball players and coaches regularly criticize home-plate umpires' decisions that negatively affect their teams.

Thanks to baseball and its use of frequent criticism, its numerous statistics, and the technology that accompanies it, it is possible to measure in objective ways verbal aggression against umpires and its impact on subsequent strike-zone decisions. The purpose of our study is therefore to explore the impact of excessive criticism on MLB umpires' decisions related to the strike zone. We argue that umpires change their decision-making after being excessively criticized. Given the theoretical background, it is unclear whether this behavior change will provide an advantage or disadvantage to the aggressor's team, also referred to as the *in-group*. To strengthen the primary hypothesis that verbal abuse alters behavior, we also conducted a series of supplementary analyses to determine whether a change in behavior would be better explained by the alternative hypothesis that umpires are trying to compensate for the negative consequences to a team of losing a teammate to ejection. Tableau 7 provides a summary of the seven models and expected outcomes that would support the effect of excessive criticism, as well as those that would support an attempt to counterbalance the consequences of an ejection.

Tableau 7

Summary of analysis from all seven models and their anticipated results according to the primary (excessive criticism) and alternative (compensate for ejection) hypotheses

Models	Factors	Participants	Multiple ejections	Excessive criticism advantage	Ejection advantage	Full model	Reduced model
<i>Primary analysis</i>							
Model 1	N/A	Players, Coaches	Yes	Before $\neq$ After	Before $\neq$ After	N/A	Ejection*Team + FE _{1,2} + RE
<i>Additional analysis</i>							
Model 2	Ejection Reason	Players, Coaches	Yes	SZ = NSZ $\neq$ ABTO	SZ = NSZ = ABTO	Ejection*Team*Reason + FE _{1,2} + RE	Ejection*Team + FE _{1,2} + RE
Model 3	Role of the ejected individual	Players, Coaches	No	P = C	P > C	Ejection*Team*Role + FE _{1,2} + RE	Ejection*Team + FE _{1,2} + RE
Model 4	All-Star status	Players	No	AS = NAS	AS > NAS	Ejection*Team*All-Star + FE ₁ + RE	Ejection*Team + FE ₁ + RE
Model 5	All-Star appearances	Players	No	AS+ = AS	AS+ > AS	Ejection*Team*N _{All-Star} + FE ₁ + RE	Ejection*Team + FE ₁ + RE
Model 6	Number of seasons	Players	No	High = Low	High > Low	Ejection*Team*N _{Seasons} + FE ₁ + RE	Ejection*Team + FE ₁ + RE
Model 7	WAR	Players	No	High = Low	High > Low	Ejection*Team*WAR + FE ₁ + RE	Ejection*Team + FE ₁ + RE

Note: Before: Before ejection; After: After ejection; SZ: Strike zone related criticism; NSZ: Non-strike zone related criticism; ABTO: Aggressive behavior towards opponent; P: Players; C: Coaches; AS: All-Star; NAS: Not All-Star; WAR: Wins above replacement. FE₁: Fixed effects for baseline strike zone probability (BCSprob); FE₂: Fixed effects for home team; RE: Random effects for umpires, catchers, batters, and pitchers.

Method

Data acquisition

The data and code utilized in the project are available at <https://osf.io/eu6v9/>. The study was not preregistered. The project relies exclusively on publicly available information and does not require ethical approval to meet ethical guidelines and legal requirements. The host institution's ethics committee has confirmed that this is the case, following receipt of an ethics declaration.

We collected data on ejections that occurred in MLB from the 2010 season through the 2019 season, listed on the Retrosheet website (<https://retrosheet.org>). The studied time frame begins in the year following implementation of the umpire-assessment system called Zone Evaluation (Mills, 2017) and ends before the commencement of spectator-free matches caused by the COVID-19 pandemic. Each ejection is accompanied by various information, including match identification, the name of the person ejected, his role on the team (player or coach), the name of the umpire who issued the ejection, and the reason for the ejection. Only the ejections performed by home-plate umpires were considered, given that it is the excessive criticism of their work and their subsequent decision-making that interested us. Then the ejections were separated into three categories—those related to criticism of the strike zone, those related to criticism of other home-plate-umpire decisions, and those related to aggressive behavior directed at the other team (i.e., behavior that did not involve criticism of umpire decisions). Reasons for ejection that may fall into more than one category (e.g., bench jockeying) and that did not indicate that the decision being criticized was solely that of the home-plate umpire (e.g., video replays) were removed from the sample. Games in which members of each of the two competing teams were ejected have also been removed from the data set. However, games in which multiple members of the same team

were ejected for the same reason were retained. The list of reasons reported for the ejections that we retained in the final sample is presented in Table S1 in the Supplemental Material.

To complete the data on ejections, we also collected information to define characteristics of the ejected players. *Wins above replacement* (WAR), an all-encompassing measure that summarizes a player's contributions to his team's success during a season, was obtained via the Fangraphs website (<https://fangraphs.com>), whereas the number of years played in the MLB and the number of All-Star Game appearances were acquired via the Baseball Reference website (<https://baseball-reference.com>). This supplementary information helps to identify the status of players relative to other players within MLB, either in terms of performance, reputation, or experience.

From games containing ejections retained, we exported the pitch-by-pitch data using the *baseballr* package in R (Petti, 2021). The data is collected by a pitch-tracking system installed in all MLB stadiums (PITCHf/x from 2010 to 2016; StatCast from 2017 to 2019) and provides accurate information about the ball's trajectory. In particular, the system identifies the horizontal and vertical position of the ball when it reaches the two-dimensional area that defines the strike zone, which makes it possible to evaluate the accuracy of the decision-making of home-plate umpires.

We then verified that the ejections appearing in the list obtained on Fangraphs are also listed in the pitch-by-pitch data. Five games with missing ejections were removed from the sample. We also removed ejections that occurred late in the game—situations in which the home-plate umpire did not have to rule on balls and strikes after the ejection ($n = 9$). The final sample is composed of 1,204 ejections listed from 1,001 games.

Then the pitch-by-pitch data was filtered to include only those pitches for which the batter does not swing, to ensure that the umpire was forced to make a decision (i.e., calling a ball or a

strike). This represented 157,415 pitches. Pitches with missing data for variables to be included in the models were removed from the sample. Specifically, 1,163 pitches were eliminated because their spatial position was missing, one pitch was eliminated because the count given prior to the pitch was impossible (4 balls and 3 strikes), and two additional pitches were eliminated because the identity of the pitcher and batter did not appear. We also removed all pitches that required virtually no subjective decision-making on the part of the umpire to be called balls—that is, those that were extremely far from the strike zone and thus evident calls to make ($n = 2,994$). As previously proposed by Mills (2014), these pitches are those that hit the ground before reaching home plate, crossed home plate at a height of more than 5 feet, or were more than 2 feet away on either side of home plate. The final sample consisted of 153,255 pitches.

Finally, we collected additional data to estimate the baseline probability that each pitch in our sample would be called a strike. To do so, we once again used the *baseballr* package to export pitches thrown during all games outside of our sample during the 2010 to 2019 seasons. These data were used to calculate the probability of a pitch being called a strike as a function of spatial position, pitcher's and batter's handedness, batter's count, season, and pitch type. Nineteen matches could not be exported because of the recording format of the source (see Supplemental Table S2). Again, pitches were removed if the umpire did not have to make a decision (because the batter swung at the ball), if data were missing, or if the pitches were obvious balls (because they were very far from the strike zone). The final data set consisted of 3,444,443 pitches.

Analyses

Baseline called-strike probabilities

According to the MLB rulebook, the only reason an umpire should call a pitch a strike is its spatial location, which is available from the pitch's horizontal and vertical coordinates recorded by the pitch-tracking system. However, previous studies have demonstrated a nonlinear relationship between the position of the pitch and the probability of its being called a strike (Albert et al., 2017; Deshpande & Wyner, 2017; Mills, 2014). To address this issue, we propose to use a variable, *baseline called-strike probability* (BCSprob), that is an estimate of the probability of a pitch being called a strike in our logistic mixed-effects models rather than the coordinates of the pitches (Deshpande & Wyner, 2017; Marchi et al., 2018; Mills & Salaga, 2018). Relying on previous studies (Albert et al., 2017; Deshpande & Wyner, 2017; Marchi et al., 2018; Mills, 2014; Mills & Salaga, 2018), we estimated a generalized additive model (GAM) with logistic link from the pitches that occurred in games without ejections to create this variable, using the following equation:

$$g(\mu_i) = f_{bp}(X_i, Y_i) + \text{Count}_i + \text{Year}_i + \text{PitchType}_i + \varepsilon_i$$

Eq (1).

The dependent variable is defined as $g(\mu_i)$, with $g(\cdot)$ representing a logit link function and μ_i representing the mean of called strikes (ball = 0; strike = 1). The GAM included the unknown smooth function $f_{bp}(X_i, Y_i)$, estimated for combined horizontal (X_i) and vertical (Y_i) pitch location. The smooth function is indexed by bp , because different surfaces are estimated for each combination of batter handedness b and pitcher handedness p . Linear additive effects are included for the 12 possible counts (Count_i), the 10 seasons (Year_i), and the 15 types of pitches (PitchType_i) registered in the database. The GAM model was fitted by penalized iteratively

reweighted least squares and by generalized cross-validation for smoothness selection (Wood et al., 2015), using the *mgcv* package (Wood, 2021) in R (R Core Team, 2022). Results are presented in Supplemental Table S3.

The model allowed us to create the variable *baseline called strike probability* (BCSprob) by estimating the probability of each pitch in our ejection-games sample to be called a strike. By identifying a 50% or greater probability that a pitch is a strike as a successful prediction when the umpire's actual decision is a strike, our GAMs correctly predicted 90.58% of the umpires' strike-zone decisions in the pitch-by-pitch sample. We also assessed the performance of the model via a receiver operating characteristic (ROC) curve, built with the *ROCR* package (Sing et al., 2005). The area under the curve is 96.98%, which is considered outstanding (Hosmer et al., 2013). The ROC curve is available in Supplemental Figure S2. Supplemental Figures S3 to S6 show the baseline probability of an umpire calling a strike, based on certain batter counts: the starting count (0 balls/0 strikes) and the two most extreme counts (0 balls/2 strikes, 3 balls/0 strikes). Supplemental Figures S7 to S10 present the baseline probability during different seasons (i.e., the most distant in time, 2010 and 2019). Finally, Supplemental Figures S11 to S14 illustrate the baseline probability for different pitch types (i.e., pitches with a different ball trajectory—the four-seam fastball and the curveball).

Logistic mixed-effects models

To estimate the influence of an MLB team member's ejection on strike calls, we created mixed-effects logistic models. First, to assess the impact of an ejection for excessive criticism of the strike-zone-related decisions, we created a model with an interaction between the ejection event and the team that lost a player or coach to ejection (described as the *ejected team*, for conciseness). Then, for all our additional analyses, we used a reduced model containing this

interaction to compare with six improved models, each including a different additional variable. The additional variable was included in a three-way interaction with ejection event and ejected team. Including a variable in a triple interaction also adds the fixed effect of that variable, as well as any possible double interactions between the three variables in the triple interaction.

Our reduced models are represented by these expressions:

$$\begin{aligned} \text{logit}(p_{ijklm}) = & \beta_0 + \beta_1 \text{Ejection}_i + \beta_2 \text{Team}_{li} + \beta_3 \text{Ejection}_i * \text{Team}_{li} + \beta_4 \text{BCSprob}_i \\ & + \beta_5 \text{Home}_{li} + b_1 \text{Umpire}_{ji} + b_2 \text{Catcher}_{ki} + b_3 \text{Batter}_{li} + b_4 \text{Pitcher}_{mi} + \varepsilon_i \\ \text{Umpire}_j \sim & N(0, \sigma^2_{\text{Umpire}}) \\ \text{Catcher}_k \sim & N(0, \sigma^2_{\text{Catcher}}) \\ \text{Batter}_l \sim & N(0, \sigma^2_{\text{Batter}}) \\ \text{Pitcher}_m \sim & N(0, \sigma^2_{\text{Pitcher}}) \end{aligned}$$

Eq (2).

Our full models are represented by these expressions:

$$\begin{aligned} \text{logit}(p_{ijklm}) = & \beta_0 + \beta_1 \text{Ejection}_i + \beta_2 \text{Team}_{li} + \beta_3 \text{Condition}_i + \beta_4 \text{Ejection}_i * \text{Team}_{li} \\ & + \beta_5 \text{Ejection}_i * \text{Condition}_i + \beta_6 \text{Team}_{li} * \text{Condition}_i + \beta_7 \text{Ejection}_i * \text{Team}_{li} \\ & * \text{Condition}_i + \beta_8 \text{BCSprob}_i + \beta_9 \text{Home}_{li} + b_1 \text{Umpire}_{ji} + b_2 \text{Catcher}_{ki} \\ & + b_3 \text{Batter}_{li} + b_4 \text{Pitcher}_{mi} + \varepsilon_i \\ \text{Umpire}_j \sim & N(0, \sigma^2_{\text{Umpire}}) \\ \text{Catcher}_k \sim & N(0, \sigma^2_{\text{Catcher}}) \\ \text{Batter}_l \sim & N(0, \sigma^2_{\text{Batter}}) \\ \text{Pitcher}_m \sim & N(0, \sigma^2_{\text{Pitcher}}) \end{aligned}$$

Eq (3).

where $\text{logit}(p_{ijklm})$ designates the log odds of a pitch i , thrown by a pitcher m to a batter l , which is caught by a catcher k and judged by an umpire j , being called a strike (ball = 0; strike = 1). The model includes a dummy variable for the moment when the pitch was thrown, named Ejection_i (before ejection = 0; after ejection = 1). For games with more than one ejection, the cut-off was established at the moment of the first ejection. We created another dummy for the team to which the batter who receives the pitch belongs, named Team_{li} (nonejected team = 0; ejected team = 1). Condition_i symbolizes the characteristics for which we want to evaluate the impact. We generated six different conditions to test six different models. For the first three conditions, we created dummy variables, 1, 2, and 3, respectively, for the reason for the ejection (strike-zone criticism = 0; nonstrike-zone criticism = 1; aggressive behavior toward opposite team = 2); ejected individuals' role in their team (players = 0; coaches or managers = 1); and all-star status of players (never been invited to the All-Star Game = 0; at least one invitation to the All-Star Game = 1). For other conditions, we used continuous variables (4, 5, 6) for player-specific characteristics, respectively the number of All-Star Game participations, the number of years played in the MLB, and players' performance level measured by WAR. The continuous variables were all centered before being included in the models. BCSprob_i is the baseline probability that a pitch thrown at a specific location, during a specific count, in a specific year, and with a specific trajectory, will be called a strike. This variable was preliminarily predicted from Equation 1. Dummy variable Home_{li} (away team = 0; home team = 1) was included to control for the home-field advantage of the team at bat. Random effects for umpires, catchers, batters, and pitchers were added to the model to control for the inconsistency of their behavior (Archsmith et al., 2018; Barr et al., 2013; Deshpande & Wyner, 2017; Huang & Hsu, 2020).

Models were fitted with the *lme4* package (Bates et al., 2014) in R, using maximum likelihood from adaptive Gauss-Hermite quadrature and *bobyqa* optimizer to avoid convergence

issues (Pinheiro & Chao, 2006; Powell, 2009). For all models, we first tested whether the mixed-effects minimal baseline model explained significantly more variance than a model without random variables by applying a likelihood ratio test ($p < .05$ for all models; Bolker et al., 2009). Standardized parameters are presented in the Results section and were obtained by fitting the model on a standardized version of the data set. We used Wald approximation to compute 95% confidence intervals (CIs) and p values. For post hoc analysis, we used the *emmeans* package (Lenth, 2021).

Results

Tableau 8 presents descriptive statistics for pitches judged by umpires in games with excessive strike-zone-criticism ejections, by ejection event (before, after), and batter's team (ejected, nonejected).

Tableau 8

Descriptive statistics and accuracy of MLB umpires by team, before and after ejection for excessive strike zone criticism

	Ejected Team				Non-ejected Team			
	Total strikes	Missed call	Correct call	False alarm	Total strikes	Missed call	Correct call	False alarm
<i>Before ejection</i>								
Number of called pitches	13,993	1,166	35,658	1,421	12,745	1,330	35,860	1,147
Percentage	36.59	3.05	93.23	3.72	33.24	3.47	93.54	2.99
<i>After ejection</i>								
Number of called pitches	7,768	788	21,651	725	7,857	735	21,826	753
Percentage	33.53	3.40	93.47	3.13	33.70	3.15	93.62	3.23
Percentage difference	-3.06	0.35	0.24	-0.59	0.46	-0.32	0.08	0.24

Note: The total number of pitches is obtained by the sum of missed calls, correct calls, and false alarms. The accuracy of the umpires' decisions was calculated based on the spatial position of the ball relative to the strike zone described in MLB rules. The coordinates of the pitches in the database are given by the center of the ball. Since a pitch should be considered a strike if any part of the ball touches the strike zone, we have added the radius of the ball on each side of home plate, so that a pitch that grazes the strike zone should be called a strike to represent a good call. The top and bottom of the strike zone, which considers the size and position of the batter, were taken from the database for each pitch. Missed calls are pitches classified as balls even if they hit the strike zone; false alarms are pitches classified as strikes even if they did not hit the strike zone. Missed calls benefit the batter; false alarms disadvantage the batter.

Primary analysis

Model 1: excessive criticism impact

To determine the impact of verbal abuse on baseball umpires' decision-making, we fitted a linear mixed-effects model, including an interaction between the ejection event and the team to which the ejected person belonged (Equation 2). The aim of the interaction was to test whether a potential change in umpire behavior after the umpire ejected someone is similar for both teams or whether it provides an advantage to one team over the other following excessive criticism related to strike-zone decisions ($n = 123,060$). Tableau 9 displays the estimates of Model 1. Inclusion of the interaction ($\beta = -0.57$, 95% CI = $[-0.67, -0.47]$, $p < .001$) significantly improved the model compared to a reduced model with ejection event and ejected team as simple effects only, likelihood ratio test $\chi^2(1) = 137.69$, $p < .001$. Tukey's honestly significant difference (HSD) post hoc analysis (Supplemental Table S4) revealed a significant decrease in the likelihood that pitches directed at the offending team's players are called strikes after an ejection (odds ratio, or $OR = 0.697$, 95% CI = $[0.636, 0.762]$, $p < .001$), giving them an advantage. In contrast, there is an increase in the probability that pitches directed at players on the team that did not experience the ejection are called strikes after ejection ($OR = 1.232$, 95% CI = $[1.124, 1.350]$, $p < 0.001$), once again offering an advantage to the aggressor's team. Interestingly, the team that excessively criticizes the umpire's decisions is indeed at a disadvantage compared to the opposing team before the ejection ($OR = 1.479$, 95% CI = $[1.366, 1.602]$, $p < .001$). The opposite effect is observed after the ejection, when players from the complaining team receive significantly fewer strikes than their opponents ($OR = 0.837$, 95% CI = $[0.756, 0.926]$, $p < .001$). The results of Model 1

indicate that the use of verbal abuse toward umpires could provide a two-sided advantage for a team by not only decreasing the likelihood of their batters' receiving strikes but also increasing the likelihood of opposing batters receiving strikes. Figure 6 and Supplemental Figure S15 illustrate this dual advantage.

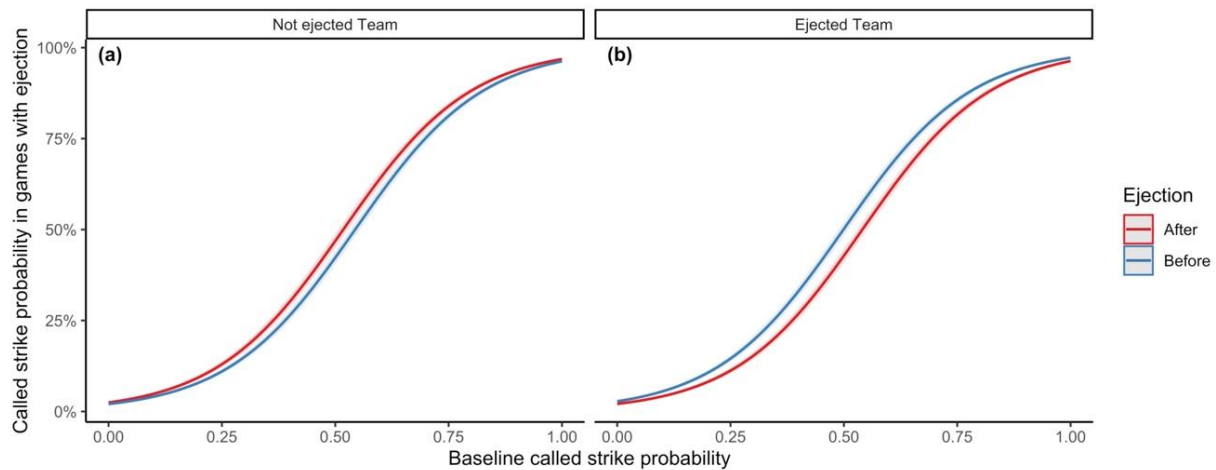
Tableau 9

Impact of ejections for strike zone related criticisms on strike decisions (model 1)

<i>Predictors</i>	<i>Odds Ratios</i>	<i>95% CI</i>	<i>p</i>
(Intercept)	0.02	[0.02-0.02]	<0.001
BCSprob	1217.22	[1134.66-1305.79]	<0.001
Home Team	1.07	[1.02-1.12]	0.006
Ejection	1.44	[1.34-1.54]	<0.001
Team	1.20	[1.11-1.29]	<0.001
Ejection*Team	0.57	[0.51-0.62]	<0.001
<i>Random effects</i>			
σ^2		3.29	
τ_{00} umpire		0.05	
τ_{00} catcher		0.07	
τ_{00} batter		0.03	
τ_{00} pitcher		0.07	
ICC		0.06	
N _{Umpire}		107	
N _{Catcher}		209	
N _{Batter}		1,641	
N _{Pitcher}		1,357	
Observations		123,060	
Marginal R ² /		0.699 / 0.717	
Conditional R ²			

Figure 6

Called strike probability in games with ejections related to excessive strike zone criticism as a function of the baseline probability observed in games without ejections and the team ejected ($n = 123,060$)



Our first model suggested that verbal aggression may cause umpires to alter their decision-making to benefit the aggressive individual's team. However, it is imperative to consider that the verbal aggression is coupled with a decision by the umpires—the ejection of the offender—that is detrimental to his team. To validate that the umpires' behavioral change is due to the verbal aggression and that it does not represent an attempt to compensate for the negative consequence of their decision to eject a team member, we performed a series of additional analyses.

Additional analysis

We conducted additional investigations to determine whether the change in the umpires' behavior was better explained as an attempt to compensate for the decision to eject someone or as a reaction to being verbally abused. Six models were created to test different variables that might

support this hypothesis. We first analyzed the impact on the strike zone of ejections for different types of misconduct (Model 2); then we tested whether the status and role of the ejected individuals could also influence the decision-making of umpires. If player status is indeed a factor, the umpires should consider the magnitude of the team's loss and adjust their decisions differently depending on the value of the individual ejected. Indeed, players have a greater impact on the outcome of a match than the coaches do (Berri et al., 2009; Demiralp et al., 2012). Also, the more talented, recognized, and experienced players are more valuable to their team than the less talented, recognized, or experienced players (Adler, 1985; Kim & King, 2014; Mills, 2014; Swaab et al., 2014). To eliminate these assumptions and reinforce the hypothesis that verbal aggression influences umpire behavior, we assessed the impact of the ejected individual on the basis of characteristics such as role (Model 3), all-star status (Model 4), number of All-Star Game appearances (Model 5), years of experience (Model 6), and performance level (Model 7).

Model 2: impact of ejection reason

As a first step in testing the likelihood that umpires are trying to balance the negative consequences of an ejection, we analyzed the impact on the strike zone of ejections for different types of misconduct (i.e., for excessively criticizing the strike zone, for excessively criticizing another decision, or for physically attacking a member of the opposing team; $n = 153,255$). If all types of ejections result in an advantage for the ejected team, the impact of the loss of a team member would seem to have a greater influence on umpire behavior than the verbal abuse itself.

The reason for the ejection was added to the interaction between the ejection event and the ejected team, thus generating a triple interaction. Supplemental Table S5 presents the estimates of Model 2. Including the triple interaction significantly improved the model compared to a reduced model with only an interaction between ejection event and the team with an ejected

player or coach, likelihood ratio test $\chi^2(8) = 70.76, p < .001$. When compared to ejections for excessive criticism of strike-zone-related decisions, the interaction effect of non-strike-zone verbal abuse ($\beta = 0.47, 95\% \text{ CI} = [0.20, 0.74], p < .001$) and that of aggressive behavior against the other team ($\beta = 0.45, 95\% \text{ CI} = [0.12, 0.78], p = .008$) on ejection event by the team with an ejected player or coach were statistically significant. Post hoc analysis (Supplemental Table S6) using Tukey's HSD pairwise comparisons suggests that, for the group of individuals ejected because of strike-zone criticism, there is a significant decrease in the likelihood that pitches directed at the offending team are called strikes after the ejection ($OR = 0.698, 95\% \text{ CI} = [0.638, 0.763], p < .001$), which provides benefits. Conversely, we observe an increase in the probability that pitches directed at opposing players were called strikes after the ejection ($OR = 1.238, 95\% \text{ CI} = [1.130, 1.356], p < .001$), again resulting in an advantage for the offending team.

Furthermore, significant differences are observed between both teams before and after the ejection, suggesting that the umpires were disfavoring the offending team over the opposing team before the ejection ($OR = 1.473, 95\% \text{ CI} = [1.362, 1.594], p < .001$), and favoring the offending team over the opposing team after the ejection ($OR = 0.831, 95\% \text{ CI} = [0.751, 0.919], p < .001$).

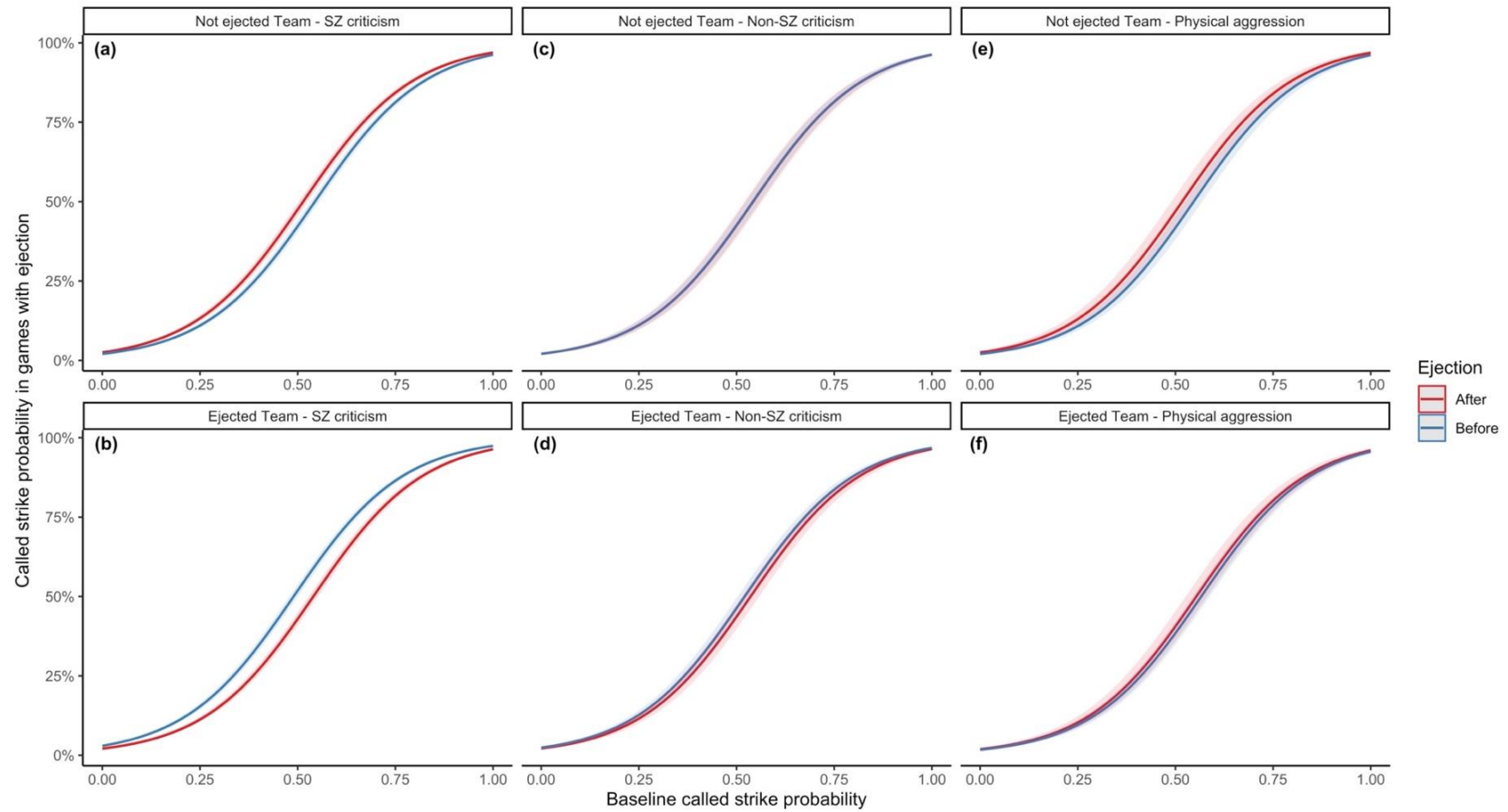
Unsurprisingly, these results are consistent with those of the first model, because they are based on the same data. For groups added to the analysis, no significant difference was observed in the probability of strike calls after an ejection, both for individuals ejected for verbal aggression unrelated to the strike zone ($OR = 0.897, 95\% \text{ CI} = [0.713, 1.129], p = .6170$ for the offending team; $OR = 0.995, 95\% \text{ CI} = [0.785, 1.262], p = 1.000$ for the opposing team) and for aggressive behavior against the other team ($OR = 1.092, 95\% \text{ CI} = [0.814, 1.465], p = .8673$ for the offending team; $OR = 1.234, 95\% \text{ CI} = [0.916, 1.662], p = .2659$ for the opposing team).

Likewise, there is no significant differences between teams before and after the ejection, suggesting that umpires, when not excessively criticized regarding the strike zone, are not

favoring any team with respect to the strike zone before the ejection ($OR = 1.163$, 95% CI = [0.947, 1.427], $p = .2339$ for non-strike-zone verbal abuse; $OR = 0.872$, 95% CI = [0.677, 1.122], $p = .5017$ for aggression directed at opponent) and after the ejection ($OR = 1.047$, 95% CI = [0.805, 1.362], $p = .9690$ for non-strike-zone verbal abuse; $OR = 0.771$, 95% CI = [0.550, 1.081], $p = .1976$ for aggression directed at opponent). These results, presented visually in Figure 7 and Supplemental Figure S16, suggest that verbal abuse influences umpires' strike-zone decisions to the advantage of the offending team, but only when the verbal abuse is specific to the strike zone. Ejections for verbal abuse not specific to the strike zone and for aggressive behavior toward opponents do not provide this advantage to the aggressive individual's team. For this reason, future models will only include data from ejections that occurred because of excessive criticism of strike-zone-related decisions, as with the first model.

Figure 7

Called strike probability in games with ejections as a function of the baseline probability observed in games without ejections and the reason for the ejection ($n = 153,255$)



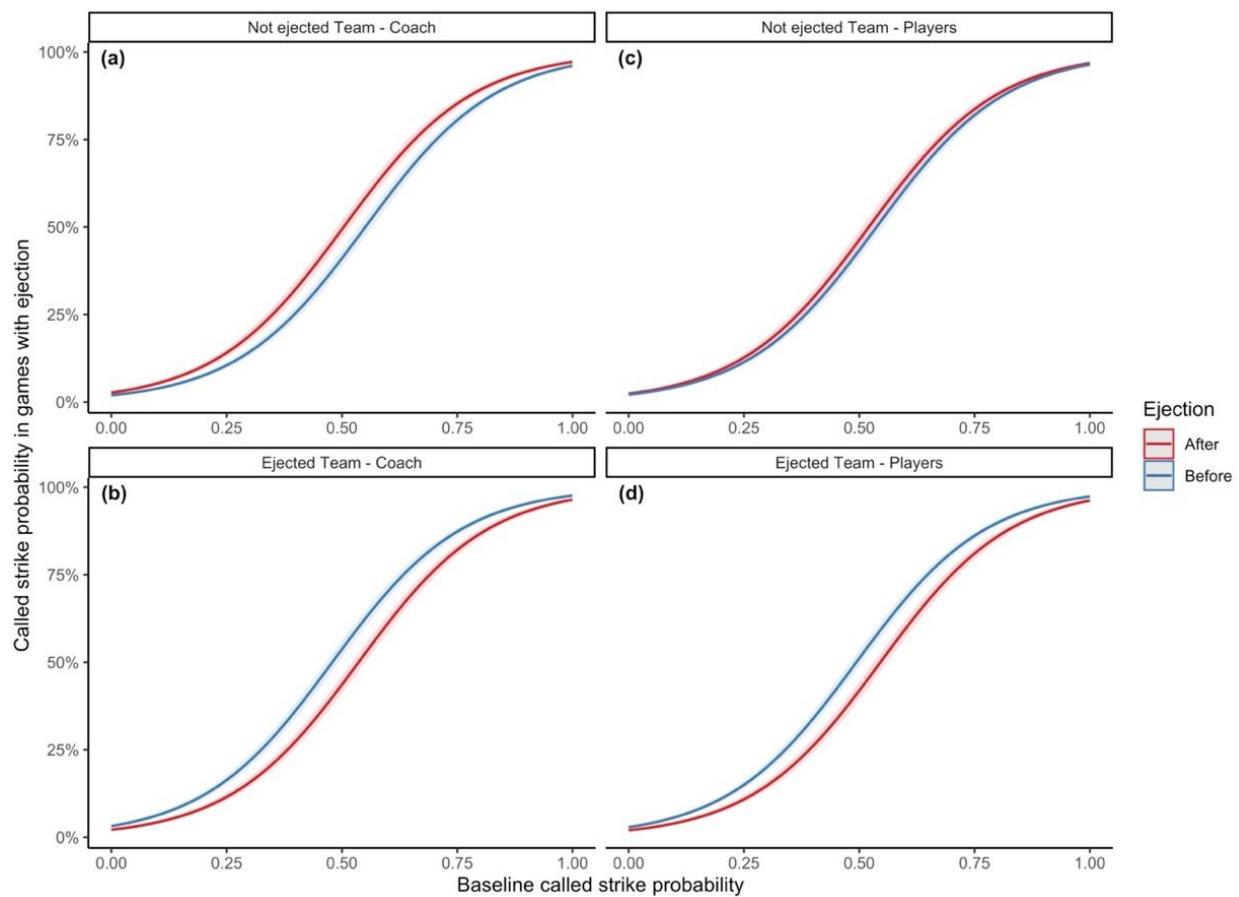
Model 3: impact of the ejected individual's role

We used the strike-zone excessive-criticism ejection data to generate a third model to evaluate the impact of the ejected person's role (player, coach) on umpire decisions. Players are directly involved in the action and represent the most important members of a team. Coaches also play a crucial role, but they are not as important as players. To isolate the impact of the unique role of the ejected person, we included only pitches from games in which a single ejection for verbal abuse related to the strike zone was recorded ($n = 103,913$). Supplemental Table S7 presents the estimates of Model 3, again fitted with and without a triple interaction, this time between the ejection event, the team that lost someone to ejection (ejected team), and the role of the ejected individual. The triple interaction significantly improved the model compared to the model with only an interaction between ejection event and the team that suffered the ejection, likelihood ratio test $\chi^2(4) = 13.56, p = .009$. Compared to ejected players, the interaction effect of coach ejection on the ejection event per ejected team was statistically significant ($\beta = 0.78, 95\% \text{ CI} = [0.63, 0.96], p = .021$). Post hoc analysis (Supplemental Table S8) illustrates, for all roles of ejected individuals, a significant decrease in the likelihood of pitches being considered strikes after an ejection for the offending team (players: $OR = 0.666, 95\% \text{ CI} = [0.577, 0.768], p < .001$; coaches: $OR = 0.576, 95\% \text{ CI} = [0.430, 0.771], p < .001$). For the opponent, significant increases in the likelihood of pitches being called strikes after the ejection are present for coaches' ejections ($OR = 1.398, 95\% \text{ CI} = [1.211, 1.613], p < .001$) but not for players' ejections ($OR = 1.128, 95\% \text{ CI} = [0.982, 1.295], p < .1138$). Comparisons between teams indicate that umpires tend to disadvantage the offending team over the opposing team prior to the ejection for both roles (players: $OR = 1.367, 95\% \text{ CI} = [1.221, 1.531], p < .001$; coaches: $OR = 1.665, 95\% \text{ CI} = [1.455, 1.905], p < .001$). After the ejection of a player or coach, umpire decisions significantly

disadvantage the opposing team compared to the offending team (players: $OR = 0.835$, 95% CI = [0.713, 0.979], $p = .0185$; coaches: $OR = 0.835$, 95% CI = [0.713, 0.979], $p < .001$). Results, presented in Figure 8 and Supplemental Figure S17, differ only in the significant increase in the number of strikes called on the team that did not experience an ejection (nonejected team) after a coach's ejection, which is not the case for a player's ejection. These findings do not support the assumption that the umpire tries to compensate for the importance of the ejected person, and they are even counterintuitive to what one would expect under this hypothesis.

Figure 8

Called strike probability in games with ejections as a function of the baseline probability observed in games without ejections and the team member ejected ($n = 103,913$)



Model 4: impact of the ejected player's all-star status

To measure the impact of ejected-player characteristics (Models 4 to 7), only pitches from games in which a single player was ejected for excessively criticizing strike-zone-related decisions were included in the analyses ($n = 57,223$). Unlike Models 1, 2 and 3, home-team advantage was not included in these models because this control variable was no longer statistically significant and decreased the fit of the models (Supplemental Table S9). Supplemental Table S10 presents the estimates of Models 4 to 7, fitted with a triple interaction between the ejection event, the team that lost someone to ejection, and the measured player feature. All those models were compared to the same reduced model including a significant interaction between the ejection event and the team that lost someone to ejection ($\beta = -0.49$, 95% CI = $[-0.63, -0.34]$, $p < .001$).

All-star status is the first feature to be included (Model 4). Each year, the best players in the league compete in the All-Star Game. The rules for selection change from year to year, but the selection of participating players is usually made by a combination of votes from fans, players, managers, the commissioner's office, or media representatives. The players selected are among the elite in their sport. Any player who has participated in the All-Star Game at least once is considered an all-star in our sample. This factor (*not an all-star* = 0; *all-star* = 1) was included in an interaction with the ejection event and the ejected team (Supplemental Fig. S18). The triple interaction ($\beta = -0.02$, 95% CI = $[-0.17, 0.13]$, $p = .779$) did not significantly improve the model compared to the model with only an interaction between ejection event and the team that lost someone to ejection, likelihood ratio test $\chi^2(4) = 1.63$, $p = .803$. The single effect of all-star status is also not significant ($\beta = 0.03$, 95% CI = $[-0.04, 0.09]$, $p = .400$).

Model 5: impact of the ejected player's All-Star Game appearances

After the all-star status, we tested the number of appearances at All-Star Games (Model 5). The best players will usually participate in several All-Star Games, in contrast to players who may have had a particularly good season but are not among the elite on a regular basis. Players with more appearances at this event are widely recognized as being more successful. We included the number of All-Star Game appearances as a continuous variable in an interaction with the ejection event and the ejected team (Supplemental Fig. S19). Again, the triple interaction ($\beta = 0.03$, 95% CI = $[-0.11, 0.17]$, $p = .674$) did not significantly improve the model compared to the model with only an interaction between ejection event and the ejected team, likelihood ratio test $\chi^2(4) = 1.35$, $p = .852$. In addition, the effect of the single variable (i.e., all-star appearances) was not significant ($\beta = 0.03$, 95% CI = $[-0.04, 0.10]$, $p = .374$).

Model 6: ejected individual experience impact

Next, we included an interaction with the number of years of experience of the ejected players (Model 6). Usually, to stay in the best baseball league in the world, an individual's performance must be quite valuable to his team. Compared to players who cannot secure a regular spot on a team's roster or who are relatively new to MLB and have not yet proven themselves, players who have more years in the league have the status of established players. We included the number of years played in MLB as a continuous variable, again in an interaction with the ejection event and the ejected team (Supplemental Fig. S20). The triple interaction ($\beta = -0.06$, 95% CI = $[-0.20, 0.08]$, $p = .428$) did not significantly improve the model compared to the model with only an interaction between ejection event and ejected team, likelihood ratio test $\chi^2(4) = 3.53$, $p = .473$. The years-of-experience single effect was not significant ($\beta = 0.001$, 95% CI = $[-0.07, 0.07]$, $p = .976$).

Model 7: impact of the ejected individual's performance

Finally, we fitted Model 7 by including a triple interaction between the ejection event, the ejected team, and the performance level of ejected players as measured by WAR. As we indicated earlier, WAR is a measure that summarizes a player's contributions to his team in one all-encompassing statistic; it acts as a good performance index. The higher a player's WAR is, the more he contributes to his team's success. We included the player's WAR as a continuous variable (Supplemental Fig. S21). As with the previous player characteristics, the triple interaction ($\beta = -0.02$, 95% CI = $[-0.16, 0.12]$, $p = .783$) did not significantly improve the model compared to the baseline model, likelihood ratio test $\chi^2(4) = 2.23$, $p = .694$. Thus, the single effect of WAR is not significant either ($\beta = 0.03$, 95% CI = $[-0.03, 0.10]$, $p = .329$).

In summary, none of the four player characteristics added to the reduced model suggest an influence on umpire behavior, leading us to infer that umpires do not actively seek to compensate for the importance of an ejected player. This further solidifies our main hypothesis.

Discussion

Strong negative emotions generated by the perception of an unfair decision may lead people to become verbally aggressive. Our study indicates that although verbal aggression can have its costs—such as ejection from a baseball game—it can also have positive consequences for the in-group (here, the aggressor's team) by influencing the umpires' decision-making in a favorable manner. Our findings suggest that verbal aggression, under certain circumstances, can help restore what could be perceived as justice.

Umpires who experience verbal abuse tend to alter their decision-making to restore a perceived fairness. They do so by making decisions that benefit the team that expressed a sense of unfairness in a verbally abusive manner. This change in decision-making is observable only

when the object of criticism is specific, because ejections that are not due to strike-zone criticism do not generate an advantage for the offending team in terms of strike-zone calls. The umpires' change in behavior does not appear to be an attempt to compensate the team for the loss of the ejected individual. Indeed, the relative importance of the ejected person does not influence the decisions of the umpires who were overly criticized. Coaches' ejections can generate a slight additional benefit compared to players' ejections, even if they are less valuable to their team. Players of all skill, recognition, and experience levels generate similar benefits for their team when they are ejected. Thus, as shown in Tableau 10, our results both support the hypothesis that verbal aggression influences MLB umpires' decision-making and also undermine the alternative hypothesis that umpires seek to balance the negative consequences brought on by the loss of a team member.

Tableau 10

Summary of results from all seven models and their interpretation according to the primary and alternative hypotheses

Models	Factors	Participants	Multiple ejections	Excessive criticism advantage	Ejection advantage	Results	Interpretation
<i>Primary analysis</i>							
Model 1	N/A	Players, Coaches	Yes	Before $\neq$ After	Before $\neq$ After	Before $\neq$ After	Benefits of excessive criticism or ejection.
<i>Additional analysis</i>							
Model 2	Ejection Reason	Players, Coaches	Yes	SZ = NSZ $\neq$ ABTO	SZ = NSZ = ABTO	SZ $\neq$ NSZ = ABTO	No benefits of non-specific verbal abuse or aggression not directed towards the umpire.
Model 3	Role of the ejected individual	Players, Coaches	No	P = C	P > C	P < C	No benefits of being more valuable to the team.
Model 4	All-Star status	Players	No	AS = NAS	AS > NAS	AS = NAS	No All-Star status benefits.
Model 5	All-Star appearances	Players	No	AS+ = AS	AS+ > AS	AS+ = AS	No All-Star appearances benefits.
Model 6	Number of seasons	Players	No	High = Low	High > Low	High = Low	No experience benefits.
Model 7	WAR	Players	No	High = Low	High > Low	High = Low	No performance level benefits.

Note: Before: Before ejection; After: After ejection; SZ: Strike zone related criticism; NSZ: Non-strike zone related criticism; ABTO: Aggressive behavior towards opponent; P: Players; C: Coaches; AS: All-Star; NAS: Not All-Star; WAR: Wins above replacement. FE₁: Fixed effects for baseline strike zone probability (BCSprob); FE₂: Fixed effects for home team; RE: Random effects for umpires, catchers, batters, and pitchers.

Umpires are first and foremost human beings, which leads them to perceive verbal aggression as a significant social punishment. An attempt to prevent this form of sanction could explain the adjustment in their decision-making. Indeed, our findings support *social-information processing theory*, which suggests that aggressively criticized individuals seek to conform to their aggressor's viewpoint in order to avoid further punishment (Germar et al., 2016; Pryor et al., 2019; Rilling et al., 2008; Rilling & Sanfey, 2011; Sanfey, 2007; Schiebener & Brand, 2015; Starcke & Brand, 2012; Theriault et al., 2021; Wasylshyn et al., 2018). No statistical model supports the theory that verbal aggression alters the relationship with the aggressor and reduces the likelihood that the victim will conform to the aggressor's viewpoint (Gass & Seiter, 2018; Guerin, 2003; Perloff, 2020; Rancer & Avtgis, 2006). Therefore, our study adds to existing knowledge about the possible effects of using verbal aggression in response to perceived injustice.

Our results are derived from a specific ecological environment, which generates certain limitations that must be acknowledged. Among other things, we find that the perceived injustice of the verbal abuser is often justified, as the umpires' decisions prior to the ejection are truly disadvantageous to the complainant's team. Laboratory studies controlling for this factor will be required to determine whether verbal aggression loses its effectiveness when the alleged injustice is unfounded. Also, MLB umpires are considered to be the best in their discipline and are the most qualified to realize their mistakes (Mills, 2017). Lower-level umpires may be less able to realize the injustices caused by their decisions and may react differently. Unfortunately, we had to limit ourselves to MLB umpires, given the lack of available data at lower levels (e.g., the reasons for ejections at the MLB level are listed). This is not the case in other professional and minor leagues. Access to this data, combined with the pitch-by-pitch data already available for several lower-level leagues, would allow other researchers to generalize our results. At present, it

is prudent to exercise caution when extrapolating the results observed within our specific context. The findings may vary in other sporting disciplines, at different levels of competition, or when involving younger or less experienced officials.

Although an ecological environment generates certain limitations, it also offers undeniable strengths for our study. First, the verbal aggression toward umpires is genuine and stems from a real alleged injustice. Second, it is perceived by the umpires as truly unacceptable behavior, because they judge that the offending individual has crossed a line and deserves to be ejected from the game. Laboratory-controlled criticism may not have the same impact on the umpires, who may interpret and process the aggression differently. Third, our study relies on a significant number of observations, collected from a sophisticated and accurate pitch-tracking system. This data provides an objective measure with which we can assess umpires' decision-making. Also, with such statistical power, it can be assumed that the additional analyses that were undertaken would have provided evidence for the alternative hypothesis, if it had been well founded. Fourth, the statistical models controlled for the many factors known to influence MLB umpires' decision-making, allowing us to isolate the impact of excessive criticism. Our set of additional analyses allows for the inclusion of several variables that have the potential to support an alternative hypothesis—that is, that the umpire is influenced by his decision to eject someone. Results of these numerous analyses have added weight to the initial hypothesis that verbal aggression is responsible for changing umpire behavior (Supplemental Table S11).

Although our findings support the effectiveness of a behavior that is subject to reprimand, in this case verbal aggression, they provide valuable insight into the impact of such behavior on an authority figure's decision-making. In the context presented here, each baseball sequence must be judged individually and must not influence the umpires' future decisions. Once a person whose verbal behavior exceeds the acceptable limits is ejected, it is the umpire's duty to keep

issuing impartial decisions, which should in no way be biased by past events. Our results suggest that the decision-making of umpires is altered following verbal aggression, and they also suggest that there is an impact on the fairness of the game by giving an advantage to the team that previously failed to respect the rules. This bias in human decision-making contributes to the debate over the use of advanced technology to assist MLB home-plate umpires during games. A new policy introducing the use of the pitch-tracking system to support strike-zone decisions could not only help improve the fairness of the game but also reduce verbal aggressions against umpires.

Outside the sporting environment, other decision makers could benefit from our results to improve the fairness of their decisions (see Supplemental Material S2 for a generalization of the results to other authority figures). Decision makers need to be aware that after being aggressively criticized, their future decisions may become unfair if the decision maker becomes more lenient toward the aggressor's in-group. This could be the case regardless of whether the decision makers overturn the decision that has been excessively criticized. For example, police officers who should cite drivers for an infraction may opt not to; judges may grant reduced sentences to criminals; or business managers may accept employees' requests for time off or additional benefits. Education geared toward understanding the potential impact of excessive criticism on one's decision-making may be beneficial for people in positions of authority, especially to help them make informed, rational, and fair decisions. By limiting the advantage provided to the verbally aggressive individuals' in-group, decision makers could help decrease the frequency of aggressive behavior and unfair decisions. Our findings should therefore be used to better understand the impact of verbal aggression in order to decrease its use, not to encourage its spread.

CHAPITRE IV

Discussion générale

L'environnement sportif, particulièrement lorsqu'on s'intéresse au sport professionnel, est un canevas idéal pour étudier le comportement humain en contexte de compétition. En raison des gains potentiels pour les individus impliqués, il est possible de percevoir des interactions sociales qui transcendent souvent les normes sociales observables dans la vie courante. C'est notamment le cas des interactions impliquant les arbitres, véritables souffre-douleurs du domaine sportif (Bright et al., 2022), qui voient leurs décisions critiquées à répétition. En effet, puisque chaque décision prise par un arbitre se solde par un avantage pour un parti (un athlète ou son équipe) et un désavantage pour un autre (l'autre athlète ou l'autre équipe), il est évident que l'arbitre ne sera jamais en mesure de plaire à tout le monde.

Si le rôle de bouc émissaire que joue un arbitre peut avoir de nombreuses conséquences négatives (Brick et al., 2022; Downward et al., 2024), il permet aussi d'étudier l'impact des éléments du contexte sur la prise de décision de l'être humain. En effet, en étant soumis à l'influence du comportement des acteurs faisant partie du contexte sportif, c'est-à-dire les joueurs, entraîneurs et spectateurs, un arbitre est confronté aux forces sociales générées par le contexte dans lequel il évolue. C'est l'influence de ces forces sur la prise de décision des arbitres qui a été étudiée dans la présente thèse doctorale.

L'influence des forces sociales sur la prise de décision des arbitres

Pour mesurer l'influence des forces sociales sur la prise de décision des arbitres sportifs, deux projets de recherche distincts ont été réalisés. Les résultats de ces deux projets suggèrent que les acteurs impliqués dans le contexte sportif exercent une influence sur la prise de décision des officiels par leur présence et leur comportement.

Plus spécifiquement, dans le chapitre II, il a été soulevé que l'implication des spectateurs peut exercer un impact sur la prise décisionnelle des arbitres. Au hockey professionnel, la présence d'une foule contribue à ce que l'équipe locale soit avantagée lorsqu'on s'intéresse au nombre de pénalités décernées. Cet avantage disparaît lorsque la foule est retirée des amphithéâtres. Les arbitres semblent plus enclins à éviter les sanctions sociales transmises par la foule, en limitant le nombre de pénalités accordées à l'équipe locale plutôt qu'en augmentant le nombre de pénalités allouées à l'équipe adverse. Ce constat est du moins apparent en séries éliminatoires, lorsque l'enjeu est le plus important et la pression placée sur les arbitres est à son apogée.

Dans un même ordre d'idées, au chapitre III, il a été possible d'observer que les joueurs et les entraîneurs de baseball ont le potentiel d'influencer les décisions des arbitres au marbre en critiquant excessivement les décisions de ces derniers. Après avoir été victimes d'agressions verbales, comportements qui représentent une forme de sanction sociale, les arbitres rendent plus de décisions avantageuses à l'endroit de l'équipe de la personne agressive. Cet avantage ne semble pas lié au statut de la personne agressive, puisque le rôle, le niveau de performance et le niveau d'expérience de cette dernière ne modifient pas la réaction des arbitres au-delà de l'influence que procure l'agression verbale.

Les mécanismes explicatifs potentiels

La pression sociale

L'explication la plus simple pour appuyer ces résultats réside dans la pression sociale exercée par les joueurs, entraîneurs et spectateurs. Effectivement, lorsqu'on compare les résultats obtenus à la définition de la pression sociale, on constate l'exercice d'une influence sur une personne (l'arbitre) par une autre personne (un joueur ou un entraîneur) ou un groupe de

personnes (les spectateurs). Toutefois, comme mentionné précédemment, pratiquement toutes les interactions sociales peuvent respecter la définition de pression sociale, ne permettant qu'une compréhension peu approfondie des mécanismes responsables de la prise de décision de l'arbitre. C'est pourquoi se contenter d'expliquer les résultats obtenus par le phénomène de pression sociale se voudrait insatisfaisant et limiterait les retombées potentielles de la présente thèse doctorale.

En regard des résultats obtenus dans les deux projets de recherche inclus dans le document actuel, il est possible de suggérer que la pression sociale que subissent les arbitres semble particulièrement influente lorsqu'elle revêt une valence négative. Plus précisément, les résultats permettent de soulever l'hypothèse que les arbitres ont prioritairement tendance à éviter les sanctions sociales plutôt qu'à rechercher l'approbation sociale. C'est le cas au baseball où les arbitres modifient leur prise de décision après avoir été excessivement critiqués, possiblement pour éviter de futures sanctions. C'est également le cas au hockey, où les arbitres semblent plus enclins à omettre de pénaliser l'équipe locale et ainsi éviter les réactions négatives de la foule à leur endroit, plutôt qu'à se commettre et pénaliser l'équipe adverse pour obtenir l'approbation des partisans.

Considérant la possibilité que la valence des forces sociales en présence permette de mieux comprendre comment la pression sociale agit sur la prise de décision des arbitres en contexte sportif, la théorie du biais de négativité est proposée comme cadre théorique alternatif à l'étude du jugement et de la prise de décision en contexte sportif.

Le biais de négativité

De manière générale, les phénomènes négatifs sont plus saillants et plus diagnostiques que les phénomènes positifs, augmentant l'importance que l'être humain leur accorde

(Baumeister et al., 2001). Cette asymétrie permet d'observer des effets plus forts des informations négatives sur la perception, la mémoire et la prise de décision (Norris, 2021), qui sont des étapes de la séquence du traitement de l'information sociale utilisée par les arbitres.

La dominance générale des éléments négatifs sur les éléments positifs se nomme le biais de négativité (Rozin & Royzman, 2001). On retrouve notamment des démonstrations de l'existence du biais de négativité en provenance de méthodes comportementales et psychophysiologiques (Baker et al., 2020; Müller-Pinzler et al., 2019; Vanderhasselt et al., 2015), électrophysiologiques (Huang & Luo, 2006; Ito et al., 1998) et d'imagerie cérébrale (Ito et al., 2017). Bien que ce biais semble généralisable chez les humains, des différences individuelles sont néanmoins observables (Norris, 2021; Norris et al., 2011).

Le biais de négativité pour expliquer le comportement des arbitres sportifs

Liens avec la théorie

Le biais de négativité pourrait influencer le comportement des arbitres de nombreuses façons. D'abord, la rétroaction négative (ou feedback négatif), qui agit à titre de sanction sociale, peut amener les arbitres à apprendre plus rapidement qu'avec la rétroaction positive, une forme de récompense sociale (Baumeister et al., 2001). Les réactions négatives des participants aux événements sportifs pourraient ainsi être associées plus rapidement aux décisions indésirables que les réactions positives aux décisions souhaitables. De plus, au niveau émotionnel, les éléments négatifs sont reconnus pour générer des réactions plus fortes, qui durent plus longtemps et qui demeurent plus saillantes dans l'esprit des gens (Brickman et al., 1978). L'effet plus prononcé de l'information négative rendrait l'évitement des émotions négatives plus important que la recherche d'émotions positives (Baumeister et al., 1994). Pour les arbitres, cela pourrait être observable par une motivation supérieure à agir de manière à éviter la critique

comparativement à la motivation à se comporter pour être félicité (Baumeister et al., 2001). Cette asymétrie motivationnelle pourrait être à la source de décisions biaisées, fournissant un avantage à ceux qui critiquent plus régulièrement ou plus fortement le travail des arbitres.

Le biais de négativité pourrait également agir à titre de protection pour les arbitres. En effet, le fait d'accorder plus d'importance aux éléments négatifs que positifs est un mécanisme d'adaptation que l'être humain a développé pour affronter son environnement physique et social (Baumeister et al., 2001; Cacioppo & Berntson, 1994; Kahneman, 2011; Rozin & Royzman, 2001). Même si l'adaptation des arbitres à leur environnement n'est pas une question de survie, la réputation qu'ils se forgent peut grandement influencer la pérennité de leur carrière. En s'adaptant principalement aux critiques négatives, les arbitres pourraient éviter d'être victime d'une mauvaise réputation, laquelle est plus facile à développer et plus difficile à perdre qu'une bonne réputation (Baumeister et al., 2001).

Liens avec les résultats des projets de recherche

Considérant les nombreux avantages pour les arbitres de prioriser le traitement de l'information négative, il est pertinent de s'intéresser à la valence des forces sociales qui l'accable, plutôt que de parler de pression sociale au sens large. Les études présentées précédemment pourraient servir à tracer le pont entre la théorie sur le biais de négativité et des études expérimentales visant à vérifier l'influence de l'information négative sur la prise de décision des arbitres. En effet, les projets de recherche proposés dans la thèse doctorale ici présentée permettent d'obtenir des premiers éléments de réponse quant à l'impact de l'information négative sur la prise de décision des arbitres, sans toutefois permettre de produire l'ensemble du portrait.

Dans l'étude sur les arbitres de la Ligue Nationale de Hockey, il a été permis d'observer que les officiels ont tendance à avantager l'équipe locale en ce qui concerne le nombre de pénalités accordées. Un tel avantage détient le potentiel de prendre deux configurations différentes : 1 - une diminution du nombre de pénalités accordée à l'équipe locale, afin d'éviter les huées de la foule, lesquelles représentent des sanctions sociales; 2 - une augmentation du nombre de pénalités accordées à l'équipe visiteuse, permettant de s'attirer les applaudissements de la foule, lesquels agissent à titre de récompenses sociales. Les résultats de l'article publié dans *PloS One* supportent, dans les conditions où la pression est la plus forte, l'hypothèse selon laquelle les arbitres chercheraient à éviter les sanctions sociales plus qu'ils ne souhaitent obtenir l'approbation sociale. Les conclusions de cet article offrent certains arguments pour appuyer l'hypothèse du biais de négativité.

L'étude sur l'impact de la violence verbale à l'endroit des arbitres de baseball a aussi permis de supporter l'influence de l'information négative dans le processus décisionnel d'un arbitre. Transmises sous forme d'agression verbale, les forces sociales négatives en provenance des joueurs et entraîneurs génèrent des bénéfices pour l'équipe de l'agresseur. Ce constat était important à vérifier, puisque la théorie sur l'agression verbale ne permettait pas d'assumer hors de tout doute la direction de l'influence de l'information négative sur la prise de décision des arbitres. En effet, des théories contradictoires proposaient des répercussions potentiellement différentes en lien avec le comportement de l'arbitre, menant à une certaine incertitude concernant l'effet de l'information négative. D'une part, certains auteurs suggéraient que l'agression verbale pouvait entraîner des conséquences psychologiques négatives chez la victime (Hamilton, 2012; Infante & Rancer, 1996; Perloff, 2020; Rancer & Avtgis, 2006), ce qui diminuerait la probabilité que cette dernière se conforme au point de vue de son agresseur (Gass & Seiter, 2018; Guerin, 2003; Perloff, 2020; Rancer & Avtgis, 2006). D'autre part, les chercheurs

s'intéressant au traitement de l'information sociale suggéraient plutôt que la victime de l'agression verbale pourrait se conformer au point de vue de son agresseur dans le but d'éviter les futures sanctions sociales (Germar et al., 2016; Pryor et al., 2019; Rilling et al., 2008; Rilling & Sanfey, 2011; Sanfey, 2007; Schiebener & Brand, 2015; Starcke & Brand, 2012; Theriault et al., 2021; Wasylyshyn et al., 2018). Les résultats du projet présenté au chapitre III supportent un effet significatif de l'information négative, lequel avantage l'équipe de la personne qui utilise cette information de manière excessive, suggérant que les arbitres cherchent à éviter les futures sanctions sociales.

Dans l'ensemble, les projets de recherche proposés dans la présente thèse soulèvent la possibilité que le biais de négativité puisse faire partie des mécanismes ayant un impact sur la prise de décision des arbitres sportifs. Cette proposition permet de tisser un lien cohérent avec le cadre conceptuel proposé par Betsch et al. (2011), lequel soutient que la situation influence la prise de décision. Le biais de négativité viendrait préciser comment la valence de l'information est traitée dans la portion « intégration de l'information », permettant de pallier une lacune des études existantes en jugement et prise de décision en contexte sportif.

Implications potentielles des résultats

Les résultats obtenus dans les projets de recherche réalisés permettent de faire progresser les connaissances en lien avec la prise de décision des arbitres sportifs. Ces connaissances ont des implications dans différentes sphères, notamment en termes théorique et pratique.

Implications théoriques

Au-delà du lien proposé avec le biais de négativité, il existe d'autres implications théoriques qui découlent des études réalisées dans le cadre de la thèse doctorale en présence. D'abord, la première étude permet de faire évoluer l'étendue des connaissances en ce qui

concerne le phénomène bien connu de l'avantage de jouer à domicile. Elle fournit également des arguments appuyant l'implication des arbitres dans ce phénomène. De plus, l'étude soutient l'influence de la pression sociale sur les décisions des arbitres de hockey, même si cette explication demeure encore une fois superficielle.

Si les éléments présentés dans le précédent paragraphe sont intéressants au niveau théorique, ils ne font que s'ajouter aux connaissances en regard de phénomènes précédemment observés et rapportés dans le domaine de l'arbitrage au hockey. Ce qui distingue l'article présenté au chapitre II des autres projets de recherche en lien avec la thématique de l'influence de la foule sur les arbitres de hockey est la présence d'une plus grande propension à éviter de pénaliser l'équipe locale plutôt qu'à pénaliser de manière plus fréquente l'équipe adverse. La source de l'avantage offert à l'équipe locale en ce qui concerne le nombre de pénalités accordées était jusque-là impossible à identifier en raison des contextes d'études, qui ne permettaient que de soulever une différence entre le nombre de pénalités accordées. Pour la première fois, il est possible de mettre en lumière que cette différence s'explique principalement par une plus faible tendance à pénaliser l'équipe locale en présence d'une foule. Les résultats ici présentés suggèrent une propension des arbitres de la Ligue Nationale de Hockey à éviter les décisions qui entraînent des critiques de la part de la foule, laquelle est supérieure à la recherche d'approbation de la part des spectateurs. L'importance des matchs doit toutefois être prise en considération, puisque cette observation n'est significative que lorsque les enjeux des parties sont le plus élevés, c'est-à-dire en séries éliminatoires.

La deuxième étude, portant sur l'impact de l'agression verbale à l'endroit des arbitres de baseball, génère elle aussi des avancées théoriques particulièrement importantes. En effet, le projet étant appuyé sur deux théories en compétition, soit que l'agression verbale peut aider ou nuire à la personne agressive, il est pertinent de réaliser que c'est la première hypothèse qui

semble dominante dans le contexte étudié. Puisque les théories sur l'agression verbale sont utiles pour prédire le comportement, les retombées théoriques du projet peuvent s'étendre au-delà du contexte du baseball professionnel. Si cette généralisation doit être réalisée avec prudence, elle permet néanmoins de surpasser le niveau de la simple intuition, puisque les conclusions de l'étude permettent d'émettre des hypothèses quant aux conséquences potentielles de la violence verbale appuyées sur des résultats quantitatifs.

Les résultats permettent également de spécifier le contexte dans lequel la théorie dominante est celle appuyant les bénéfices de l'agression verbale. Il importe que la situation implique un sentiment d'injustice perçu. De plus, l'agression verbale doit être spécifique à cette injustice, puisqu'un excès de colère lié à une thématique non spécifique n'entraînera probablement pas les bénéfices spécifiques escomptés. Également, il est important de constater que les retombées positives observables à la suite d'une agression verbale peuvent être dirigées vers le groupe d'appartenance, plutôt que vers la personne agressive elle-même. Dans le cas du baseball, c'est l'équipe de l'agresseur qui se voit profiter de l'agression puisque ce dernier a été éjecté de la partie. Il est donc possible que les bénéfices pour l'agresseur soient indirects, par l'entremise de son sentiment d'appartenance à son équipe, pendant que ses coéquipiers profitent des bénéfices directs.

Le projet proposé au chapitre III permet également des retombées théoriques supplémentaires en écartant la vraisemblance d'une théorie alternative qui aurait pu expliquer l'avantage offert à l'équipe d'une personne agressive. Le postulat repose principalement sur le fait que l'impact observé dans la prise de décision d'un arbitre après qu'il eut éjecté une personne pourrait être expliqué par une tentative de compenser l'effet négatif de l'éjection, un comportement connu sous le nom de la théorie de la décision de compensation (Hamilton, 2011), plutôt que par une réaction à l'agression verbale. Cette théorie aurait pu être appuyée par une

modification de la prise de décision d'un arbitre qui corrèle avec la magnitude de la conséquence négative de l'éjection d'une personne. Comme chaque membre d'une équipe n'a pas le même impact sur la probabilité d'accéder à la victoire, la validité de la théorie aurait pu être vérifiée par l'observation de décisions plus avantageuses lorsqu'un arbitre éjecte les membres les plus importants d'une équipe. Ce n'est toutefois pas ce que les résultats proposent, suggérant plutôt que le statut de la personne éjectée n'a pas d'impact significatif sur les décisions de l'arbitre. Ce constat permet de faire évoluer les connaissances en écartant la théorie de la décision de compensation comme mécanisme explicatif des résultats.

Pour renforcer la contribution théorique de la recherche présentée au chapitre III, il est à noter que les résultats de ce projet ont fait l'objet d'une initiative de reproduction à grande échelle par l'entremise d'une collaboration entre la revue *Psychological Science* et l'*Institute for Replication*. Cette initiative a mené à la conclusion que les résultats rapportés dans l'article original sont entièrement reproductibles en termes informatiques sur la base des données et du code informatique fourni en accès ouvert par les auteurs. Le rapport permet également de constater que le résultat principal de l'article est robuste à toutes les spécifications alternatives préenregistrées par le responsable de la reproduction, offrant un appui supplémentaire aux conclusions rapportées. Le rapport de reproduction est présenté à l'annexe C.

Implications pratiques

La plus grande implication pratique qui découle des résultats de la présente thèse est assurément la possibilité d'améliorer la formation destinée aux arbitres sportifs. En ayant exposé l'impact des forces sociales sur la prise de décision des officiels, il sera maintenant possible de déterminer des stratégies visant à mieux outiller ces derniers pour faire face aux sources

potentielles de biais générées par la présence et le comportement des joueurs, entraîneurs et spectateurs.

Rendre les arbitres conscients de l'influence de l'information négative est une première étape. Sans les connaissances liées aux possibles biais décisionnels engendrés par la critique de leur travail, qu'elle soit réelle ou anticipée, il est impossible pour les arbitres de concevoir l'importance de porter attention à cet élément. Les formateurs et superviseurs d'arbitres pourront également revenir sur les situations où des comportements négatifs ont été dirigés vers les officiels et demander à ces derniers de réfléchir à l'impact que cela a eu sur leurs décisions post-agression. L'utilisation de la pleine conscience pourrait être une avenue intéressante, stratégie qui a été éprouvée dans d'autres contextes comme étant en mesure de réduire le biais de négativité (Kiken et Shook, 2011). Évidemment, cette stratégie devra faire l'objet d'études démontrant des retombées positives avant d'être implantée dans le cursus de formation des arbitres.

Des situations d'entraînement plus immersives pourraient également être créées afin d'inclure des périodes d'exposition à l'information négative. Par exemple, la création d'environnements virtuels représentant des situations de match où des agressions verbales ont lieu pourraient servir de méthode d'entraînement pour les arbitres. Le contexte virtuel procure un contrôle sur les réactions des participants aux compétitions sportives, dans le cas qui nous intéresse les joueurs, entraîneurs et spectateurs. Une rétroaction offerte aux arbitres après les avoir exposés à ce type d'environnement pourrait leur permettre de diminuer l'impact des forces sociales générées par les acteurs du domaine sportif sur leur prise de décision. À noter qu'il existe déjà certains systèmes d'entraînement en réalité virtuelle pour les arbitres de baseball et de softball (Kaiser et al., 2024; Yanase et al., 2021). Des environnements similaires pourraient être créés pour les arbitres de hockey et pour ceux d'autres sports pouvant bénéficier des conclusions

des études susmentionnées. Ces environnements virtuels pourraient aider à rendre plus réalistes les séances pratiques et à maximiser les retombées de l'entraînement des officiels.

Si la recherche sur la prise de décision des arbitres vise d'abord et avant tout à améliorer le fruit de leur travail, elle permet tout de même d'influencer d'autres acteurs du domaine sportif. Ainsi, les dirigeants des diverses ligues et associations pourraient utiliser l'information présentée dans la présente thèse pour modifier les politiques sportives de manière à décourager les comportements d'agression. En étant conscient que certains avantages peuvent être obtenus par l'entremise de comportements négatifs, il est peut-être nécessaire d'associer ces bénéfices avec des conséquences négatives plus importantes que celles présentement octroyées. Cela permettrait de rendre moins attrayantes les critiques envers les arbitres. Par exemple, des suspensions plus sévères pour les joueurs et entraîneurs qui critiquent le travail des officiels, ou des expulsions des stades pour les spectateurs qui dénigrent les décisions des arbitres de manière exagérée. Bien que plusieurs disciplines pénalisent déjà les comportements d'agression, l'amplitude de leurs punitions ne prend probablement pas en considération que ces comportements peuvent engendrer des avantages pour les agresseurs, tel que suggèrent les résultats présentés ci-haut.

Finalement, les conclusions des études proposées dans la thèse en présence peuvent aussi servir aux joueurs, entraîneurs et spectateurs. Puisque certains avantages peuvent découler de l'utilisation d'information négative, ces personnes pourraient adopter des stratégies visant à influencer les décisions des arbitres à leur avantage. Si une telle conduite demeure moralement discutable, il n'en demeure pas moins que les connaissances acquises par l'intermédiaire des projets de recherche précédemment exposés peuvent être appliquées à des fins stratégiques en contexte de compétitions sportives. Ces retombées ne sont pas nécessairement enviables, mais elles font appel à la logique pour une personne qui considère la victoire plus importante que le respect des conventions sociales et des êtres humains qui pratiquent le métier d'arbitre.

Contexte et limites des résultats

Les résultats préalablement exposés suggèrent d'importantes implications, autant sur le plan théorique que pratique. Cependant, les retombées doivent être mises en contexte, puisque certaines limites accompagnent évidemment les études réalisées.

Une des principales limites à adresser est la généralisation des résultats. Les deux projets de recherche ont été réalisés dans des contextes d'études précis, lesquels incluent les meilleures ligues sportives des sports sélectionnés et, du même coup, les arbitres considérés comme étant les meilleurs de leur discipline. Ce contexte appelle à la prudence en ce qui concerne la généralisation des résultats à des arbitres ne possédant pas les mêmes caractéristiques que ceux étudiés. En effet, il est possible que les forces sociales affectent différemment les arbitres moins expérimentés et ceux œuvrant dans des ligues de niveaux inférieurs. De plus, les projets de recherche ont été réalisés sur des populations masculines, ne permettant pas de présumer l'impact de l'information négative sur des arbitres féminines. Des différences culturelles pourraient également passer sous le radar des projets actuels, puisque les ligues étudiées sont toutes situées en sol nord-américain. Même si les arbitres peuvent provenir de nationalités variées, tout comme les entraîneurs, joueurs et spectateurs qui les côtoient, les attentes qui viennent avec le rôle d'officiel dans la MLB, la NHL et la CHL sont grandement ancrées dans la culture nord-américaine.

De manière hypothétique, il est possible de généraliser les résultats aux personnes en position d'autorité, puisque les arbitres possèdent ce rôle dans le contexte sportif. C'est d'ailleurs ce qui a été proposé dans le deuxième article, avec des appuis théoriques élaborés dans le matériel supplémentaire 2 présenté à l'annexe D. Même s'il existe de bonnes raisons de mettre de l'avant la comparaison entre les arbitres et d'autres figures d'autorité, il faut encore une fois être prudent avec cette tentative de généralisation. En effet, la comparaison se veut surtout une

proposition à vérifier dans d'autres projets de recherche et non une certitude que la prise de décision observée chez les arbitres sera aussi observée chez des juges, des policiers ou des dirigeants d'entreprise, par exemple.

Une autre sphère dans laquelle il faut faire attention avant de transposer les résultats des études illustrées dans la thèse doctorale en présence est celle du sport individuel. Cet appel à la prudence repose notamment sur les résultats de l'article réalisé au baseball, où les retombées positives d'une agression verbale sont palpables auprès du groupe d'appartenance de la personne agressive plutôt que directement à l'endroit de celle-ci. Cette possibilité se résorbe dans le contexte du sport individuel, où les conséquences d'une agression verbale ne pourraient pas se solder par des décisions favorables à l'endroit des coéquipiers de la personne agressive.

Les résultats doivent aussi considérer les enjeux de la compétition, puisqu'il semble y avoir une influence plus marquée des forces sociales sur les arbitres lorsque l'importance du match augmente. C'est ce que suggère l'étude sur les arbitres de hockey, qui semble plus enclins à éviter les réactions négatives de la foule en séries éliminatoires qu'en saison régulière. L'impact de l'information négative doit donc être considéré avec réserve lorsqu'on observe des arbitres en contexte de compétition où les enjeux sont peu élevés.

Une autre limite à considérer dans les études présentées est la possibilité d'un biais de sélection dans l'embauche des arbitres professionnels. Si les projets de recherche ont inclus tous les arbitres des ligues sélectionnées, il n'en demeure pas moins que les personnes qui les ont recrutés peuvent elles-mêmes être victimes de certains biais, lesquels viennent influencer les résultats des études. En effet, si les employeurs des arbitres sont eux-mêmes victimes du biais de négativité, il est possible qu'ils aient tendance à sélectionner les arbitres qui se font le moins critiquer par les joueurs, entraîneurs ou spectateurs. Ce biais dans la sélection se transposerait dans la prise de décision de la majorité des arbitres professionnels étudiés, générant du même

coup les résultats observés. Les arbitres moins sujets à l'influence de l'information négative seraient donc exclus de l'échantillon dans une proportion suffisamment élevée pour modifier les résultats. Cette possibilité est à considérer avant de généraliser les résultats à l'ensemble de la population des arbitres sportifs.

Finalement, une limite à ne pas passer sous silence se veut la nature écologique des contextes d'études. Bien que des études en situations réelles d'observation offrent de nombreux avantages, elles sont assurément accompagnées d'inconvénients significatifs. Le désavantage majeur d'un contexte écologique est l'absence de contrôle sur les variables à l'étude. En plus de devoir éliminer d'emblée un effet causal des différentes forces sociales sur la prise de décision des arbitres sportifs, il est possible que ce contexte implique une ou plusieurs variables non considérées qui seraient mieux à même d'expliquer les résultats observés. Des études expérimentales seront nécessaires avant de statuer du rôle des forces sociales sur la prise de décision des arbitres.

Directions futures

Comme mentionné à la section précédente, il est primordial d'utiliser les résultats présentés jusqu'ici avec discernement. Si ces résultats appellent à la prudence, ils permettent toutefois de faire évoluer à leur manière les connaissances en lien avec la prise de décision des arbitres, en plus de soulever de nouvelles possibilités d'études qui permettraient de pallier certaines incertitudes laissées en place par les limites susmentionnées.

Les projets en réalité virtuelle

Une première proposition se veut la création d'études en laboratoires permettant de contrôler les nombreuses variables confondantes faisant partie des environnements naturels utilisés dans les deux projets détaillés précédemment. Pour ce faire, la réalité virtuelle pourrait ici

aussi servir d'outils de prédilection. Avec les arbitres de baseball, il serait possible de générer des tâches de prise de décisions en lien avec la zone de prise où la seule variable qui fluctue est la rétroaction fournie aux arbitres, cette dernière pouvant prendre la forme de commentaires négatifs, positifs ou neutres. Cette méthode permettrait notamment d'éliminer la variation observée en prémisse des attaques verbales dirigées à l'endroit des arbitres au marbre dans le projet présenté au chapitre III. En effet, la principale lacune de ce projet est que les sanctions sociales observées découlent d'une réaction à une injustice réelle, les arbitres au marbre ayant défavorisé l'équipe de la personne agressive préalablement à son éjection. Le contrôle sur cette variable permettrait de déterminer si les avantages liés à l'agression verbale le sont même lorsque la sanction sociale est injustifiée. Un scénario en réalité virtuelle permettrait également de mesurer l'impact des agressions verbales qui ne mènent pas à une éjection, c'est-à-dire d'intensité moindre, chose qui n'a pu être réalisée dans le contexte antérieur d'étude. De plus, des paramètres pourraient être rendus équivalents entre les conditions plutôt que contrôlés statistiquement, comme les caractéristiques du match, des lancers et des joueurs.

Le même exercice pourrait être réalisé pour générer des études dans le contexte de l'arbitrage au hockey. Le nombre et les types de gestes possiblement répréhensibles sous forme de pénalités à juger par un arbitre varient d'un match à l'autre et la création de scénarios virtuels identiques, où la seule variable changeante est la réaction de la foule, permettrait d'étayer encore plus la compréhension de l'impact des forces sociales de différentes valences. Ici aussi, plutôt que de contrôler statistiquement certains éléments liés aux matchs et aux joueurs, il serait possible de les rendre uniformes, en modifiant uniquement les réactions de la foule dans les différentes conditions.

L'influence des caractéristiques des participants

Puisque les études proposées dans la présente thèse doctorale ont été effectuées auprès d'arbitres professionnels, il est possible que des arbitres possédant des caractéristiques différentes réagissent autrement au comportement des joueurs, entraîneurs et spectateurs (Sors et al., 2019). Pour vérifier cette hypothèse, il serait approprié d'étudier des arbitres possédant des profils différents.

D'abord, la tenue de projets de recherche possédant des méthodologies similaires à ceux présentés aux chapitres II et III pourraient être réalisés, mais cette fois avec des arbitres provenant de l'extérieur de l'Amérique du Nord. Comme mentionné précédemment, l'environnement culturel pourrait avoir un impact sur la réaction aux forces sociales dans le contexte de l'arbitrage, ce qui limite la généralisation des résultats obtenus dans les ligues situées au Canada et aux États-Unis à l'ensemble des arbitres. Des ligues professionnelles de baseball au Japon et de hockey en Europe rendent disponibles des données qui permettraient de répéter les projets présentés dans la présente thèse doctorale, à quelques détails près.

Au-delà de la culture, le genre des arbitres pourrait lui aussi jouer un rôle dans la réaction aux forces sociales. En effet, certaines études ont suggéré des différences d'amplitude du biais de négativité en fonction du genre (Ashare et al., 2013; Braund et al., 2019; Norris, 2021). Pour évaluer la possibilité que des réactions différentes puissent être observables chez les hommes et les femmes, des études réalisées sur des femmes arbitres sont nécessaires. Même si elles demeurent présentes dans une proportion moindre que les hommes, les femmes arbitrant dans le sport professionnel masculin sont en augmentation (de Haan et Warner, 2020). Ainsi, des comparaisons à l'intérieur de mêmes ligues pourraient être réalisées entre les arbitres des deux genres. Évidemment, le sport professionnel féminin est aussi un environnement où l'on peut

observer des femmes arbitres, permettant également de proposer des études dans ce contexte (Drury et al., 2022).

Toujours dans l'optique de généraliser l'effet des forces sociales sur la prise de décision des arbitres à plus grande échelle, il serait intéressant de mener des projets de recherche dans le contexte du sport individuel. En effet, tel que suggéré dans la section sur les limites des résultats, il est difficile de prévoir si les gains observés chez le groupe d'appartenance dans les sports d'équipe se remarquent aussi chez un individu, et si oui, dans quelle mesure. Un sport comme le tennis, où la technologie et les règles créent un parallèle intéressant avec la zone de prise au baseball, pourrait faire l'objet d'études. Sans être répertoriés au même titre que les éjections au baseball, les épisodes de violence pourraient être observés par l'entremise d'analyses vidéo de matchs ou d'observations directes, afin de déterminer l'influence des sanctions sociales sur la prise de décision.

Une autre thématique pertinente à aborder réside dans les différences individuelles chez les arbitres. Des projets où la personnalité des arbitres est mise en relation avec leur tendance à éviter les sanctions sociales ou à rechercher l'approbation sociale pourraient aussi faire avancer les connaissances du domaine de la prise de décision en contexte sportif. Des recherches sur cette thématique permettraient potentiellement d'illustrer un type de personnalité dominant chez les arbitres comparativement aux non-arbitres, ainsi qu'à voir si les meilleurs arbitres ont des personnalités différentes des arbitres de moins haut calibre. L'évitement des sanctions sociales pourrait être une motivation plus importante chez certaines personnes en fonction de leur tempérament, rendant ce genre d'études pertinentes à réaliser dans le contexte de l'arbitrage sportif (Corr et McNaughton, 2012; Elliot et Thrash, 2002).

L'impact du biais de négativité

Si la présente thèse doctorale permet de tracer des associations entre le comportement d'arbitres sportifs et le biais de négativité, il demeure essentiel de poursuivre la recherche sur cette thématique pour renforcer le lien théorique qui peut unir ces deux éléments.

Dans le même ordre d'idées que la possibilité d'observer des différences individuelles au niveau de la personnalité risque d'avoir des impacts sur les décisions des arbitres, des différences individuelles de sensibilité au biais de négativité pourraient aussi être liées à la prise de décision de ces derniers. Des méthodes comportementales de mesures du biais de négativité pourraient être utilisées, pour ensuite comparer les scores à la tendance des arbitres sportifs à éviter les sanctions sociales en situation réelle d'arbitrage (Fazio et al., 2004; Miao et al., 2022). Des méthodes plus complexes, comme la mesure de potentiels liés à l'événement en électroencéphalographie ou certaines techniques d'imagerie par résonance magnétique fonctionnelle, pourraient aussi être pertinentes à utiliser (Norris et al., 2011).

L'intérêt théorique de la relation entre le biais de négativité et le comportement des arbitres pourrait aussi ouvrir la voie à des projets plus appliqués, lesquels pourraient mesurer l'impact de stratégies visant à réduire le biais de négativité sur la prise de décision des arbitres. Par exemple, comme proposée auparavant, la pleine conscience pour diminuer le biais de négativité pourrait être envisagée comme stratégie d'entraînement des arbitres (Ford et al., 2023; Gibb et al., 2022; Kiken et Shook, 2011). L'efficacité de la stratégie pourrait être mesurée, faisant l'objet d'un ou de plusieurs projets de recherche. D'autres stratégies de gestion de stress pourraient aussi être considérées. De surcroît, des méthodes d'amorçage positif pourraient être insérées dans le type d'études suggérées ici. Ces dernières ont déjà fait l'objet de recommandations favorables pour aider à limiter l'impact de l'information négative, mais doivent être considérées avec prudence, car elles ont récemment été remise en question dans certaines

conditions (Zayas et al., 2022). Finalement, des exercices de conscience de soi pourraient être effectués par les arbitres afin de déterminer leur influence sur leur prise de décision. Xu et al. (2021), sur la base d'un projet utilisant des méthodes électrophysiologiques, suggèrent que la conscience de soi pourrait moduler les rétroactions négatives. En contexte d'arbitrage, ces résultats pourraient avoir le potentiel de diminuer l'impact des critiques faites par les joueurs, entraîneurs et spectateurs sur la prise de décision.

Les sous-catégories de pression sociale

Parmi les lacunes des écrits scientifiques en lien avec l'impact de la pression sociale exercée sur les arbitres sportifs, le manque de précision du concept de pression sociale a été soulevé, représentant une limite majeure de l'évolution des connaissances dans l'étude du jugement et de la prise de décision en contexte sportif. Pour pallier cette lacune, des études visant à évaluer les sous-catégories de pression sociale mériteraient d'être mises sur pied. Il serait intéressant de comprendre comment les trois sous-catégories de pression sociale, c'est-à-dire l'influence informationnelle, l'influence normative et l'influence interpersonnelle, peuvent influencer le travail des arbitres.

Pour explorer les conséquences potentielles de ces types d'influence, des projets où la pression exercée sur les arbitres provient des autres arbitres seraient appropriés. En effet, l'influence normative s'observe principalement au niveau des normes du groupe d'appartenance, dans le cas présent, l'ensemble des autres arbitres. Utiliser une procédure expérimentale pour placer des arbitres dans une position où leurs décisions seraient soit jugées a posteriori par d'autres arbitres (groupe expérimental), soit exemptées de jugement (groupe contrôle), permettrait de voir si certains biais sont causés par l'influence normative.

Dans une même ligne de pensée, l'étude des commentaires des superviseurs des arbitres pourrait être étudiée à titre d'influence informationnelle. Ces commentaires sont habituellement beaucoup plus rationnels que ceux en provenance des joueurs, entraîneurs ou spectateurs, qui cherchent communément à obtenir un avantage pour leur équipe quand ils s'expriment. Des protocoles similaires à celui proposé pour l'influence normative, mais en remplaçant les arbitres du groupe d'appartenance qui forment le groupe expérimental par les superviseurs des arbitres et la condition de jugement par des rétroactions, permettrait d'analyser des comportements qui rencontrent la définition de l'influence informationnelle.

En ce qui concerne l'influence interpersonnelle, une bonne manière de l'étudier serait de contrôler les commentaires de la foule, des joueurs ou des entraîneurs dans des études expérimentales. Si les contextes étudiés dans la présente thèse doctorale peuvent porter à croire que la pression sociale en provenance de ces personnes est principalement causée par l'influence interpersonnelle, il n'est pas possible de l'assumer hors de tout doute, puisqu'aucun contrôle n'était exercé sur le contenu des messages prodigués aux arbitres dans les études écologiques en situations réelles de compétition. Les commentaires des participants pouvaient autant être rationnels qu'être des menaces ou des récompenses, ce qui ne permet pas de distinguer l'influence informationnelle de l'influence interpersonnelle. Limiter les stimuli présentés aux arbitres à des menaces ou des récompenses permettrait de rencontrer la définition de l'influence interpersonnelle et de mesurer son impact sur les processus décisionnels des arbitres sportifs.

Synthèse

Les projets de recherche proposés dans la présente thèse doctorale permettent de faire évoluer les connaissances dans le domaine de la prise de décision des arbitres sportifs.

L'information sociale en présence dans le contexte sportif, plus particulièrement l'influence des

forces sociales, a été illustrée comme ayant le potentiel de biaiser le processus décisionnel des officiels. En prenant racine dans des contextes réels de compétitions sportives, les articles présentés apportent une forte validité écologique et un potentiel intéressant de généralisation des résultats. Les lacunes qui viennent avec des designs non expérimentaux permettent de proposer des études à réaliser dans le futur, afin de poursuivre le travail de clarification du processus décisionnel des arbitres et de vérifier la proposition d'enraciner la recherche sur la prise de décision des arbitres dans la théorie du biais de négativité. Les retombées immédiates de la thèse doctorale ici présentée sont donc perceptibles, mais ne sont qu'un premier pas vers l'amélioration du travail des arbitres sportifs.

CHAPITRE V

Conclusion

Le milieu sportif est un environnement de choix pour observer les interactions sociales entre les êtres humains. En plein cœur de l'écosystème sportif, on retrouve les arbitres, qui ont le mandat de faire respecter les règles des différentes disciplines. Affublés d'une quantité innombrable d'information de nature sociale, les arbitres sont assurément à risque de commettre des erreurs. En effet, ces derniers sont exposés à de nombreuses forces sociales, lesquelles ont le pouvoir d'influencer leur prise de décision. C'est particulièrement le cas de l'information négative, qui est reconnue pour avoir un impact supérieur à l'information positive, lorsque traitée par un être humain.

La thèse doctorale présentée dans ce document visait à comprendre comment l'information sociale retrouvée dans l'environnement sportif influence la prise de décision des arbitres. Afin de rendre les résultats généralisables, la thèse a été inspirée par deux disciplines sportives différentes, permettant d'observer l'influence de deux comportements humains distincts. Plus précisément, le chapitre II présente une étude consistant à mesurer l'impact du comportement de la foule sur l'appel de pénalités au hockey. Pour sa part, le chapitre III a permis l'analyse de l'influence de la violence verbale produite par les joueurs et entraîneurs de baseball.

Comme anticipé, les résultats soutiennent l'influence des forces sociales sur la prise de décision des arbitres sportifs, en amenant ces derniers à prendre des décisions biaisées. En s'appuyant sur un contexte écologique, le projet de thèse doctoral présenté ici permet de faire avancer les connaissances dans le domaine du jugement et de la prise de décision non seulement des arbitres sportifs, mais également des êtres humains en position d'autorité lorsque ceux-ci sont exposés aux contingences de diverses forces sociales.

Les retombées du projet se retrouvent dans trois champs d'action principaux. D'abord, la proposition d'étudier le jugement et la prise de décision des arbitres en insérant cette thématique dans la théorie du biais de négativité pourrait offrir une compréhension plus étoffée des processus cognitifs impliqués lorsqu'un arbitre prend une décision. Ensuite, les connaissances acquises devraient permettre de mieux former les arbitres afin de limiter l'impact des contingences environnementales sur leur prise de décision. Dans un troisième temps, l'amélioration de la prise de décision des arbitres pourrait entraîner une diminution des comportements inadéquats dans le monde du sport. En effet, en connaissant mieux l'effet des comportements inacceptables comme la critique excessive des décisions des arbitres, ces derniers pourront adapter leurs décisions afin de diminuer les avantages potentiels ou maximiser les inconvénients associés à ces comportements. Cet ajustement pourrait décourager les différents acteurs des compétitions sportives à adopter des comportements socialement inacceptables, favorisant ainsi un environnement de pratique sportive plus respectueux et accessible.

RÉFÉRENCES

- Adler M. (1985). Stardom and talent. *The American Economic Review*, 75(1), 208–212. <https://www.jstor.org/stable/1812714>
- Albert, J., Glickman, M. E., Swartz, T. B., & Koning, R. H. (Eds.). (2017). *Handbook of statistical methods and analyses in sports*. CRC Press.
- Allen, K. (2019, 6 mai). Five-alarm fire: NHL playoffs show officiating, review process need significant overhaul. USA Today. <https://www.usatoday.com/story/sports/nhl/columnist/kevinallen/2019/05/16/nhl-playoffs-2019-officiating-review/3691445002/>
- American Psychological Association. (2018). *APA Dictionary of Psychology: social pressure*. <https://dictionary.apa.org/social-pressure>
- Anderson, K. J., & Pierce, D. A. (2009). Officiating bias: The effect of foul differential on foul calls in NCAA basketball. *Journal of sports sciences*, 27(7), 687-694. <https://doi.org/10.1080/02640410902729733>
- Archsmith J., Heyes A., & Saberian S. (2018). Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation. *Journal of the Association of Environmental and Resource Economists*, 5(4), 827–863. <https://doi.org/10.1086/698728>
- Ashare, R. L., Norris, C. J., Wileyto, E. P., Cacioppo, J. T., & Strasser, A. A. (2013). Individual differences in positivity offset and negativity bias: Gender-specific associations with two serotonin receptor genes. *Personality and individual differences*, 55(5), 469-473. <https://doi.org/10.1016/j.paid.2013.04.009>
- Baker, C., Pawling, R., & Fairclough, S. (2020). Assessment of threat and negativity bias in virtual reality. *Scientific reports*, 10(1), 1-10. <https://doi.org/10.1038/s41598-020-74421-1>
- Bar-Eli, M., Plessner, H., & Raab, M. (2011). *Judgment, decision-making and success in sport*. John Wiley & Sons.

- Baron, R. A., & Richardson, D. R. (1994). *Human aggression*. Springer Science & Business Media.
- Barr D. J., Levy R., Scheepers C., & Tily H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Barrett, D. (2021). Star player bias in the nba: A quantile regression approach. *Empirical Economic Bulletin, An Undergraduate Journal*, 14(1), 9. <https://digitalcommons.bryant.edu/eeb/vol14/iss1/9/>
- Bates D., Mächler M., Bolker B., & Walker S. (2014). *Fitting linear mixed-effects models using lme4*. <https://doi.org/10.48550/arXiv.1406.5823>
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of general psychology*, 5(4), 323-370. <https://doi.org/10.1037/1089-2680.5.4.323>
- Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1994). Losing control: How and why people fail at self-regulation. *San Diego: Academic*.
- Beaudoin, D., Schulte, O., & Swartz, T. B. (2016). Biased penalty calls in the National Hockey League. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 9(5), 365-372. <https://doi.org/10.1002/sam.11320>
- Beaudoin, D., & Swartz, T. B. (2010). Strategies for pulling the goalie in hockey. *The American Statistician*, 64(3), 197-204. <https://doi.org/10.1198/tast.2010.09147>
- Berri D. J., Leeds M., Leeds E. M., & Mondello M. (2009). The role of managers in team performance. *International Journal of Sport Finance*, 4(2), 75–93.
- Betsch, T., Funke, J., & Plessner, H. (2011). *Allgemeine Psychologie für Bachelor: Denken-Urteilen, Entscheiden, Problemlösen. Lesen, Hören, Lernen im Web*. Springer-Verlag.

- Blackburn, P. (2019, 16 mai). 2019 NHL Playoffs: Examining the controversial officiating that has plagued the playoffs. CBS Sports. <https://www.cbssports.com/nhl/news/2019-nhl-playoffs-examining-the-controversial-officiating-that-has-plagued-the-playoffs/>
- Blanchette, I., & Richards, A. (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition & Emotion*, 24(4), 561-595. <https://doi.org/10.1080/02699930903132496>
- Bless, H., Fiedler, K., & Strack, F. (2004). *Social Cognition: How Individuals Construct Social Reality*. Psychology Press.
- Bolker B. M., Brooks M. E., Clark C. J., Geange S. W., Poulsen J. R., Stevens M. H. H., & White J. S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, 24(3), 127–135. <https://doi.org/10.1016/j.tree.2008.10.008>
- Boyko, R. H., Boyko, A. R., & Boyko, M. G. (2007). Referee bias contributes to home advantage in English Premiership football. *Journal of sports sciences*, 25(11), 1185-1194. <https://doi.org/10.1080/02640410601038576>
- Boylen, R. (2019, 7 juin). A look at the most controversial calls of the Stanley Cup Playoffs. Sportsnet. <https://www.sportsnet.ca/hockey/nhl/look-controversial-calls-stanley-cup-playoffs/>
- Braund, T. A., Palmer, D. M., Tillman, G., Hanna, H., & Gordon, E. (2019). Increased chronic stress predicts greater emotional negativity bias and poorer social skills but not cognitive functioning in healthy adults. *Anxiety, Stress, & Coping*, 32(4), 399–411. <https://doi.org/10.1080/10615806.2019.1598555>

- Brick, N. E., Breslin, G., Shevlin, M., & Shannon, S. (2022). The impact of verbal and physical abuse on distress, mental health, and intentions to quit in sports officials. *Psychology of Sport and Exercise*, 63, 102274. <https://doi.org/10.1016/j.psychsport.2022.102274>
- Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative?. *Journal of personality and social psychology*, 36(8), 917.
- Bright, V., Warner, S., & Zvosec, C. (2022). Refereeing as a postathletic career option. *Journal of Sport Management*, 36(6), 548-558. <https://doi.org/10.1123/jsm.2021-0268>
- Brimberg, J., & Hurley, W. J. (2009). Are National Hockey League referees Markov?. *OR Insight*, 22(4), 234-243. <https://doi.org/10.1057/ori.2009.12>
- Bryson, A., Dolton, P., Reade, J. J., Schreyer, D., & Singleton, C. (2021). Causal effects of an absent crowd on performances and refereeing decisions during Covid-19. *Economics Letters*, 198, 109664. <https://doi.org/10.1016/j.econlet.2020.109664>
- Buraimo, B., Forrest, D., & Simmons, R. (2010). The 12th man?: refereeing bias in English and German soccer. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 173(2), 431-449. <https://doi.org/10.1111/j.1467-985X.2009.00604.x>
- Buraimo, B., Simmons, R., & Maciaszczyk, M. (2012). Favoritism and referee bias in European soccer: Evidence from the Spanish League and the UEFA Champions League. *Contemporary Economic Policy*, 30(3), 329-343. <https://doi.org/10.1111/j.1465-7287.2011.00295.x>
- Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological bulletin*, 115(3), 401. DOI: 10.1037/0033-2909.115.3.401.

- Canary, D. J., Spitzberg, B. H., & Semic, B. A. (1996). The experience and expression of anger in interpersonal settings. In *Handbook of communication and emotion* (pp. 189-213). Academic Press.
- Caudill, S. B., Mixon JR, F. G., & Wallace, S. (2014). Life on the red carpet: Star players and referee bias in the national basketball association. *International Journal of the Economics of Business*, 21(2), 245-253. <https://doi.org/10.1080/13571516.2014.903110>
- Corr, P. J., & McNaughton, N. (2012). Neuroscience and approach/avoidance personality traits: A two stage (valuation–motivation) approach. *Neuroscience & Biobehavioral Reviews*, 36(10), 2339-2354. <https://doi.org/10.1016/j.neubiorev.2012.09.013>
- Coulomb-Cabagno, G., Rasclé, O., & Souchon, N. (2005). Players' gender and male referees' decisions about aggression in French soccer: A preliminary study. *Sex Roles*, 52, 547-553. <https://doi.org/10.1007/s11199-005-3720-z>
- Courneya, K. S., & Carron, A. V. (1992). The home advantage in sport competitions: a literature review. *Journal of Sport & Exercise Psychology*, 14(1). <https://doi.org/10.1123/jsep.14.1.13>
- Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. *Journal of personality assessment*, 91(2), 121-136. <https://doi.org/10.1080/00223890802634175>
- Crockett, M. J., Özdemir, Y., & Fehr, E. (2014). The value of vengeance and the demand for deterrence. *Journal of Experimental Psychology: General*, 143(6), 2279-2286. <https://doi.org/10.1037/xge0000018>
- Cueva, C. (2020). Animal spirits in the beautiful game. Testing social pressure in professional football during the COVID-19 lockdown. <https://osf.io/preprints/osf/hczkj>

- Davis, A. P., Leppard, T. R., & Kinney, A. B. (2022). A shame of inches: Are teams with black head coaches more heavily penalized in Division 1 college football?. *Social Science Quarterly*, 103(2), 259-273. <https://doi.org/10.1111/ssqu.13117>
- Dawson, P., & Dobson, S. (2010). The influence of social pressure and nationality on individual decisions: Evidence from the behaviour of referees. *Journal of Economic Psychology*, 31(2), 181-191. <https://doi.org/10.1016/j.joep.2009.06.001>
- Debanne, T. (2014). Techniques used by coaches to influence referees in professional team handball. *International Journal of Sports Science & Coaching*, 9(3), 433-446. <https://doi.org/10.1260/1747-9541.9.3.433>
- de Haan, D., & Warner, S. (2020). Coaches and officials. In *Developing sport for women and girls* (pp. 185-196). Routledge.
- DeLong, J. E. (2017). Neuropsychology behind the plate. *Sport, Ethics and Philosophy*, 11(3), 385-395. <https://doi.org/10.1080/17511321.2017.1311367>
- Demiralp B., Colburn C., & Koch J. V. (2012). The effects of age, experience and managers upon baseball performance. *Journal of Economics and Finance*, 36(2), 481–498. <https://doi.org/10.1007/s12197-010-9141-z>
- Deshpande, S. K., & Wyner, A. (2017). A hierarchical Bayesian model of pitch framing. *Journal of Quantitative Analysis in Sports*, 13(3), 95-112. <https://doi.org/10.1515/jqas-2017-0027>
- Di Corrado, D., Pellarin, E., & Agostini, T. A. (2011). The phenomenon of social influence on the football pitch: Social pressure from the crowd on referees' decisions. *Review of Psychology*, 18(1), 33-36. <https://hrcak.srce.hr/78215>
- Dix, A. (2023). Indications of Referee Bias in Division I Women's College Volleyball: Testing Expectancy Violations and Examining Nonverbal Communication. *International Journal of Sport Communication*, 16(4), 414-422. <https://doi.org/10.1123/ijsc.2023-0050>

- Dohmen, T. J. (2005). Social pressure influences decisions of individuals: Evidence from the behavior of football referees. *Disponible à SSRN 725541*.
<https://ssrn.com/abstract=725541> or <http://dx.doi.org/10.2139/ssrn.725541>
- Dohmen, T. J. (2008). The influence of social forces: Evidence from the behavior of football referees. *Economic inquiry*, 46(3), 411-424. <https://doi.org/10.1111/j.1465-7295.2007.00112.x>
- Dohmen, T., & Sauermann, J. (2016). Referee bias. *Journal of Economic Surveys*, 30(4), 679-695. <https://doi.org/10.1111/joes.12106>
- Downward, P., & Jones, M. (2007). Effects of crowd size on referee decisions: Analysis of the FA Cup. *Journal of sports sciences*, 25(14), 1541-1545.
<https://doi.org/10.1080/02640410701275193>
- Downward, P., Webb, T., & Dawson, P. (2024). Referee abuse, intention to quit, and well-being. *Research quarterly for exercise and sport*, 95(1), 207-217.
<https://doi.org/10.1080/02701367.2023.2184459>
- Drury, S., Stride, A., Fitzgerald, H., Hyett-Allen, N., Pylypiuk, L., & Whitford-Stark, J. (2022). “I’m a Referee, Not a Female Referee”: The Experiences of Women Involved in Football as Coaches and Referees. *Frontiers in sports and active living*, 3, 789321.
<https://doi.org/10.3389/fspor.2021.789321>
- Elliot, A. J., & Thrash, T. M. (2002). Approach-avoidance motivation in personality: approach and avoidance temperaments and goals. *Journal of personality and social psychology*, 82(5), 804. <https://doi.org/10.1037/0022-3514.82.5.804>
- Endrich, M., & Gesche, T. (2020). Home-bias in referee decisions: Evidence from “Ghost Matches” during the Covid19-Pandemic. *Economics Letters*, 197, 109621.
<https://doi.org/10.1016/j.econlet.2020.109621>

- Faltings, R., Krumer, A., & Lechner, M. (2023). Rot-Jaune-Verde: On linguistic bias of referees in Swiss soccer. *Kyklos*, 76(3), 380-406. <https://doi.org/10.1111/kykl.12332>
- Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology*, 87(3), 293–311. <https://doi.org/10.1037/0022-3514.87.3.293>
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137-140. <https://doi.org/10.1038/415137a>
- Findlay, L. C., & Ste-Marie, D. M. (2004). A reputation bias in figure skating judging. *Journal of Sport and Exercise Psychology*, 26(1), 154-166. <https://doi.org/10.1123/jsep.26.1.154>
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. McGraw-Hill Book Company.
- Folger, R., & Cropanzano, R. (2001). Fairness theory: Justice as accountability. *Advances in organizational justice*, 1(1-55), 12. Stanford University Press.
- Ford, C. G., Kiken, L. G., Haliwa, I., & Shook, N. J. (2023). Negatively biased cognition as a mechanism of mindfulness: a review of the literature. *Current Psychology*, 42(11), 8946-8962. <https://doi.org/10.1007/s12144-021-02147-y>
- Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological bulletin*, 118(3), 392. <https://doi.org/10.1037/0033-2909.118.3.392>
- Garicano, L., Palacios-Huerta, I., & Prendergast, C. (2005). Favoritism under social pressure. *Review of Economics and Statistics*, 87(2), 208-216. <https://doi.org/10.1162/0034653053970267>
- Gass, R. H., & Seiter, J. S. (2018). *Persuasion: Social influence and compliance gaining*. Routledge.

- Germar, M., Albrecht, T., Voss, A., & Mojzisch, A. (2016). Social conformity is due to biased stimulus processing: Electrophysiological and diffusion analyses. *Social Cognitive and Affective Neuroscience*, 11(9), 1449-1459. <https://doi.org/10.1093/scan/nsw050>
- Gibb, A., Wilson, J. M., Ford, C., & Shook, N. J. (2022). Does mindfulness reduce negative interpretation bias?. *Cognition and Emotion*, 36(2), 284-299. <https://doi.org/10.1080/02699931.2021.2008322>
- Gift, P., & Rodenberg, R. M. (2014). Napoleon complex: height bias among national basketball association referees. *Journal of Sports Economics*, 15(5), 541-558. <https://doi.org/10.1177/1527002514535168>
- Goumas, C. (2014). Home advantage and referee bias in European football. *European journal of sport science*, 14, S243-S249. <https://doi.org/10.1080/17461391.2012.686062>
- Graf, R. G., Yabko, B. A., & Christensen, P. N. (2009). Gender effects in the assessment of technical fouls among high school basketball officials and collegiate proxies. *Journal of Sport Behavior*, 32(2), 175+. https://link.gale.com/apps/doc/A199684543/AONE?u=crepuq_bishop&sid=googleScholar&xid=1d538776
- Greenberg, J., & Cohen, R. L. (Eds.). (2014). *Equity and justice in social behavior*. Academic Press.
- Guerin, B. (2003). Language use as social strategy: A review and an analytic framework for the social sciences. *Review of General Psychology*, 7(3), 251-298. <https://doi.org/10.1037/1089-2680.7.3.251>
- Guillén, F., & Feltz, D. L. (2011). A conceptual model of referee efficacy. *Frontiers in psychology*, 2, 8562. <https://doi.org/10.3389/fpsyg.2011.00025>

- Hamilton, M. (2011). The moral ambiguity of the makeup call. *Journal of the Philosophy of Sport*, 38(2), 212-228. <https://doi.org/10.1080/00948705.2011.10510423>
- Hamilton, M. A. (2012). Verbal aggression: Understanding the psychological antecedents and social consequences. *Journal of Language and Social Psychology*, 31(1), 5-12. <https://doi.org/10.1177/0261927X11425032>
- Harris, A. R., & Roebber, P. J. (2019). NBA team home advantage: Identifying key factors using an artificial neural network. *PloS one*, 14(7), e0220630. <https://doi.org/10.1371/journal.pone.0220630>
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Edwins, G., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2006). Costly punishment across human societies. *Science*, 312(5781), 1767-1770. <https://doi.org/10.1126/science.1127333>
- Hopfensitz, A., & Reuben, E. (2009). The importance of emotions for the effectiveness of social punishment. *The Economic Journal*, 119(540), 1534-1559. <https://doi.org/10.1111/j.1468-0297.2009.02288.x>
- Hosmer D. W. Jr., Lemeshow S., & Sturdivant R. X. (2013). *Applied logistic regression* (Vol. 398). John Wiley.
- Huang J., & Hsu H. J. (2020). Approximating strike zone size and shape for baseball umpires under different conditions. *International Journal of Performance Analysis in Sport*, 20(2), 133–149. <https://doi.org/10.1080/24748668.2020.1726156>
- Huang, Y. X., & Luo, Y. J. (2006). Temporal course of emotional negativity bias: an ERP study. *Neuroscience letters*, 398(1-2), 91-96. <https://doi.org/10.1016/j.neulet.2005.12.074>

- Inan, T. (2020). The effect of crowd support on home-field advantage: Evidence from European football. *Annals of Applied Sport Science*, 8(3), 0-0. <http://aassjournal.com/article-1-806-en.html>
- Infante, D. A., & Rancer, A. S. (1996). Argumentativeness and verbal aggressiveness: A review of recent theory and research. *Annals of the International Communication Association*, 19(1), 319-352. <https://doi.org/10.1080/23808985.1996.11678934>
- Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *Journal of personality and social psychology*, 75(4), 887. <https://doi.org/10.1037/0022-3514.75.4.887>
- Ito, T., Yokokawa, K., Yahata, N., Isato, A., Suhara, T., & Yamada, M. (2017). Neural basis of negativity bias in the perception of ambiguous facial expression. *Scientific reports*, 7(1), 1-9. <https://doi.org/10.1038/s41598-017-00502-3>
- Jamieson, J. P. (2010). The home field advantage in athletics: A meta-analysis. *Journal of Applied Social Psychology*, 40(7), 1819-1848. <https://doi.org/10.1111/j.1559-1816.2010.00641.x>
- Jones, M. B. (2018). Differences in home advantage between sports. *Psychology of Sport and Exercise*, 34, 61-69. <https://doi.org/10.1016/j.psychsport.2017.07.012>
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Kaiser, K., Walters, S., Sheehy, K., Murray, E., & Spencer, K. (2024). Implementing the technology shift from 2D to 3D: insights and suggestions for umpire educators. *Frontiers in Virtual Reality*, 5, 1368648. <https://doi.org/10.3389/frvir.2024.1368648>

- Kenning, P., & Plassmann, H. (2005). NeuroEconomics: An overview from an economic perspective. *Brain research bulletin*, 67(5), 343-354.
<https://doi.org/10.1016/j.brainresbull.2005.07.006>
- Khurshudyan, I. (2019, 7 juin). Officiating gaffes continue to plague the Stanley Cup playoffs. The Washington Post. <https://www.washingtonpost.com/sports/2019/06/07/officiating-gaffes-continue-plague-stanley-cup-playoffs/>
- Kiken, L. G., & Shook, N. J. (2011). Looking up: Mindfulness increases positive judgments and reduces negativity bias. *Social Psychological and Personality Science*, 2(4), 425-431.
<https://doi.org/10.1177/1948550610396585>
- Kim J. W., & King B. G. (2014). Seeing stars: Matthew effects and status bias in Major League Baseball umpiring. *Management Science*, 60(11), 2619-2644.
<https://doi.org/10.1287/mnsc.2014.1967>
- Lakens, D., & Evers, E. R. (2014). Sailing from the seas of chaos into the corridor of stability: Practical recommendations to increase the informational value of studies. *Perspectives on psychological science*, 9(3), 278-292. <https://doi.org/10.1177/1745691614528520>
- Larrick R. P., Timmerman T. A., Carton A. M., & Abrevaya J. (2011). Temper, temperature, and temptation: Heat-related retaliation in baseball. *Psychological Science*, 22(4), 423-428. <https://doi.org/10.1177/0956797611399292>
- Leard, B., & Doyle, J. M. (2011). The effect of home advantage, momentum, and fighting on winning in the National Hockey League. *Journal of Sports Economics*, 12(5), 538-560.
<https://doi.org/10.1177/152700251038986>
- Lenth, R. V. (2020). emmeans: Estimated marginal means, aka least-squares means (Version 1.5.2-1) [R package]. <https://CRAN.R-project.org/package=emmeans>

Lenth R. V. (2021). *emmeans: Estimated marginal means, aka least-squares means* (R Package Version 1.7.1-1). <https://CRAN.R-project.org/package=emmeans>

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual review of psychology*, 66(1). <https://doi.org/10.1146/annurev-psych-010213-115043>

Lex, H., Pizzera, A., Kurtes, M., & Schack, T. (2015). Influence of players' vocalisations on soccer referees' decisions. *European journal of sport science*, 15(5), 424-428. <https://doi.org/10.1080/17461391.2014.962620>

Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. Dans RJ Davidson, KR Scherer et HH Goldsmith (dir.), *Handbook of affective sciences* (p. 619–642). Oxford University Press.

MacMahon, C., & Mildenhall, B. (2012). A practical perspective on decision making influences in sports officiating. *International Journal of Sports Science & Coaching*, 7(1), 153-165. <https://doi.org/10.1260/1747-9541.7.1.153>

MacMahon, C., Helsen, W. F., Starkes, J. L., & Weston, M. (2007a). Decision-making skills and deliberate practice in elite association football referees. *Journal of sports sciences*, 25(1), 65-78. <https://doi.org/10.1080/02640410600718640>

MacMahon, C., & Starkes, J. L. (2008). Contextual influences on baseball ball-strike decisions in umpires, players, and controls. *Journal of sports sciences*, 26(7), 751-760. <https://doi.org/10.1080/02640410701813050>

MacMahon, C., Starkes, J. L., & Deakin, J. (2007b). Referee decision making in a video-based infraction detection task: Application and training considerations. *International Journal of Sports Science & Coaching*, 2(3), 257-265. <https://doi.org/10.1260/174795407782233164>

- Magistro, B., & Wack, M. (2023). Racial bias in fans and officials: Evidence from the Italian Serie A. *Sociology*, 57(6), 1302-1321. <https://doi.org/10.1177/00380385221138332>
- Maher, A., Quinlan, P., & Davis, J. (2014). Cheating in sport: contra-normative or desirable deviance?. In, Palmer, C. (Ed.) *The sports monograph: critical perspectives on socio-cultural sport, coaching and Physical Education*. SSTO Publications, Preston, UK.
- Marchi, M., Albert, J., & Baumer, B. S. (2018). *Analyzing baseball data with R*. Chapman and Hall/CRC.
- McCarrick, D., Brewer, G., Lyons, M., Pollet, T. V., & Neave, N. (2020). Referee height influences decision making in British football leagues. *BMC psychology*, 8, 1-10. <https://doi.org/10.1186/s40359-020-0370-4>
- Miao, K., Liu, X., Zhang, X., Li, Y., Liao, X., Zhang, R., Zhengzhi, F., & Chen, Z. (2022). Revision and psychometric properties of the negative cognitive processing bias scale. *Frontiers in Psychiatry*, 13, 1013108. <https://doi.org/10.3389/fpsy.2022.1013108>
- Mills, B. M. (2014). Social pressure at the plate: Inequality aversion, status, and mere exposure. *Managerial and Decision Economics*, 35(6), 387-403. <https://doi.org/10.1002/mde.2630>
- Mills, B. M. (2017). Technological innovations in monitoring and evaluation: Evidence of performance impacts among Major League Baseball umpires. *Labour Economics*, 46, 189-199. <https://doi.org/10.1016/j.labeco.2016.10.004>
- Mills, B. M., & Salaga, S. (2018). A natural experiment for efficient markets: Information quality and influential agents. *Journal of Financial Markets*, 40, 23-39. <https://doi.org/10.1016/j.finmar.2018.07.002>
- Mongeon, K., & Longley, N. (2015). Testing for ethnicity discrimination among NHL referees: A duration model approach. *Eastern Economic Journal*, 41, 86-101. <https://doi.org/10.1057/eej.2013.41>

- Moskowitz, T., & Wertheim, L. J. (2011). *Scorecasting: The hidden influences behind how sports are played and games are won*. Crown Archetype.
- Müller-Pinzler, L., Czekalla, N., Mayer, A. V., Stolz, D. S., Gazzola, V., Keysers, C., Paulus, F. M., & Krach, S. (2019). Negativity-bias in forming beliefs about own abilities. *Scientific reports*, 9(1), 1-15. <https://doi.org/10.1038/s41598-019-50821-w>
- Myers, T. D. (2014). Achieving external validity in home advantage research: generalizing crowd noise effects. *Frontiers in psychology*, 5, 532. <https://doi.org/10.3389/fpsyg.2014.00532>
- Myers, T., & Balmer, N. (2012). The impact of crowd noise on officiating in Muay Thai: achieving external validity in an experimental setting. *Frontiers in Psychology*, 3, 346. <https://doi.org/10.3389/fpsyg.2012.00346>
- Naqvi, N., Shiv, B., & Bechara, A. (2006). The role of emotion in decision making: A cognitive neuroscience perspective. *Current directions in psychological science*, 15(5), 260-264. <https://doi.org/10.1111/j.1467-8721.2006.00448.x>
- Nelissen, R. M. (2008). The price you pay: Cost-dependent reputation effects of altruistic punishment. *Evolution and Human Behavior*, 29(4), 242-248. <https://doi.org/10.1016/j.evolhumbehav.2008.01.001>
- 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(4), 261-272. [https://doi.org/10.1016/S1469-0292\(01\)00033-4](https://doi.org/10.1016/S1469-0292(01)00033-4)
- Nevill, A. M., & Holder, R. L. (1999). Home advantage in sport: An overview of studies on the advantage of playing at home. *Sports Medicine*, 28, 221-236. <https://doi.org/10.2165/00007256-199928040-00001>

- NHL Public Relations (2021, 24 mars). Referee Peel no longer working NHL games. Press release. <https://www.nhl.com/news/referee-tim-peel-no-longer-working-nhl-games/c-322856334>
- Norris, C. J. (2021). The negativity bias, revisited: Evidence from neuroscience measures and an individual differences approach. *Social neuroscience*, 16(1), 68-82.
<https://doi.org/10.1080/17470919.2019.1696225>
- Norris, C. J., Larsen, J. T., Crawford, L. E., & Cacioppo, J. T. (2011). Better (or worse) for some than others: Individual differences in the positivity offset and negativity bias. *Journal of Research in Personality*, 45(1), 100-111. <https://doi.org/10.1016/j.jrp.2010.12.001>
- Perloff, R. M. (2020). *The Dynamics of Persuasion: Communication and Attitudes in the Twenty-First Century*. Routledge.
- Pettersson-Lidbom, P., & Priks, M. (2010). Behavior under social pressure: Empty Italian stadiums and referee bias. *Economics Letters*, 108(2), 212-214.
<https://doi.org/10.1016/j.econlet.2010.04.023>
- Petti B. (2021). *baseballr: Functions for acquiring and analyzing baseball data*. <https://github.com/BillPetti/baseballr/>
- Picazo-Tadeo, A. J., González-Gómez, F., & Guardiola, J. (2017). Does the crowd matter in refereeing decisions? Evidence from Spanish soccer. *International Journal of Sport and Exercise Psychology*, 15(5), 447-459. <https://doi.org/10.1080/1612197X.2015.1126852>
- Pillutla, M. M., & Murnighan, J. K. (1996). Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational behavior and human decision processes*, 68(3), 208-224. <https://doi.org/10.1006/obhd.1996.0100>

- Pinheiro J. C., & Chao E. C. (2006). Efficient Laplacian and adaptive Gaussian quadrature algorithms for multilevel generalized linear mixed models. *Journal of Computational and Graphical Statistics*, 15(1), 58–81. <https://doi.org/10.1198/106186006X96962>
- Plessner, H., Ermark, F., Schütz, L. M., & Schweizer, G. (2023). Sports performance judgments - an update from a social cognitive perspective. *Asian Journal of Sport and Exercise Psychology*, 3(1), 13-23. <https://doi.org/10.1016/j.ajsep.2023.01.002>
- Plessner, H., & Haar, T. (2006). Sports performance judgments from a social cognitive perspective. *Psychology of sport and exercise*, 7(6), 555-575. <https://doi.org/10.1016/j.psychsport.2006.03.007>
- Ponzo, M., & Scoppa, V. (2018). Does the home advantage depend on crowd support? Evidence from same-stadium derbies. *Journal of Sports Economics*, 19(4), 562-582. <https://doi.org/10.1177/1527002516665794>
- Pope, B. R., & Pope, N. G. (2015). Own-nationality bias: evidence from UEFA Champions League football referees. *Economic Inquiry*, 53(2), 1292-1304. <https://doi.org/10.1111/ecin.12180>
- Powell M. J. (2009). *The BOBYQA algorithm for bound constrained optimization without derivatives* (Cambridge NA Report NA2009/06). University of Cambridge.
- Price, J., Remer, M., & Stone, D. F. (2012). Subperfect game: Profitable biases of NBA referees. *Journal of Economics & Management Strategy*, 21(1), 271-300. <https://doi.org/10.1111/j.1530-9134.2011.00325.x>
- Price, J., & Wolfers, J. (2010). Racial discrimination among NBA referees. *The Quarterly journal of economics*, 125(4), 1859-1887. <https://doi.org/10.1162/qjec.2010.125.4.1859>

- Pryor, C., Perfors, A., & Howe, P. D. (2019). Even arbitrary norms influence moral decision-making. *Nature human behaviour*, 3(1), 57-62. <https://doi.org/10.1038/s41562-018-0489-y>
- Raab, M., Bar-Eli, M., Plessner, H., & Araujo, D. (2019). The past, present and future of research on judgment and decision making in sport. *Psychology of Sport and Exercise*, 42, 25-32. <https://doi.org/10.1016/j.psychsport.2018.10.004>
- Rancer, A. S., & Avtgis, T. A. (2006). *Argumentative and aggressive communication: Theory, research, and application*. Sage.
- R Core Team. (2022). *R: A language and environment for statistical computing* (Version 2022.12.0) [Computer software]. <http://www.R-project.org>
- Ribeiro, H. V., Mukherjee, S., & Zeng, X. H. T. (2016). The advantage of playing home in NBA: Microscopic, team-specific and evolving features. *PloS one*, 11(3), e0152440. <https://doi.org/10.1371/journal.pone.0152440>
- Rilling, J. K., King-Casas, B., & Sanfey, A. G. (2008). The neurobiology of social decision-making. *Current opinion in neurobiology*, 18(2), 159-165. <https://doi.org/10.1016/j.conb.2008.06.003>
- Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual review of psychology*, 62(1), 23-48. <https://doi.org/10.1146/annurev.psych.121208.131647>
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and social psychology review*, 5(4), 296-320. https://doi.org/10.1207/S15327957PSPR0504_2
- Sanfey, A. G. (2007). Social decision-making: insights from game theory and neuroscience. *Science*, 318(5850), 598-602. DOI: 10.1126/science.1142996.

- Schiebener, J., & Brand, M. (2015). Decision making under objective risk conditions—a review of cognitive and emotional correlates, strategies, feedback processing, and external influences. *Neuropsychology review*, 25(2), 171-198. <https://doi.org/10.1007/s11065-015-9285-x>
- Scoppa, V. (2008). Are subjective evaluations biased by social factors or connections? An econometric analysis of soccer referee decisions. *Empirical Economics*, 35(1), 123-140. <https://doi.org/10.1007/s00181-007-0146-1>
- Scoppa, V. (2021). Social pressure in the stadiums: Do agents change behavior without crowd support?. *Journal of economic psychology*, 82, 102344. <https://doi.org/10.1016/j.joep.2020.102344>
- Seip, E. C., Van Dijk, W. W., & Rotteveel, M. (2014). Anger motivates costly punishment of unfair behavior. *Motivation and Emotion*, 38(4), 578-588. <https://doi.org/10.1007/s11031-014-9395-4>
- Seymour, B., Singer, T., & Dolan, R. (2007). The neurobiology of punishment. *Nature Reviews Neuroscience*, 8(4), 300-311. <https://doi.org/10.1038/nrn2119>
- Sing T., Sander O., Beerenwinkel N., & Lengauer T. (2005). ROCR: Visualizing classifier performance in R. *Bioinformatics*, 21(20), 3940–3941. <https://doi.org/10.1093/bioinformatics/bti623>
- Smith, M. D. (1983). Violence and sport. *Violence and sport*. Butterworths.
- Sors, F., Grassi, M., Agostini, T., & Murgia, M. (2020). The sound of silence in association football: Home advantage and referee bias decrease in matches played without spectators. *European journal of sport science*, 21(12), 1597-1605. <https://doi.org/10.1080/17461391.2020.1845814>

- Sors, F., Tomé Lourido, D., Parisi, V., Santoro, I., Galmonte, A., Agostini, T., & Murgia, M. (2019). Pressing crowd noise impairs the ability of anxious basketball referees to discriminate fouls. *Frontiers in psychology*, *10*, 498770. <https://doi.org/10.3389/fpsyg.2019.02380>
- Souchon, N., Fontayne, P., Livingstone, A., Maio, G. R., Mellac, N., & Genolini, C. (2013a). External influences on referees' decisions in judo: The effects of coaches' exclamations during throw situations. *Journal of Applied Sport Psychology*, *25*(2), 223-233. <https://doi.org/10.1080/10413200.2012.713440>
- Souchon, N., Geneviève, C., Rascle, O., Traclet, A., Dosseville, F., & Maio, G. R. (2009). Referees' decision making about transgressions: The influence of player gender at the highest national level. *Psychology of Women Quarterly*, *33*(4), 445-452. <https://doi.org/10.1111/j.1471-6402.2009.01522.x>
- Souchon, N., Livingstone, A. G., & Maio, G. R. (2013b). 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. <https://doi.org/10.1123/jsep.35.6.585>
- Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of experimental social psychology*, *27*(1), 76-105. [https://doi.org/10.1016/0022-1031\(91\)90011-T](https://doi.org/10.1016/0022-1031(91)90011-T)
- Starcke, K., & Brand, M. (2012). Decision making under stress: a selective review. *Neuroscience & Biobehavioral Reviews*, *36*(4), 1228-1248. <https://doi.org/10.1016/j.neubiorev.2012.02.003>
- Ste-Marie, D. M., & Valiquette, S. M. (1996). Enduring memory-influenced biases in gymnastic judging. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1498. <https://doi.org/10.1037/0278-7393.22.6.1498>

- Swaab R. I., Schaerer M., Anicich E. M., Ronay R., & Galinsky A. D. (2014). The too-much-talent effect: Team interdependence determines when more talent is too much or not enough. *Psychological Science*, 25(8), 1581–1591.
<https://doi.org/10.1177/0956797614537>
- Szabó, D. Z. (2022). The impact of differing audience sizes on referees and team performance from a North American perspective. *Psychology of Sport and Exercise*, 60, 102162.
<https://doi.org/10.1016/j.psychsport.2022.102162>
- Theriault, J. E., Young, L., & Barrett, L. F. (2021). The sense of should: A biologically-based framework for modeling social pressure. *Physics of Life Reviews*, 36, 100-136.
<https://doi.org/10.1016/j.plrev.2020.01.004>
- Turillo, C. J., Folger, R., Lavelle, J. J., Umphress, E. E., & Gee, J. O. (2002). Is virtue its own reward? Self-sacrificial decisions for the sake of fairness. *Organizational behavior and human decision processes*, 89(1), 839-865. [https://doi.org/10.1016/S0749-5978\(02\)00032-8](https://doi.org/10.1016/S0749-5978(02)00032-8)
- Unkelbach, C., & Memmert, D. (2010). Crowd noise as a cue in referee decisions contributes to the home advantage. *Journal of Sport and Exercise Psychology*, 32(4), 483-498.
<https://doi.org/10.1123/jsep.32.4.483>
- Vanderhasselt, M. A., Remue, J., Ng, K. K., Mueller, S. C., & De Raedt, R. (2015). The regulation of positive and negative social feedback: A psychophysiological study. *Cognitive, Affective, & Behavioral Neuroscience*, 15(3), 553-563.
<https://doi.org/10.3758/s13415-015-0345-8>
- Van Quaquebeke, N., & Giessner, S. R. (2010). How embodied cognitions affect judgments: Height-related attribution bias in football foul calls. *Journal of Sport and Exercise Psychology*, 32(1), 3-22. <https://doi.org/10.1123/jsep.32.1.3>

- Wagner-Egger, P., Gyga, P., & Ribordy, F. (2012). Racism in soccer? Perception of challenges of black and white players by white referees, soccer players, and fans. *Perceptual and motor skills*, 114(1), 275-289. <https://doi.org/10.2466/05.07.17.PMS.114.1.275-289>
- Wasylyshyn, N., Hemenway Falk, B., Garcia, J. O., Cascio, C. N., O'Donnell, M. B., Bingham, C. R., Simons-Morton, B., Vettel, J. M., & Falk, E. B. (2018). Global brain dynamics during social exclusion predict subsequent behavioral conformity. *Social Cognitive and Affective Neuroscience*, 13(2), 182-191. <https://doi.org/10.1093/scan/nsy007>
- Wood, S. N. (2021) *mgcv : Mixed GAM Computation Vehicle with Automatic Smoothness Estimation* (R package version 1.8-38) [Computer software]. <https://cran.r-project.org/src/contrib/Archive/mgcv/>
- Wood, S. N., Goude, Y., & Shaw S. (2015). Generalized additive models for large data sets. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 64(1), 139–155. <https://doi.org/10.1111/rssc.12068>
- Xiao, E., & Houser, D. (2005). Emotion expression in human punishment behavior. *Proceedings of the National Academy of Sciences*, 102(20), 7398-7401. <https://doi.org/10.1073/pnas.0502399102>
- Xu, M., Liu, B., Gu, R., Yang, S., Wang, H., & Zhu, X. (2021). Self-awareness buffers the consequences of negative feedback: Evidence from an ERP study. *International Journal of Psychophysiology*, 164, 9-16. <https://doi.org/10.1016/j.ijpsycho.2021.02.004>
- Yanase, K., Muto, S., Masuko, K., Nagami, T., & Ijiri, T. (2021, December). A system for practicing ball/strike judgment in VR environment. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology* (pp. 1-2). <https://doi.org/10.1145/3489849.3489931>

- Zayas, V., Wang, A. M., & McCalla, J. D. (2022). Me as good *and* me as bad: Priming the self triggers positive *and* negative implicit evaluations. *Journal of Personality and Social Psychology*, 122(1), 106–134. <https://doi.org/10.1037/pspp0000332>
- Zhang, D. (2020). Rsq: R-Squared and Related Measures. R package. <https://CRAN.R-project.org/package=rsq>

ANNEXE A – RÉPONSES DU COMITÉ D'ÉTHIQUE À LA RECHERCHE DE L'UQO À LA SUITE
DES DÉCLARATIONS DE PROJET DE RECHERCHE

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

ANNEXE B - SAGE CREATIVE COMMONS LICENSE: CONTRIBUTOR' S PUBLISHING
AGREEMENT

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

ANNEXE C – REPORT FOR THE PSYCHOLOGICAL SCIENCE MASS REPRODUCTION
INITIATIVE

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

Cette page a été volontairement laissée vide pour inclure les annexes en PDF

ANNEXE D – MATÉRIEL SUPPLÉMENTAIRE DE L'ARTICLE 2

Supplementary material 1

“Outside of wartime, sport is perhaps the only setting in which acts of interpersonal aggression are not only tolerated but enthusiastically applauded by large segments of society.” (Russell, 1993, p.181).

What behaviors trigger ejection?

According to the *Standards for removal from the game* listed in the Major League Baseball umpire manual (Major League Baseball, n.d.), several behaviors can lead to an ejection. For example, throwing equipment or shoving an umpire are physical aggressions subject to ejection. As far as verbal aggression is concerned, there are two concrete examples of unacceptable behavior:

- “Use of profanity specifically directed at an umpire or vulgar personal insults of an umpire are grounds for ejection” (Major League Baseball, n.d., p. 2);
- “Refusal to stop arguing, and further delaying the game after the umpire has provided a player or manager adequate opportunity to make a point, is a ground for ejection. The umpire should warn the player or manager that he has been heard and that he should return his position or be ejected” (Major League Baseball, n.d., p. 2).

The instructions given to MLB umpires suggest that players, managers, and coaches can argue with the official's decisions, as long as the criticism is not personal or persistent. These types of aggression can cause psychological harm to umpires, notably by damaging their self-esteem and reputation, and preventing them from carrying out their duties.

This is why an ejection for strike zone criticism cannot be attributed to a simple expression of disagreement with the decision of an authority figure. To support this point, here are a few quotes from professional baseball umpires that demonstrate that argumentation is tolerated, as long as it doesn't become excessive.

We suggest dividing the quotes into different themes: 1- tolerance towards criticism; 2- intolerance towards personal criticism; 3- intolerance towards persistent criticism. Obviously, the categories are not mutually exclusive, and some quotes could have been placed in more than one sub-section. Also, for the benefit of the neophyte reader, the expressions “running someone” and “dumping” means to eject the person from the game.

1- Tolerance towards criticism

- “Arguing with managers and players is just part of the game. There’s nothing wrong with that. Every fan wants to see his home team manager come out and argue. When Lasorda came out and argued, he knew what he could say and what he couldn’t say. He was a big league guy. And still is.” (Clark & Schlossberg, 2014, p. 102).
 - Umpire Al Clark

- “And people ask, “How could you take the swearing?” Are you kidding me? [...] Swearing is the language of baseball. [...] The word *fuck* becomes an all-purpose weapon: a noun, verb, adjective, and adverb all rolled into one. [...] I can’t dump a player for something like that. The last thing an umpire wants is a cheap ejection. After all, players are allowed to argue a call or do a minimal amount of bitching from the dugout.” (Postema & Wojciechowski, 2003, p. 197-198).
- Umpire Pam Postema
- “It’s really cut and dry. As long as you’re just arguing the pitch and you’re saying that you think I’m wrong and it’s this -- you can use the foulest language in the world and that doesn’t matter [...] It’s baseball. And it’s grown men playing baseball. We’re not in church. We’re on a baseball field. So you can say whatever you wanna say about that pitch, that call, that anything. It’s when you direct it and get personal. Typically, it’s when you put the word you in front of it.” (Darling & Rollins, 2022)
- Umpire Fieldin Culbreth
- “Bobby may have used salty language but I don’t think most umpires cared. That’s part of our culture. Four- letter words and combinations of are part of the culture. It’s not necessarily church out there on the field. It’s competition. We all are of a competitive mind. We wear athletic supporters. And with that comes salty language. If a player says something to the effect of, “That’s a horseshit call. How could you make a stupid fucking call like that?” that’s no big deal. That’s not an issue whatsoever— at least it wasn’t for me and most of the umpires I worked with. What is a big deal is when it becomes personal. So if a guy wants to swear, that’s fine. If a guy wants to swear about a situation, that’s fine. I have no issue with that. But if he gets personal, and swears at me, he will be ejected.” (Clark & Schlossberg, 2014, p. 31-32).
- Umpire Al Clark
- “Some guys will spend a whole game whining about this pitch, whining about that pitch, making umpires insane. They don’t do enough to get run, but they’re always in your ear.” (Scott & Neyer, 2022, p. 422).
- Umpire Dale Scott
- “If Williams had a gripe with an umpire, he’d quietly look at the ground and say, “You can do better than that.” Or, “That was an inch outside.” He never stood up an umpire and so he never got run.” (Nowlin, 2020, p. 234).
- Author Bill Nowlin
- “He would come out and argue, but he was one of those managers who never wanted to come out of the ball game. He wanted to be on the field, so he kept his cool to keep from getting ejected.” (Harvey & Golenbock, 2014, p 202).
- Umpire Doug Harvey

- “We don’t want to make too much of ejections. They happen. They offer a little drama. It’s almost always obvious why ejection occurred.” (Nowlin, 2020, p. 259).
 - Umpire Vic Carapazza
- “Vic probably spoke for most umpires when he said, “I don’t like throwing anybody out of any game, but I’ll do it if I have to do it. You warn him a couple of times and then, you know, you leave it in *their* hands. Let them decide what they want to do”. (Nowlin, 2020, p. 261).
 - Umpire Vic Carapazza

2- Intolerance towards personal criticism

- “Fans ask all the time, “What is the magic word a player, coach, or manager says to be ejected?” The answer is not what the fans might think. Most think that anytime a player uses a vulgarity he’s in danger of ejection. That might be true but it is not automatic. What is automatic during the course of an argument is when a player, coach, or manager prefaces anything with the word “you.” The umpire’s awareness heightens because anything that follows “you” in an argument is going to be personal. Without doubt an ejection will follow every time.” (Clark & Schlossberg, 2014, p. 31).
 - Umpire Al Clark
- “People always want to know whether there’s a magic word. Is it ‘cocksucker’? Is it ‘asshole’? No. The magic word is ‘you.’” (Weber, 2009, p. 99).
 - Umpire Gary Cederstrom
- “A player, manager, or coach can basically cuss all they want (with a few exceptions), but when they personalize their comments, they risk being ejected. Starting a complaint with You’re [...] is often all it takes. In other words, he might think and say the call is full of shit, but when he says I am? Buh-bye.” (Scott & Neyer, 2022, p. 215).
 - Umpire Dale Scott
- “MLB umpires are pretty generous with warnings before they throw someone out, but when it comes to foul language—especially if it’s an insult directed at someone specific—they’ll occasionally skip the heads-up. “Anything following ‘You’ is very important. If someone says ‘You,’ whatever they say after that can be pretty bad,” Everitt said. “That’d be an automatic ejection.” (Gutoskey, 2021).
 - Umpire Mike Everitt
- “Getting more frustrated as the argument went on, he yelled, “You’re all a bunch of cocksuckers!” Three of the four umpires simultaneously ejected him, each gesturing in a different direction. Fregosi immediately pointed his finger at the one umpire who didn’t throw him out and screamed, “It must be you! You’re the cocksucker!” (Scott & Neyer, 2022, p. 422-423).
 - Umpire Dale Scott

- “The way he argued and the way he disagreed really was degrading, so I ejected him immediately.” (Clark & Schlossberg, 2014, p. 40).
- Umpire Al Clark
- “Well, the next day in the Jackson paper, Mitchell was quoted extensively: “The pitch was obviously low, so I just told him there’s no way you can call that pitch a strike.” [...] like he’d been calmly discussing pitch location and I was an autocratic jerk [...] to throw him out. He’d made up the entire conversation. What Mitchell really said was the pitch was fucking low, I’d been fucking him all day, fuck you; so I ejected him.” (Scott & Neyer, 2022, p. 223-224).
- Umpire Dale Scott
- “Davis stormed out of the dugout and headed right toward me. Instead of talking or yelling about the play, he said, “You’re nothing but a fucking Hitler.” I ejected him immediately, then walked away, leaving him to argue with no one. He finally left the field and we finished the game. The next day, Davis said to my partner, Jerry Young, “Why did Clark run me so fast last night? All I did was call him ‘a fucking Hitler.’” Young started to laugh uncontrollably, much to Davis’s dismay. After regaining his composure Young told Davis I was in fact Jewish. It took Davis another two months, but eventually he did apologize and say, “That kind of shit is out of bounds.” (Clark & Schlossberg, 2014, p. 91).
- Umpire Al Clark

3- Intolerance towards persistent criticism

- “More commonly, he explained, ejections happen when a manager simply won’t stop arguing. “You’d say, ‘Hey, Tom, listen, I understand where you’re coming from. I’ve heard your argument—now we’re starting to repeat ourselves. We’re not gonna reverse the call here. I’ve heard your argument and we need to move on. This conversation’s gonna end.” If Tom refuses to end the conversation, he might find himself getting ushered off the field post-haste.” (Gutoskey, 2021).
- Umpire Mike Everitt
- “Short of an ejection, you need to let chirpers know early: We’re not doing this for nine innings, or somebody is leaving early, and I’m here all day. They’ve been warned, so now it’s their decision. If the whining continues, they ran themselves.” (Scott & Neyer, 2022, p. 423).
- Umpire Dale Scott
- “[...] But as far as arguing the call, of there’s any kind of language we pretty much let go that kind of stuff. My thing is, and I always tell guys, I always try to let the initial emotion of a play go – like especially now – mostly balls and strikes. So if I call a guy out on strike 3 and he gives me a shout or says something, like that F’ing pitch was off the plate or something like that and now he walks back to the dugout, no problem. Once he gets back at the dugout he needs to let it go. [...] if he wants

to pop off again he's probably going to get ejected because I gave him his say, put a warning in, and alright, that's it, let's go. If he goes back to the dugout and he wants to continue, now I'm trying to go back to work and get the next pitch right, and I've got him screaming at us from the dugout." (Nowlin, 2020, p. 261).

- Umpire Sam Holbrook

- "Finally, after signaling to the press box the game was under protest, I said to Gibbons, "John, the run scores, your protest is lodged, and now it's time to get this thing going." He continued to argue, so I had to ask him, "Do you want to get run? It's time to move on." During the entire argument, John never cussed me or got personal, nothing to get ejected. But there are times when there's nothing left to argue, you've explained what you're going to do, and the game needs to continue. Now it's either get run or get back in your dugout. Because we're done here." (Scott & Neyer, 2022, p. 507).

- Umpire Dale Scott

Why is an ejection for arguing balls and strikes considered to be aggression?

Players and coaches have the right to challenge umpires' decisions, as long as the arguments do not exceed the limits set out in the MLB umpire's manual. Arguing behaviors considered off-limits meet the definition of aggression, as recognized in the field of social psychology.

Actually, the concept of aggression, defined in different ways by the scientific community over time, seems to have recently found a common definition among researchers in social psychology: "A behavior that is intended to harm another person who is motivated to avoid that harm" (Allen & Anderson, 2017, p. 1). Aggression can be expressed in a number of different response modalities, such as verbal aggression, which is the subject of this research project. Examples of verbal aggression include name calling, swearing, or screaming (Allen & Anderson, 2017).

Aggression can also be subdivided according to the objective pursued by the perpetrator. Hostile aggression is motivated by the desire to hurt someone, and is impulsive behavior; instrumental aggression, on the other hand, is premeditated behavior used as a means to achieve a specific goal (Bushman & Anderson, 2001). According to Krahé (2020), both forms of aggression can be observed in sports competitions. Participants use hostile aggression when their negative emotions get the better of them, while they use instrumental aggression to improve their chances of winning.

When players and coaches excessively criticize umpires' decisions, it can be difficult to determine the type of aggression used. Indeed, participants may direct insults at the umpire or be unable to stop criticizing their work because they have lost control of their emotions. That would be considered hostile aggression. Participants may also insult the umpire or continue their arguments at extended length in a thoughtful, controlled manner, with the aim of influencing the umpire's future decisions (Warneke & Ogden, 2012). In that case, it would be considered instrumental aggression.

The manner in which ejection reasons are listed in the MLB database and on the *Restrosheet* website (“Called third strike”, “Balls and strikes”, “Arguing strike one call”) makes it impossible to determine the type of aggression that led to the ejection. Although it is an interesting question, the distinction between the two types of aggression is not quite relevant within the context of our study nor does it impact its outcome. The aim is to determine whether verbal aggression, regardless of the perpetrator's objective, leads to an ejection and whether that ejection influences the umpires' subsequent decision-making. In the context of this study, all that needs to be considered is the definition of aggression, and the idea that an ejection for contesting a strike zone decision meets that definition, whether it be hostile or instrumental in nature.

References

- Allen, J. J., & Anderson, C. A. (2017). Aggression and violence: Definitions and distinctions. *The Wiley handbook of violence and aggression*, 1-14.
- Bushman, B. J., & Anderson, C. A. (2001). Is it time to pull the plug on hostile versus instrumental aggression dichotomy?. *Psychological review*, 108(1), 273.
<https://doi.org/10.1037/0033-295X.108.1.273>
- Clark, A., & Schlossberg, D. (2014). *Called Out But Safe: A Baseball Umpire's Journey*. U of Nebraska Press.
- Darling, R., & Rollins, J. (Hosts). (2022, July 6). How to talk to Umpires feat. Fieldin Culbreth (Episode 7) [Audio podcast episode]. In *Unwritten: Behind Baseball's Secret Rules*. Audacy. <https://podcasts.apple.com/us/podcast/episode-7-how-to-talk-to-umpires-feat-fieldin-culbreth/id1624770386?i=1000568975074>
- Gutoskey, E. (2021, May 17). 10 Behind-the-Plate Secrets of MLB Umpires. *Mental Floss*. <https://www.mentalfloss.com/article/646427/mlb-umpires-secrets>
- Harvey, D., & Golenbock, P. (2014). *They Called Me God: The Best Umpire Who Ever Lived*. Simon and Schuster.
- Krahé, B. (2020). *The social psychology of aggression*. Routledge.
- Major League Baseball. (n.d.). Conduct and Responsibilities of Umpires In *Major League Baseball umpire manual*. New York: Office of the Commissioner of Baseball.
- Nowlin, B. (2020). *Working a “Perfect Game”: Conversations With Umpires*. Summer Game Books Edition.
- Postema, P., & Wojciechowski, G. (2003). *You've got to have balls to make it in this league: My life as an umpire*. U of Nebraska Press.

- Russel, G. W. (1993). *The social psychology of sport*. New York: Springer.
- Scott, D., & Neyer, R. (2022). *The Umpire Is Out: Calling the Game and Living My True Self*. U of Nebraska Press.
- Warneke, K., & Ogden, D. (2012). The right call: Baseball coaches' attempts to influence umpires. *Great Plains Research*, 137-145. <https://www.jstor.org/stable/23780025>
- Weber, B. (2009). *As they see'em: A fan's travels in the land of umpires*. Simon and Schuster.

Supplementary material 2

Generalizing the sports environment to other contexts

The sporting environment is an ideal ecological context in which to observe a variety of human behaviors. In certain fields, such as management, sports scenarios are used extensively to study phenomena related to human relations. To this end, a recent literature review of 249 studies carried out over the last 50 years establishes the numerous advantages of relying on the sporting context as a living laboratory for testing and developing theories related to human behavior (Fonti, Ross & Aversa, 2023).

The sporting context can also be used in other fields, such as social psychology. Sports performance judgment, a concrete example of human social decision-making, is one of the many behaviors that can be studied in sports environments. According to Plessner and Haar (2006), “ample anecdotal and empirical evidence indicates that sports performance judgments are at least as prone to systematic errors (biases) as other social judgments” (p. 555). According to these authors, since sports decision-makers use social information processing to analyze their environment, biases associated with this process are generalizable to a variety of social situations.

It is thus possible to raise the hypothesis that people with a roles and responsibilities similar to that of baseball umpires are at risk of being affected by the same biases but in their own social situation (outside of the sporting context). To give weight to our hypothesis, we propose to draw a parallel with actors who share similarities with baseball umpires. We considered that people in positions of authority who must apply the rules impartially may be subject to the same social decision-making bias as baseball umpires. We will first clarify the role of umpires as authority figures, and then explain how this role is comparable to that of other decision-makers.

Baseball umpire as authority figures.

Baseball umpire act as authority figure on the baseball field. According to the Major League Baseball official rules, “[t]he umpires shall be responsible for the conduct of the game in accordance with these official rules and for maintaining discipline and order on the playing field during the game” (Major League Baseball, 2022, p. 95). It is their duty to be impartial and apply the rules uniformly to everyone involved in a baseball game. For a baseball umpire, this means that each play must be considered as a unique situation and that all other external elements (i.e. the recurrence of interaction with the same participants over the course of a game or a season) must be ignored in their decision-making. To this extent, MLB umpires are considered “masters of leaving the past in the past and focusing on whatever’s right in front of them, be it the next game or the very next pitch.” (Gutoskey, 2021, p. 1).

Three key examples of this philosophy can be found in the MLB Official Rules and the MLB Umpire Manual.

- “MLB Umpires should be focused on every pitch of every game without regard to factors such as the score, inning, weather, or standing of the teams involved.” (Major League Baseball, n.d., p. 1).
- “You no doubt are going to make mistakes, but never attempt to “even up” after having made one. Make all decisions as you see them and forget which is the home or visiting Club.” (Major League Baseball, 2022, p. 100).
- “MLB Umpires should make every decision based on the circumstances and facts presented, regardless of an umpire’s past history with a particular player or club. No umpire should ever threaten a player, manager, coach, or club with future retaliation”. (Major League Baseball, n.d., p. 5).

Nevertheless, as suggested in the main manuscript, umpires are human beings and are at risk of cognitive and social bias. This is also the case for other people in positions of authority, who are tasked with making impartial decisions but are confronted with the limits of human cognition. Those responsible for applying the rules are expected to treat each situation uniquely, based solely on the rules in place and not on previous events. Like umpires, they find themselves in a role where expectations are utopian: to make perfectly impartial decisions to ensure justice between the people affected by those decisions. We propose that three other decision-makers with roles similar to those of the umpires are confronted with this utopia and may be affected by the bias proposed in our main manuscript.

Similarities between umpiring and other jobs

Judges

“Judges are often called the umpires of the law because they make sure their courtrooms are fair for everyone who comes there seeking justice” (Aulbur, n.d., p. 1). This metaphor has been widely used, ever since John G. Roberts, Jr. was nominated Chief Justice of the United States in 2005. In his testimony, Chief Roberts stated: “Judges are like umpires. Umpires don't make the rules, they apply them. The role of an umpire and a judge is critical. They make sure everybody plays by the rules.” (Administrative Office of the U.S. Courts, 2005, p.1). The image was so strong that the metaphor almost became a judicial golden rule (McKee, 2006).

While most agree that the functions of judge and umpire are not identical, they do have many points in common. Indeed, according to Kavanaugh (2015, pp. 685-692), they are similar in at least 10 different ways (see the full article for a detailed explanation of each comparison):

1. They can't be partisan.
2. They must follow established rules and principles.
3. They must be consistent.
4. They must apply the rules, not rewrite them.
5. They must have some backbone.

6. They must ignore crowd noise.
7. They must have an open mind.
8. They must behave appropriately.
9. They need to be collegial.
10. They need to explain their decisions clearly.

More importantly, judges and umpires are constrained by the boundaries of human cognition (Peer & Gamliel, 2013). Although the question of whether judges and umpires have exactly the same biases remains empirical, awareness of potential biases is essential to improve decision-making. Our results identify one of these potential biases and could thus be used to improve decision-making outside the sporting context.

Police officers

“Some umpires like to think of themselves as a police officer responsible for keeping order in a no-nonsense manner” (Porter, n.d., p. 1). Ed Hickox, former MLB umpire and police officer, recognizes that the two professions are not so different. He said that “[t]here are times in law enforcement you have to make split-second decisions, you have to write reports, you have to diffuse situations, you have to handle people” (Metz, 2021).

The comparison extends to decisions such as ejecting a player or detaining a civilian. Retired MLB umpire Jim Evans suggested that umpires, like police officers, don't necessarily like to resort to punishment, but sometimes have to when people step out of line: “[w]e don't get paid any extra to throw someone out; they're going to appeal it to the union, for one thing. So, there's a delicate balance there. Sort of like a police officer: when do you ticket a person or jail him?” (Chapman, 2016, p. 1).

Like umpires and judges, police officers don't make the laws but they must enforce them. The public expects police officers to be impartial and unbiased (Jackson & Pósch, 2019; National Research Council, 2004). Just as impartial and unbiased umpires seem utopian, police officers are likely to make mistakes, especially when faced with split-second decisions. Implicit racial prejudice is a good example: “A police officer's decision to stop another vehicle, for example, often requires a split-second decision. Several studies claim that due to implicit attitudes, police officers are more likely to stop African-American drivers than white drivers. Just like the police officer's decision to stop another driver, the umpire's decision to whistle a foul leaves very little time to think” (MCann, 2009, p. 4).

Business managers

Similar to umpires, business managers are also in a position of authority. They don't set the rules, because they're not at the top of the company, but they do have to apply them to the employees below them in the hierarchy. Both are equally accountable: they are evaluated by their superiors on how they apply company rules to employees (Capelle, 2013) or league rules to game participants (Gutoskey, 2021; Rogers, 2022).

This undoubtedly leads to dissatisfaction and questions about the impartiality of decisions. Just as the public expects judges, police officers, and umpires to be fair and impartial, employees also expect their managers to be just (Colquitt & al., 2001; Robbins, Ford & Tetrick, 2012). As human decision-makers, business leaders are subject to human biases similar to those of umpires.

Mills (2014) proposes that MLB umpires can be affected by inequality aversion, status, and mere exposure biases. He claims that his findings extend to the workplace, i.e. to relations between employees and managers, with a potential impact on the fairness of pay and promotion decisions. Parsons et al. (2011) found racial bias among MLB umpires and also suggest that this is generalizable to the employee-supervisor relationship in the labor market.

Conclusion

Like other authority figures, umpires are only human, limited by human cognition (Peer & Gamliel, 2013). We believe that the nature of the job of judges, police officers and managers may make them susceptible to the same decision-making biases as baseball umpires. It is therefore plausible that, after being verbally abused, these people seek to avoid experiencing this social punishment again, and are more likely to be lenient, making unjust decisions. However, this possibility requires further study to validate the hypothesis.

Even after being aggressively criticized by someone, authority figures have a responsibility to forget what happened and apply the rules impartially to others. These people need to be aware of the bias of MLB umpires in order to avoid being influenced in the same way. It is worse to try balancing decisions and appearing even more unfair than forgetting or accepting your mistake (whether real or not but perceived as such by someone else) and try to become impartial again. The following quote is a good illustration of how injustice is amplified when umpires (or people in positions of authority) try to correct a perceived injustice:

“Don’t try to “even up” a bad call. Occasionally you’re going to accidentally make a bad call, and you’ll know it. [...] When you’ve kicked a call, even if the coach comes out and gives you an earful, you have to live with the bad call. Everybody inadvertently makes a mistake now and then. However, the worst thing you can do is deliberately make another bad call to favor the other team to “even things up.” You just got rid of the first ranting coach. Now you’ll have the other team’s coach out there angry with you, too! Most reasonable coaches expect that the umpires will miss one occasionally, but nobody expects or wants to see an umpire to do it willfully.” (Martingrove Baseball, n.d., p. 1).

References

Administrative Office of the U.S. Courts. (2005). *Chief Justice Roberts Statement - Nomination Process*. <https://www.uscourts.gov/educational->

resources/educational-activities/chief-justice-roberts-statement-nomination-process

- Aulbur, M. (n.d.). *Judges – Umpires of the law*. Missouri Bar.
<https://missourilawyershelp.org/lessons-plans/judges-as-umpires-of-the-law/#:~:text=Judges%20are%20often%20called%20the,who%20comes%20there%20seeking%20justice>
- Capelle, R. G. (2013). *Optimizing organization design: A proven approach to enhance financial performance, customer satisfaction and employee engagement*. John Wiley & Sons.
- Chapman, M. M. (2016, August 28). Hey, ump! Why'd you eject that player or manager? *Detroit Free Press*.
<https://www.freep.com/story/sports/mlb/tigers/2016/08/28/major-league-umpiring/89516532/>
- Fonti, F., Ross, J. M., & Aversa, P. (2023). Using sports data to advance management research: A review and a guide for future studies. *Journal of Management*, 49(1), 325-362. <https://doi.org/10.1177/01492063221117525>
- Gutoskey, E. (2021, May 17). 10 Behind-the-Plate Secrets of MLB Umpires. *Mental Floss*. <https://www.mentalfloss.com/article/646427/mlb-umpires-secrets>
- Gerlach, L. R., & Nowlins, B. (2017). *The SABR book of umpires and umpiring*. Society for American Baseball Research, Inc.
- Harvey, D., & Golenbock, P. (2014). *They Called Me God: The Best Umpire Who Ever Lived*. Simon and Schuster.
- Jackson, J., & Pósch, K. (2019). New directions of research in fairness and legal authority. *Social Psychology and Justice (Frontiers of Social Psychology Series, Chapter 9)*. Routledge.
- Kavanaugh, B. M. (2015). The judge as umpire: Ten principles. *Cath. UL Rev.*, 65, 683.
<https://scholarship.law.edu/lawreview/vol65/iss3/12>
- Major League Baseball. (2022). *Official Baseball Rules*. New York: Office of the Commissioner of Baseball.
- Major League Baseball. (n.d.). Conduct and Responsibilities of Umpires In *Major League Baseball umpire manual*. New York: Office of the Commissioner of Baseball.

- Martingrove Baseball. (n.d.). *50 Tips for a New Umpire*. <https://www.martingrove.ca/50-tips-for-a-new-umpire/>
- McCann, M. (2009). Using Social Psychology to Evaluate Race and Law in Sports. *Vermont Law School Research Paper*, (10-04). <https://ssrn.com/abstract=1440938>
- McKee, T. A. (2006). Judges as Umpires. *Hofstra L. Rev.*, 35, 1709. <https://scholarlycommons.law.hofstra.edu/hlr/vol35/iss4/3>
- Metz, C. (2021, December 20). Full-time MLB umpire, part-time Daytona Beach Shores detective. *Hearst Television Inc.* <https://www.wesh.com/article/umpire-part-time-daytona-beach-shores-detective/38571461>
- National Research Council. (2004). *Fairness and effectiveness in policing: The evidence*. National Academies Press.
- Parsons, C. A., Sulaeman, J., Yates, M. C., & Hamermesh, D. S. (2011). Strike three: Discrimination, incentives, and evaluation. *American economic review*, 101(4), 1410-1435. <https://doi.org/10.1257/aer.101.4.1410>
- Peer, E., & Gamliel, E. (2013). Heuristics and biases in judicial decisions. *Ct. Rev.*, 49, 114. <https://digitalcommons.unl.edu/ajacourtreview/422>
- Plessner, H., & Haar, T. (2006). Sports performance judgments from a social cognitive perspective. *Psychology of sport and exercise*, 7(6), 555-575. <https://doi.org/10.1016/j.psychsport.2006.03.007>
- Porter, J. (n.d.). Many Roles of the Sports Officials. *Mid-Atlantic Collegiate Umpires*. <https://macumpires.org/many-roles-of-the-sports-official/>
- Rogers, J. (2022, August 17). MLB Umpires Association, adamant that crews are simply upholding rule, address home-plate collision rule in statement. *ESPN Entreprises*. https://www.espn.com/mlb/story/_/id/34415630/mlb-umpires-association-adamant-crews-simply-upholding-rule-address-home-plate-collision-rule-statement

Fig. S1. Major League Baseball strike zone. Source: Dimensions.com.

Dimensions.Guide | Baseball Strike Zone

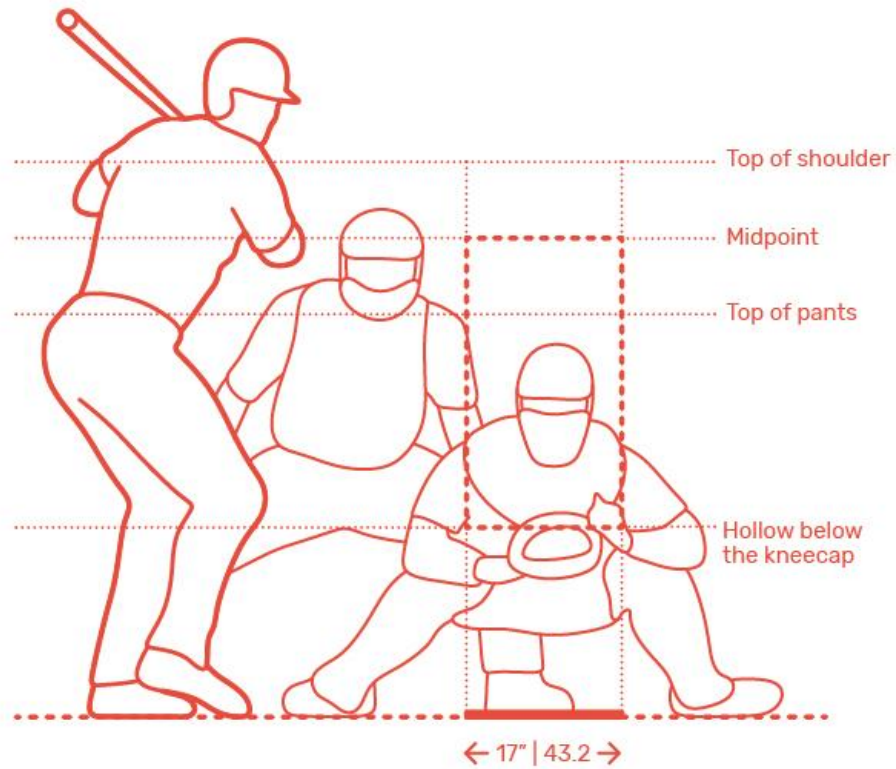


Fig. S2. Receiver operating characteristics (ROC) curve from baseline GAM. The area under the curve (AUC) is 96.98%.

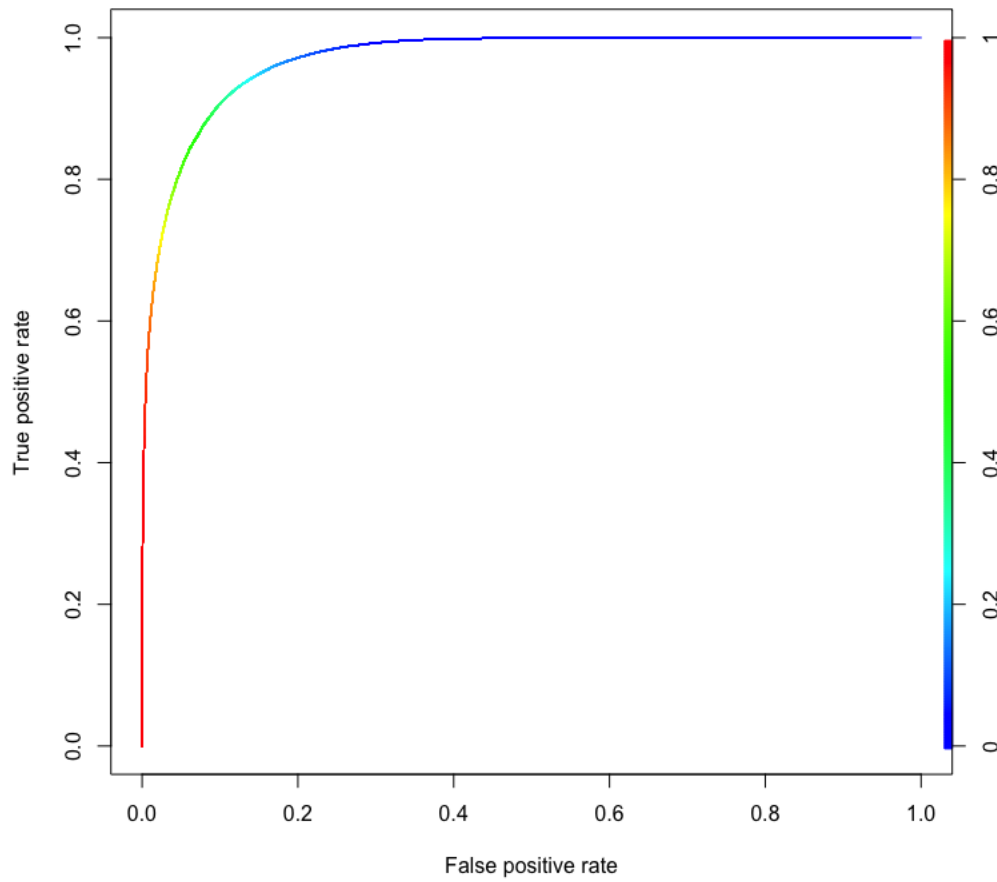


Fig. S3. RHB vs LHP called strike probability for different counts. RHB: Right-handed batter; LHP: Left-handed pitcher. Counts are presented in the upper left corner in the following way: Numbers of balls-Number of strikes. The strike zone as defined by the MLB rules is shown in dotted line. The x-axis represents the horizontal position of the ball when it reaches home plate. The x-axis represents the horizontal position of the ball, in inches, when it reaches home plate. Reference point 0 is the centre of home plate, as seen from the umpire's position. The y-axis represents the vertical position of the ball, in inches, when it reaches home plate. Reference point 0 (not visible in the figure) is the ground.

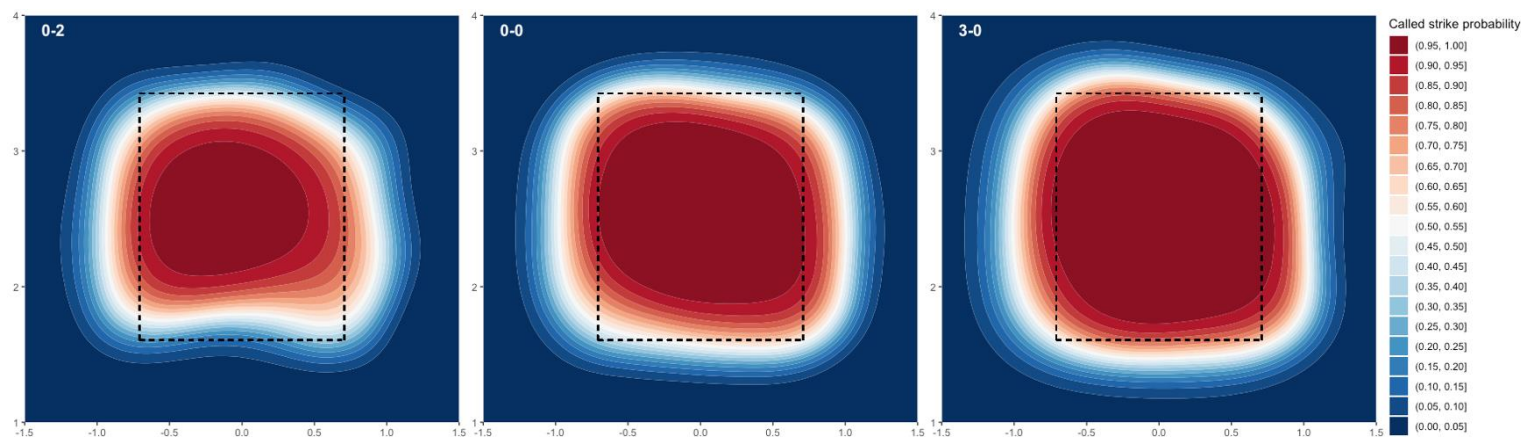


Fig. S4. RHB vs RHP called strike probability for different counts. RHB: Right-handed batter; LHP: Left-handed pitcher. Counts are presented in the upper left corner in the following way: Numbers of balls-Number of strikes. The strike zone as defined by the MLB rules is shown in dotted line. The x-axis represents the horizontal position of the ball when it reaches home plate. The x-axis represents the horizontal position of the ball, in inches, when it reaches home plate. Reference point 0 is the centre of home plate, as seen from the umpire's position. The y-axis represents the vertical position of the ball, in inches, when it reaches home plate. Reference point 0 (not visible in the figure) is the ground.

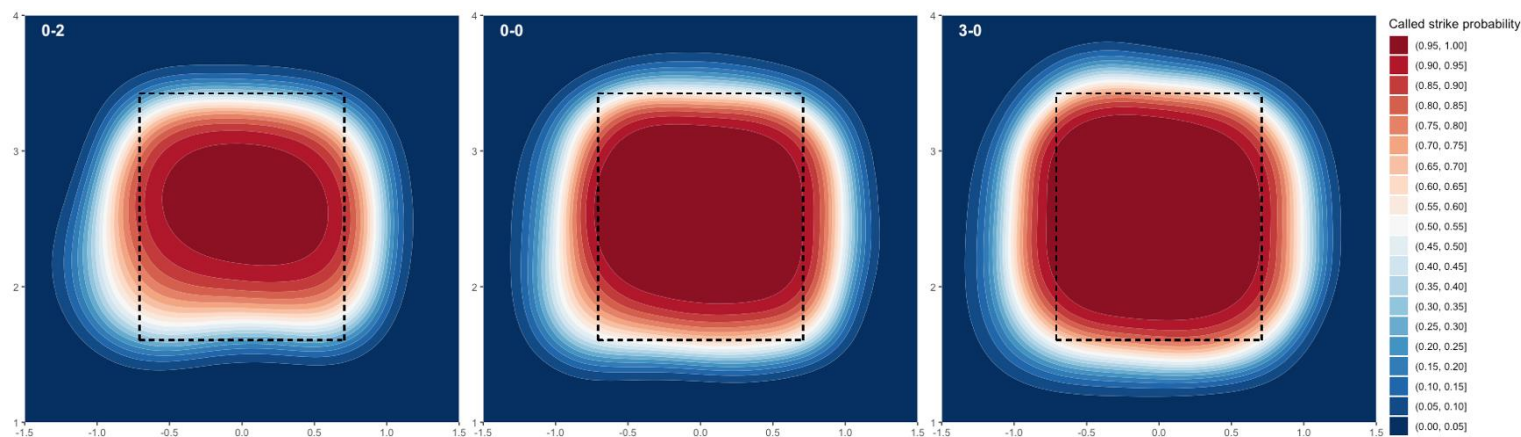


Fig. S5. LHB vs LHP called strike probability for different counts. RHB: Right-handed batter; LHP: Left-handed pitcher. Counts are presented in the upper left corner in the following way: Numbers of balls-Number of strikes. The strike zone as defined by the MLB rules is shown in dotted line. The x-axis represents the horizontal position of the ball when it reaches home plate. The x-axis represents the horizontal position of the ball, in inches, when it reaches home plate. Reference point 0 is the centre of home plate, as seen from the umpire's position. The y-axis represents the vertical position of the ball, in inches, when it reaches home plate. Reference point 0 (not visible in the figure) is the ground.

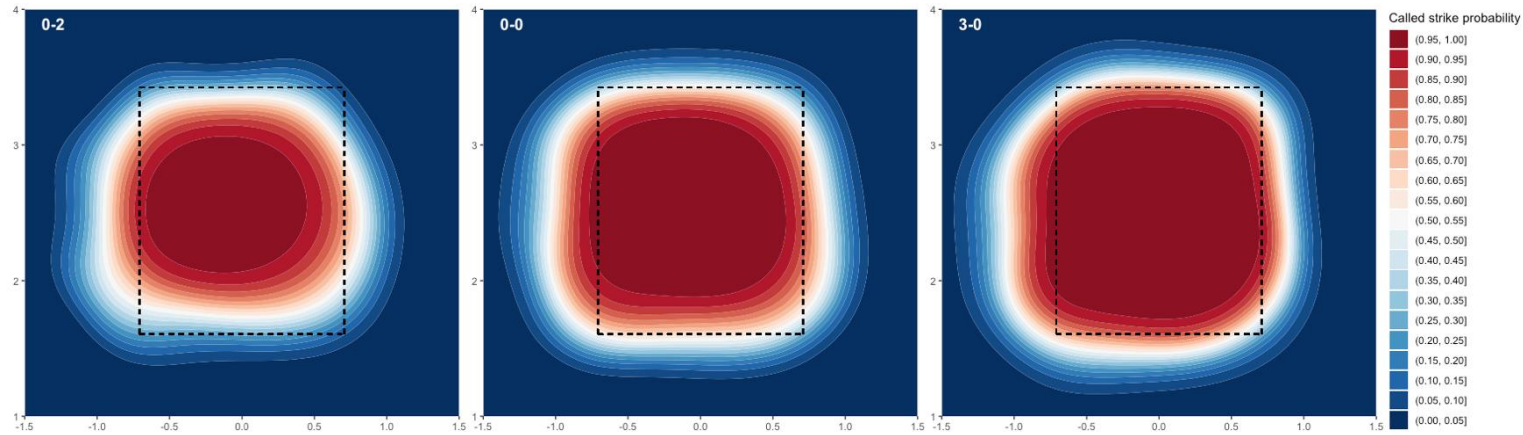


Fig. S6. LHB vs RHP called strike probability for different counts. RHB: Right-handed batter; LHP: Left-handed pitcher. Counts are presented in the upper left corner in the following way: Numbers of balls-Number of strikes. The strike zone as defined by the MLB rules is shown in dotted line. The x-axis represents the horizontal position of the ball when it reaches home plate. The x-axis represents the horizontal position of the ball, in inches, when it reaches home plate. Reference point 0 is the centre of home plate, as seen from the umpire's position. The y-axis represents the vertical position of the ball, in inches, when it reaches home plate. Reference point 0 (not visible in the figure) is the ground.

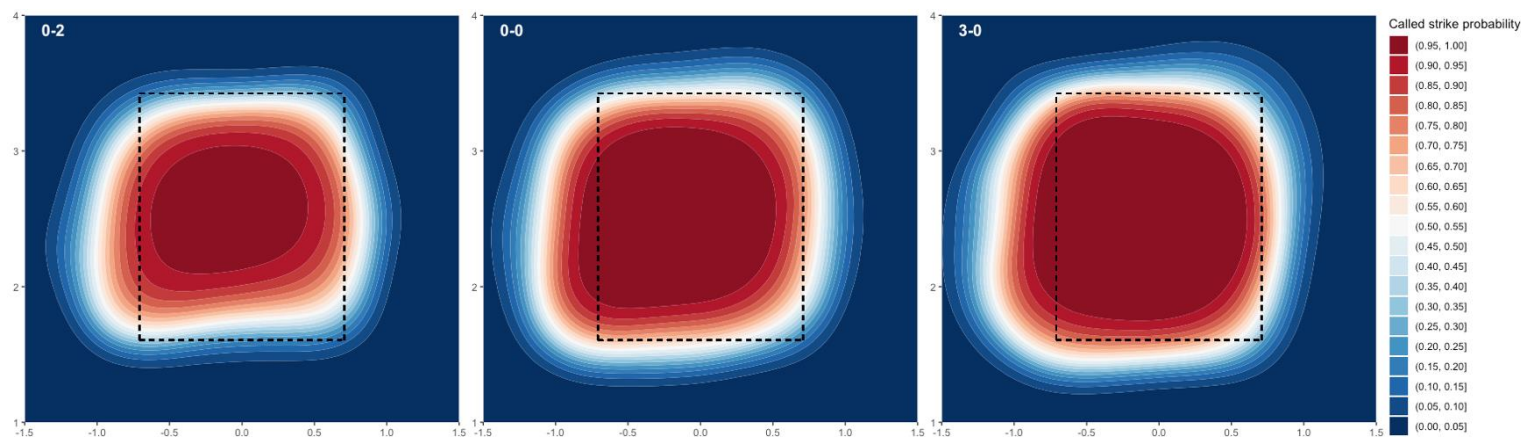


Fig. S7. RHB vs LHP called strike probability for different years. RHB: Right-handed batter; LHP: Left-handed pitcher. Seasons are presented in the upper left corner. The strike zone as defined by the MLB rules is shown in dotted line. The x-axis represents the horizontal position of the ball when it reaches home plate. The reference point is the centre of home plate, as seen from the umpire's position. The x-axis represents the horizontal position of the ball, in inches, when it reaches home plate. Reference point 0 is the centre of home plate, as seen from the umpire's position. The y-axis represents the vertical position of the ball, in inches, when it reaches home plate. Reference point 0 (not visible in the figure) is the ground.

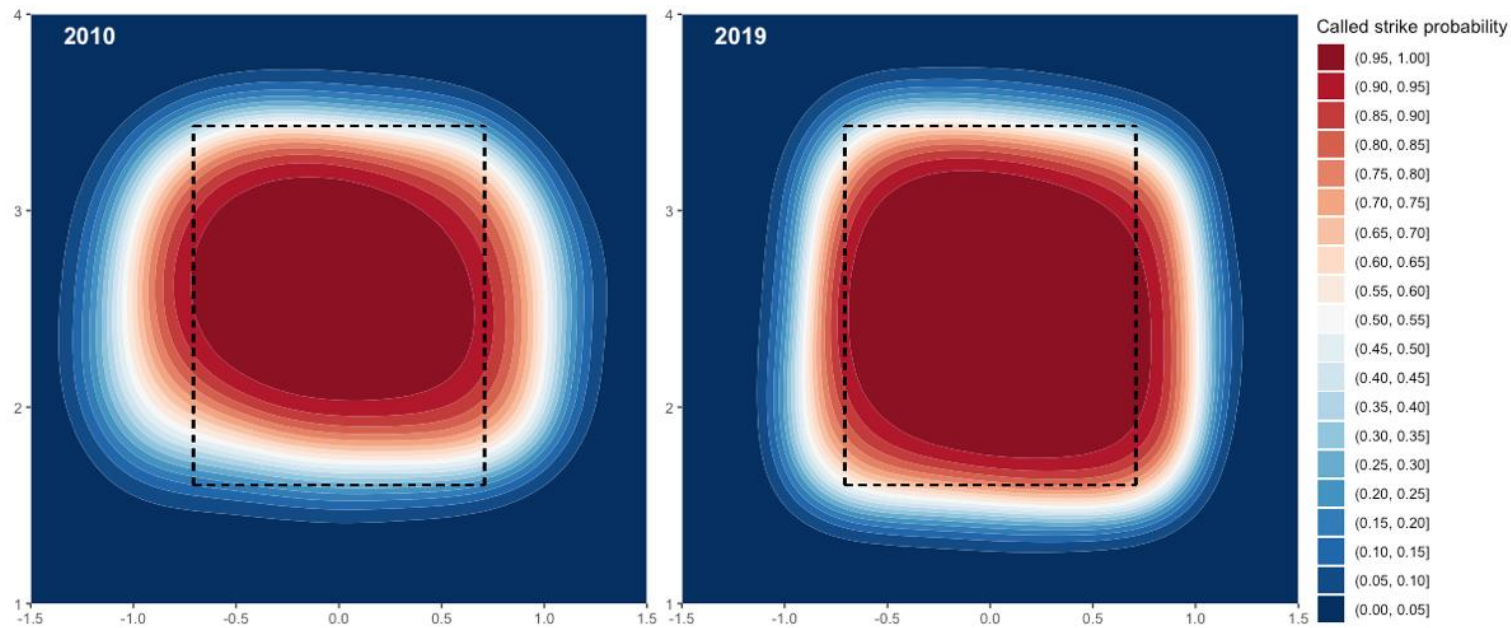


Fig. S8. RHB vs RHP called strike probability for different years. RHB: Right-handed batter; LHP: Left-handed pitcher. Seasons are presented in the upper left corner. The strike zone as defined by the MLB rules is shown in dotted line. The x-axis represents the horizontal position of the ball when it reaches home plate. The reference point is the centre of home plate, as seen from the umpire's position. The x-axis represents the horizontal position of the ball, in inches, when it reaches home plate. Reference point 0 is the centre of home plate, as seen from the umpire's position. The y-axis represents the vertical position of the ball, in inches, when it reaches home plate. Reference point 0 (not visible in the figure) is the ground.

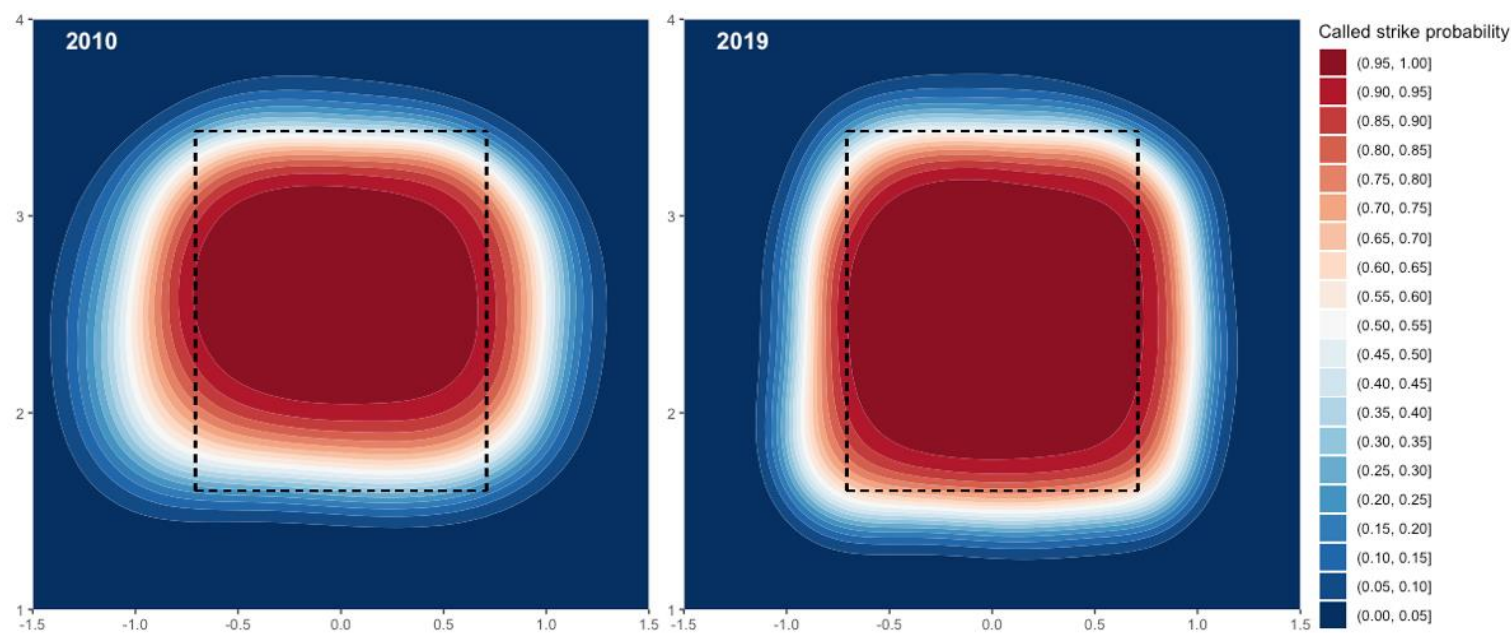


Fig. S9. LHB vs LHP called strike probability for different years. RHB: Right-handed batter; LHP: Left-handed pitcher. Seasons are presented in the upper left corner. The strike zone as defined by the MLB rules is shown in dotted line. The x-axis represents the horizontal position of the ball when it reaches home plate. The reference point 0 is the centre of home plate, as seen from the umpire's position. The x-axis represents the horizontal position of the ball, in inches, when it reaches home plate. Reference point 0 is the centre of home plate, as seen from the umpire's position. The y-axis represents the vertical position of the ball, in inches, when it reaches home plate. Reference point 0 (not visible in the figure) is the ground.

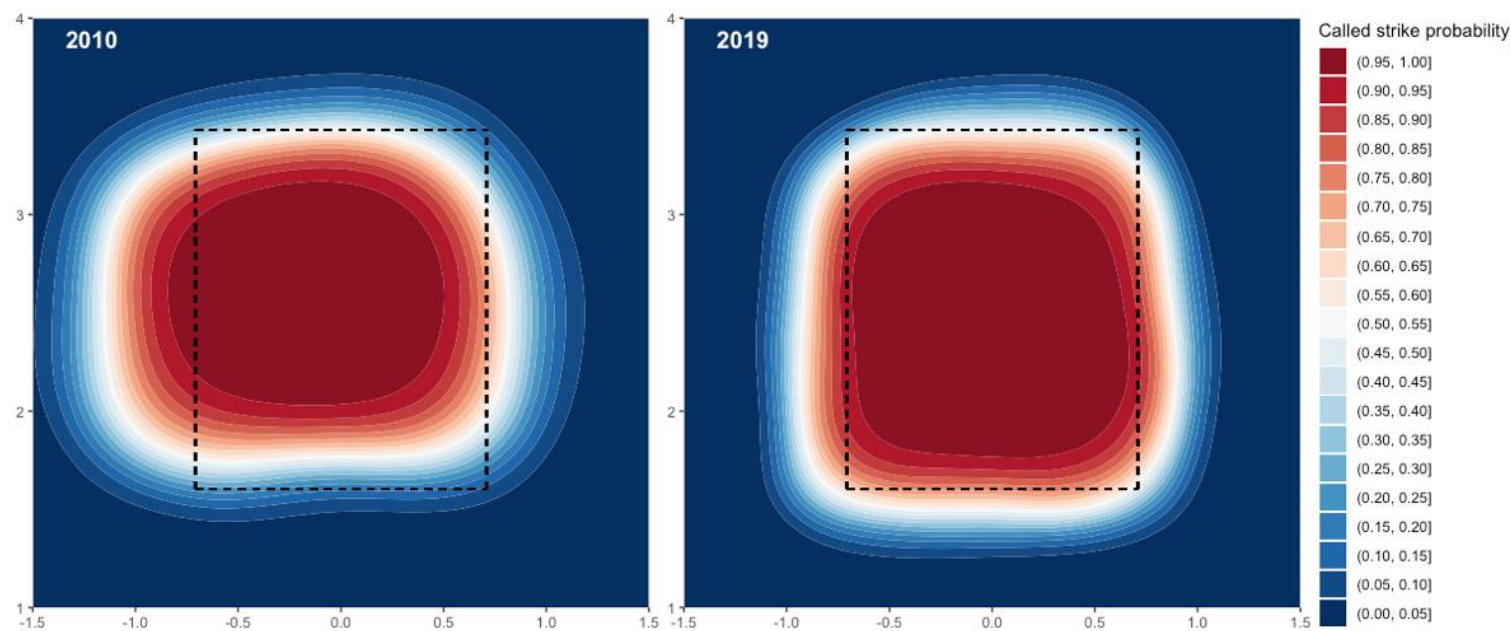


Fig. S10. LHB vs RHP called strike probability for different years. RHB: Right-handed batter; LHP: Left-handed pitcher. Seasons are presented in the upper left corner. The strike zone as defined by the MLB rules is shown in dotted line. The x-axis represents the horizontal position of the ball when it reaches home plate. The reference point 0 is the centre of home plate, as seen from the umpire's position. The x-axis represents the horizontal position of the ball, in inches, when it reaches home plate. Reference point 0 is the centre of home plate, as seen from the umpire's position. The y-axis represents the vertical position of the ball, in inches, when it reaches home plate. Reference point 0 (not visible in the figure) is the ground.

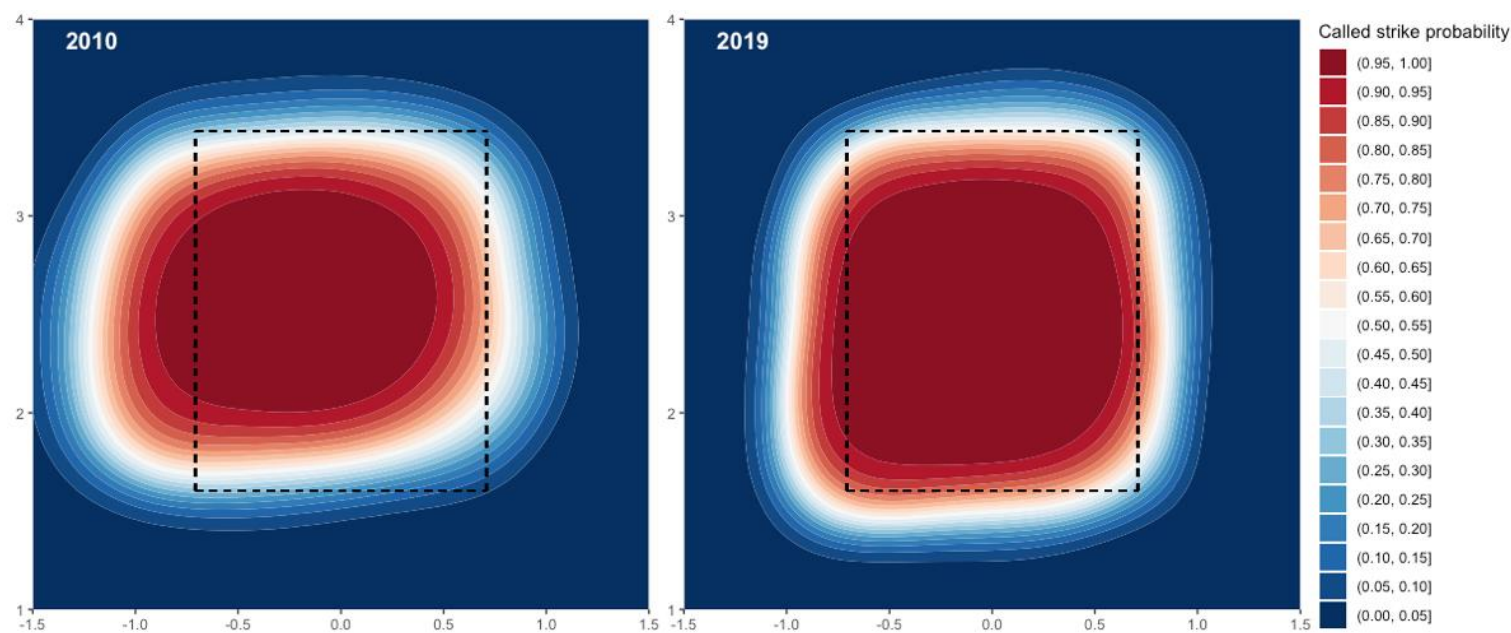


Fig. S11. RHB vs LHP called strike probability for different pitch type. RHB: Right-handed batter; LHP: Left-handed pitcher. FF: Four-seam fastball; CU: Curveball. The strike zone as defined by the MLB rules is shown in dotted line. The x-axis represents the horizontal position of the ball, in inches, when it reaches home plate. Reference point 0 is the centre of home plate, as seen from the umpire's position. The y-axis represents the vertical position of the ball, in inches, when it reaches home plate. Reference point 0 (not visible in the figure) is the ground.

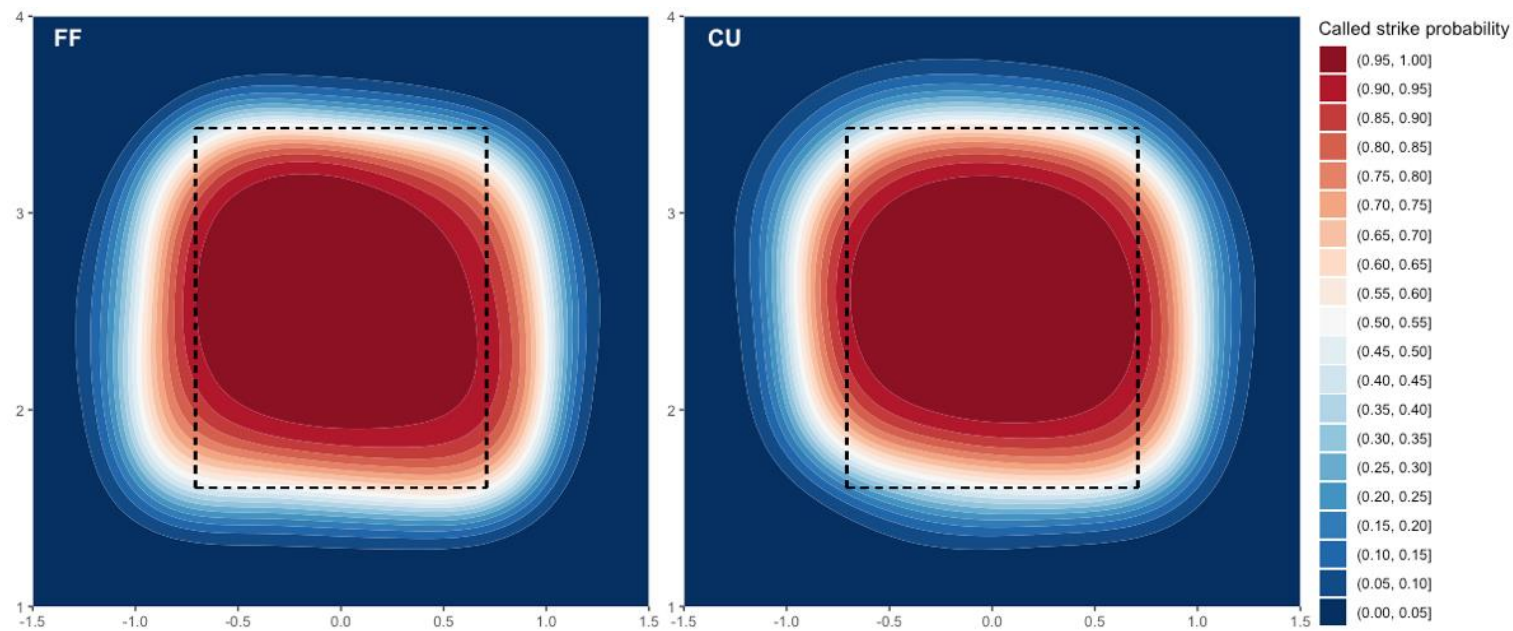


Fig. S12. RHB vs RHP called strike probability for different pitch type. RHB: Right-handed batter; LHP: Left-handed pitcher. FF: Four-seam fastball; CU: Curveball. The strike zone as defined by the MLB rules is shown in dotted line. The x-axis represents the horizontal position of the ball, in inches, when it reaches home plate. Reference point 0 is the centre of home plate, as seen from the umpire's position. The y-axis represents the vertical position of the ball, in inches, when it reaches home plate. Reference point 0 (not visible in the figure) is the ground.

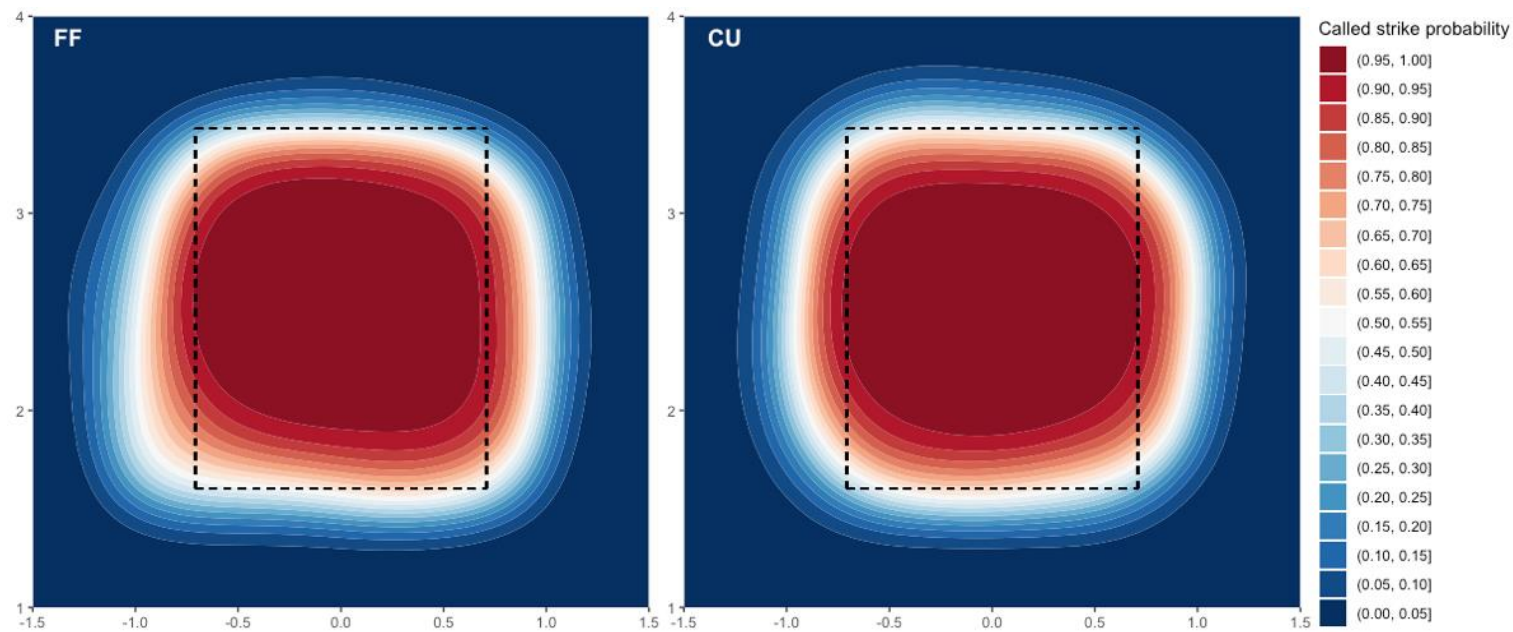


Fig. S13. LHB vs LHP called strike probability for different pitch type. RHB: Right-handed batter; LHP: Left-handed pitcher. FF: Four-seam fastball; CU: Curveball. The strike zone as defined by the MLB rules is shown in dotted line. The x-axis represents the horizontal position of the ball, in inches, when it reaches home plate. Reference point 0 is the centre of home plate, as seen from the umpire's position. The y-axis represents the vertical position of the ball, in inches, when it reaches home plate. Reference point 0 (not visible in the figure) is the ground.

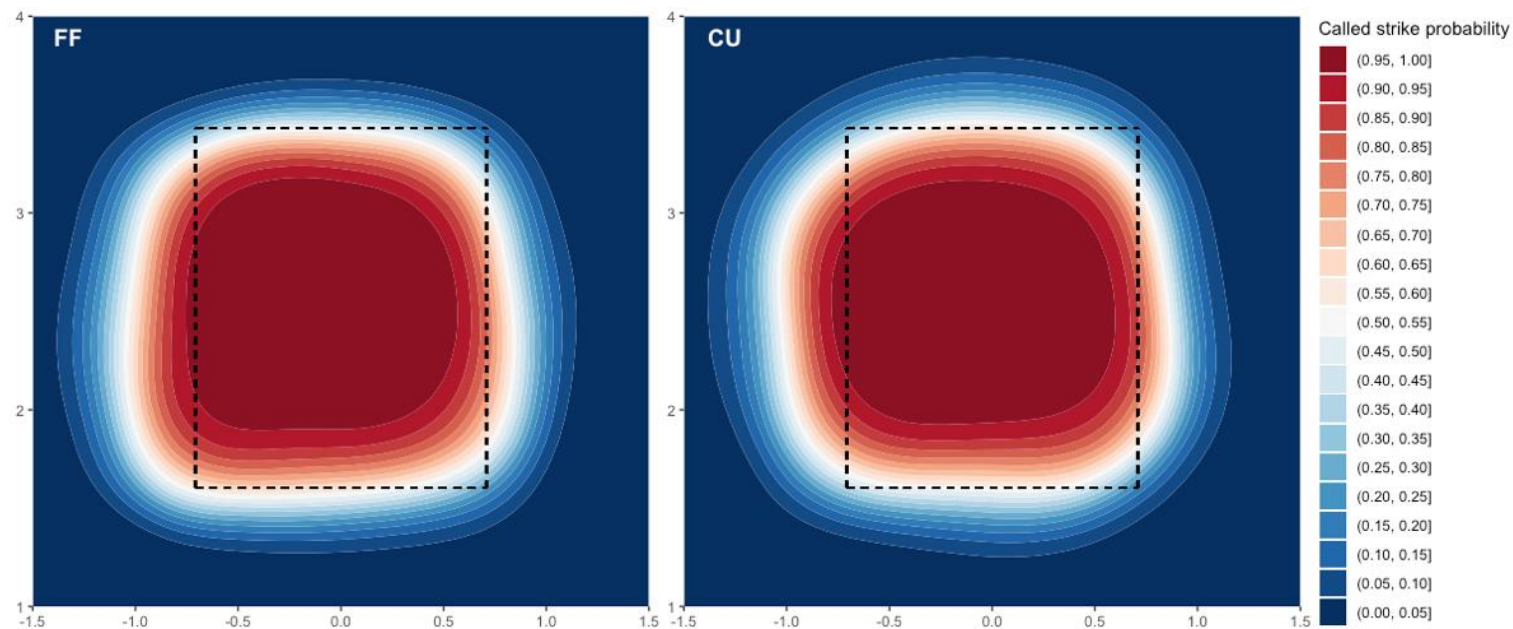


Fig. S14. LHB vs RHP called strike probability for different pitch type. RHB: Right-handed batter; LHP: Left-handed pitcher. FF: Four-seam fastball; CU: Curveball. The strike zone as defined by the MLB rules is shown in dotted line. The x-axis represents the horizontal position of the ball, in inches, when it reaches home plate. Reference point 0 is the centre of home plate, as seen from the umpire's position. The y-axis represents the vertical position of the ball, in inches, when it reaches home plate. Reference point 0 (not visible in the figure) is the ground.

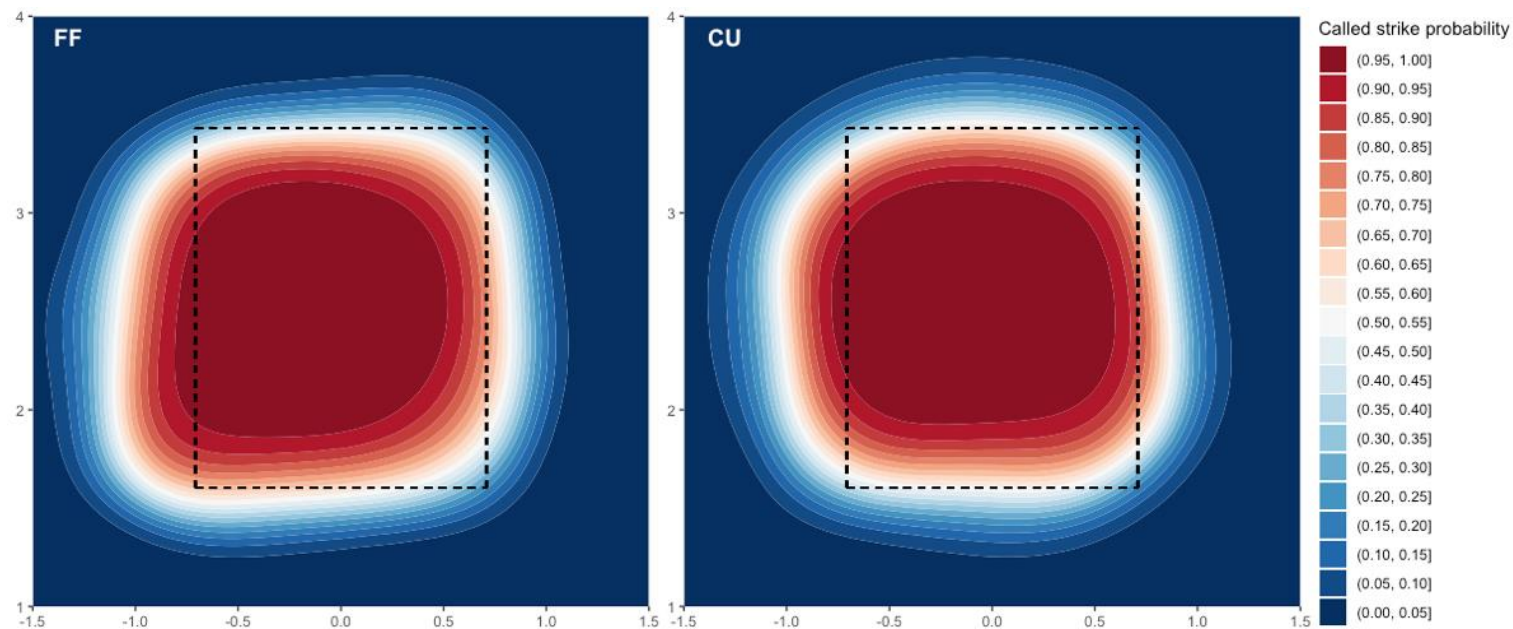


Fig. S15. Mean called strike probability by ejection event and ejected team (model 1). Fig. S15 is generated from the model 1, with the variable baseline called strike probability (BCSprob) set to 0.5 (represented by the dotted line) and the variable Home team set to 1 (home). Error bars represent confidence intervals.

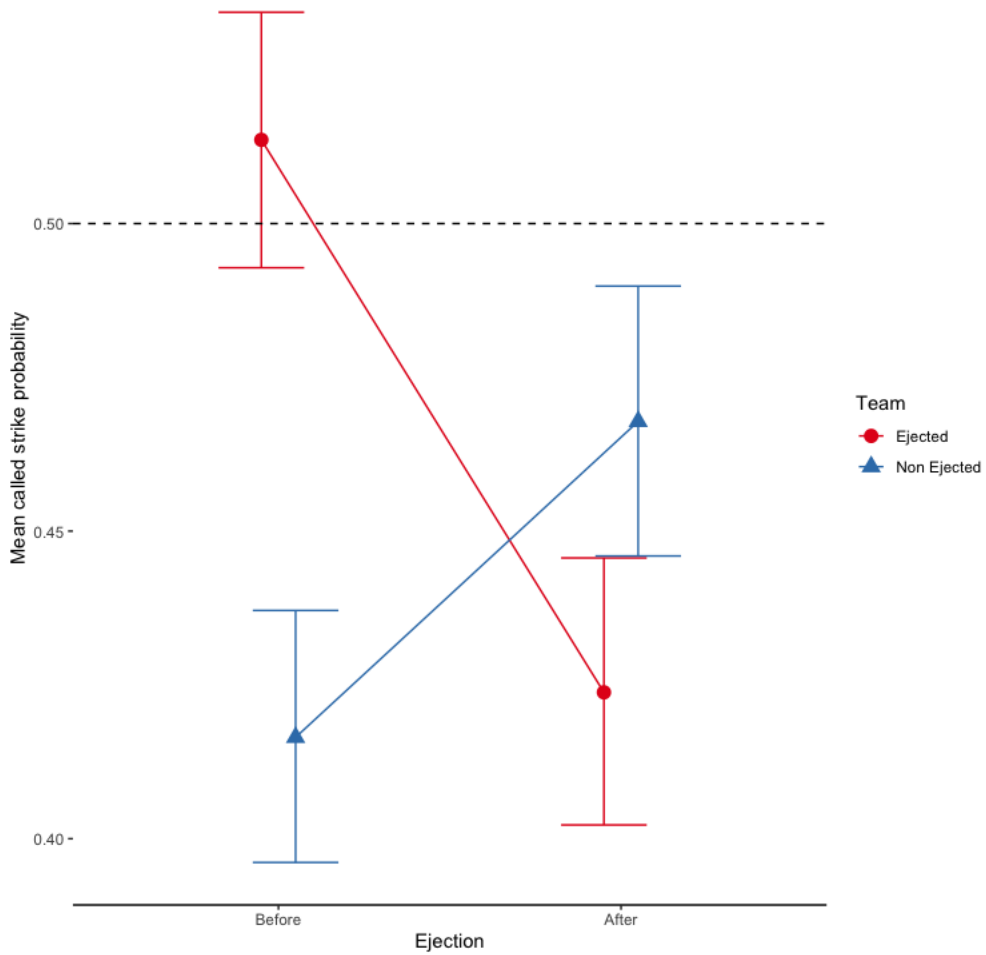


Fig. S16. Mean called strike probability by Ejection reasons (model 2). Fig. S16 is generated from the model 2, with the variable baseline called strike probability (BCSprob) set to 0.5 (represented by the dotted line) and the variable Home team set to 1 (home). Error bars represent confidence intervals.

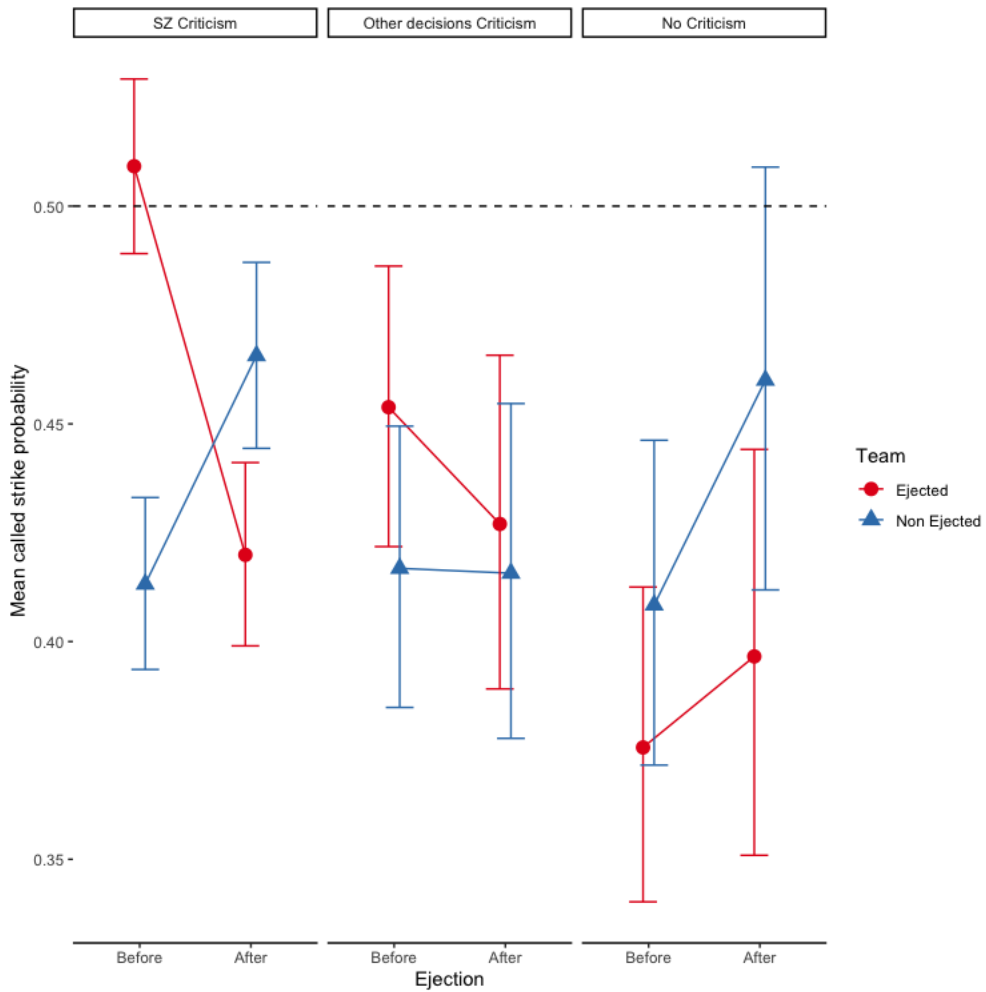


Fig. S17. Mean called strike probability by ejected individual role (model 3). Fig. S17 is generated from the model 3, with the variable baseline called strike probability (BCSprob) set to 0.5 (represented by the dotted line) and the variable Home team set to 1 (home). Error bars represent confidence intervals.

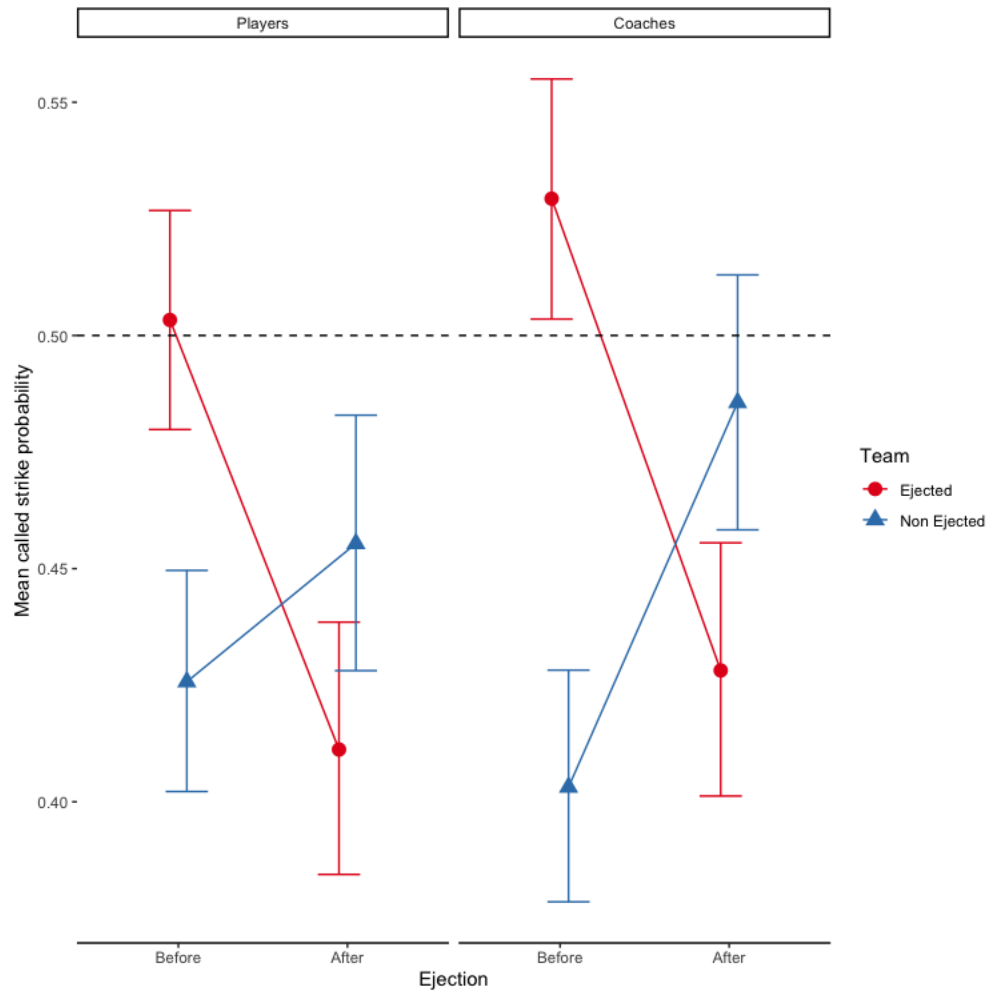


Fig. S18. Mean called strike probability by All-Star status (model 4). Fig. S18 is generated from model 4, with the variable BCSprob set to 0.5 (represented by the dotted line). Triple interaction is not significant. Error bars represent confidence intervals.

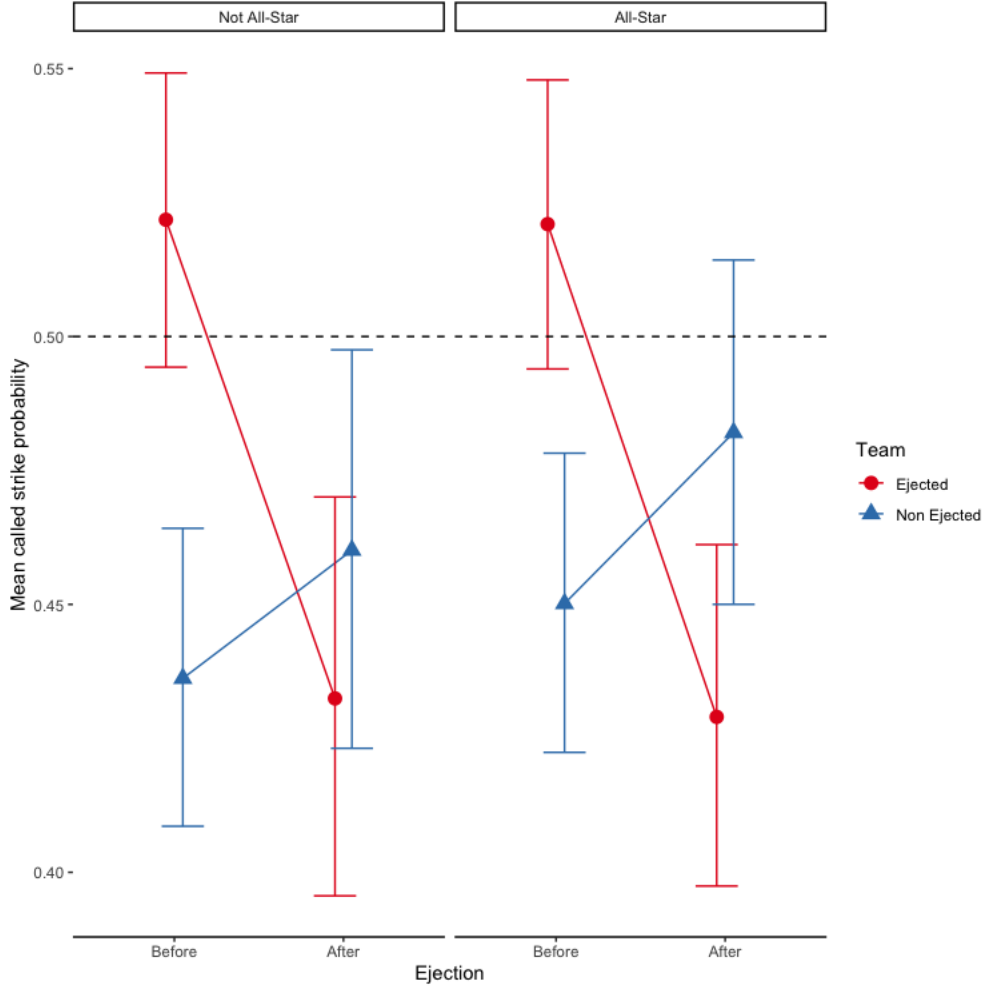


Fig. S19. Mean called strike probability by number of ejected player All-Star appearances (model 5). Fig. S19 is generated from model 5, with the variable BCSprob set to 0.5 (represented by the dotted line). Triple interaction is not significant.

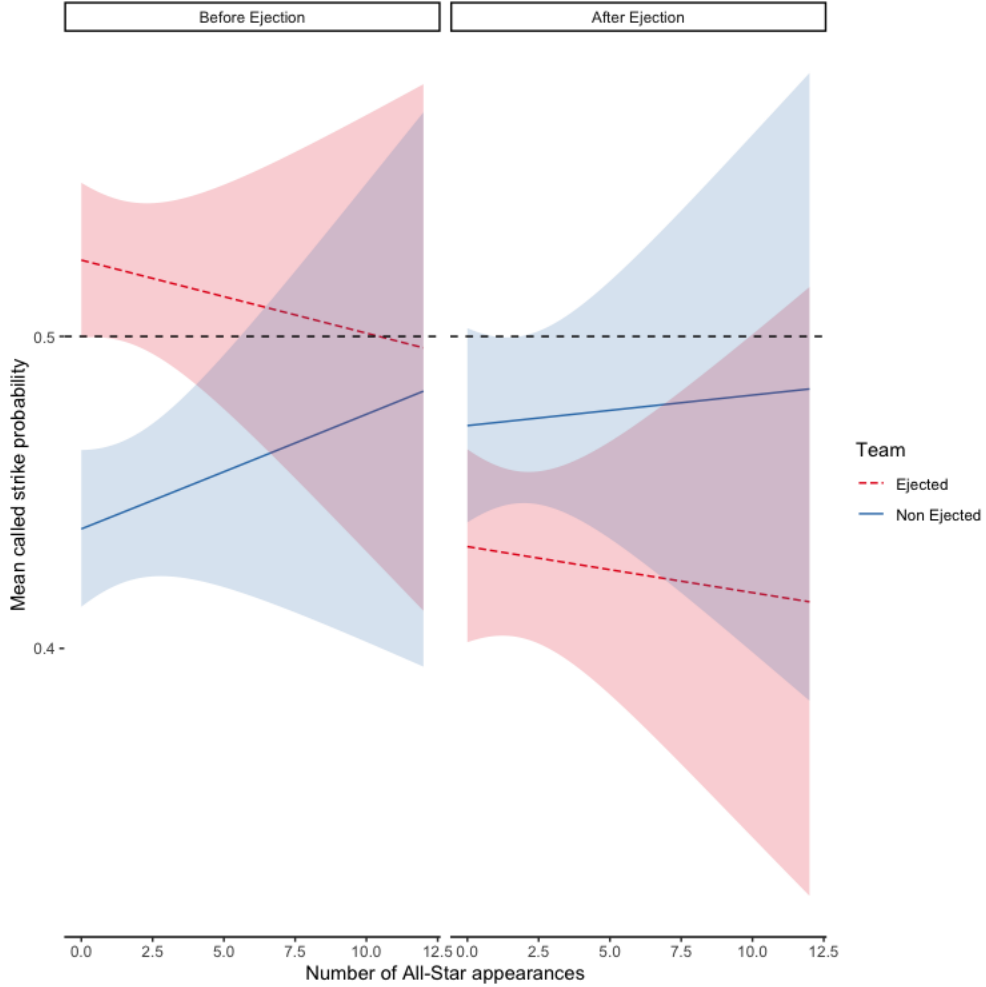


Fig. S20. Mean called strike probability by number of ejected player MLB seasons (model 6). Fig. S20 is generated from model 6, with the variable BCSprob set to 0.5 (represented by the dotted line). Triple interaction is not significant.

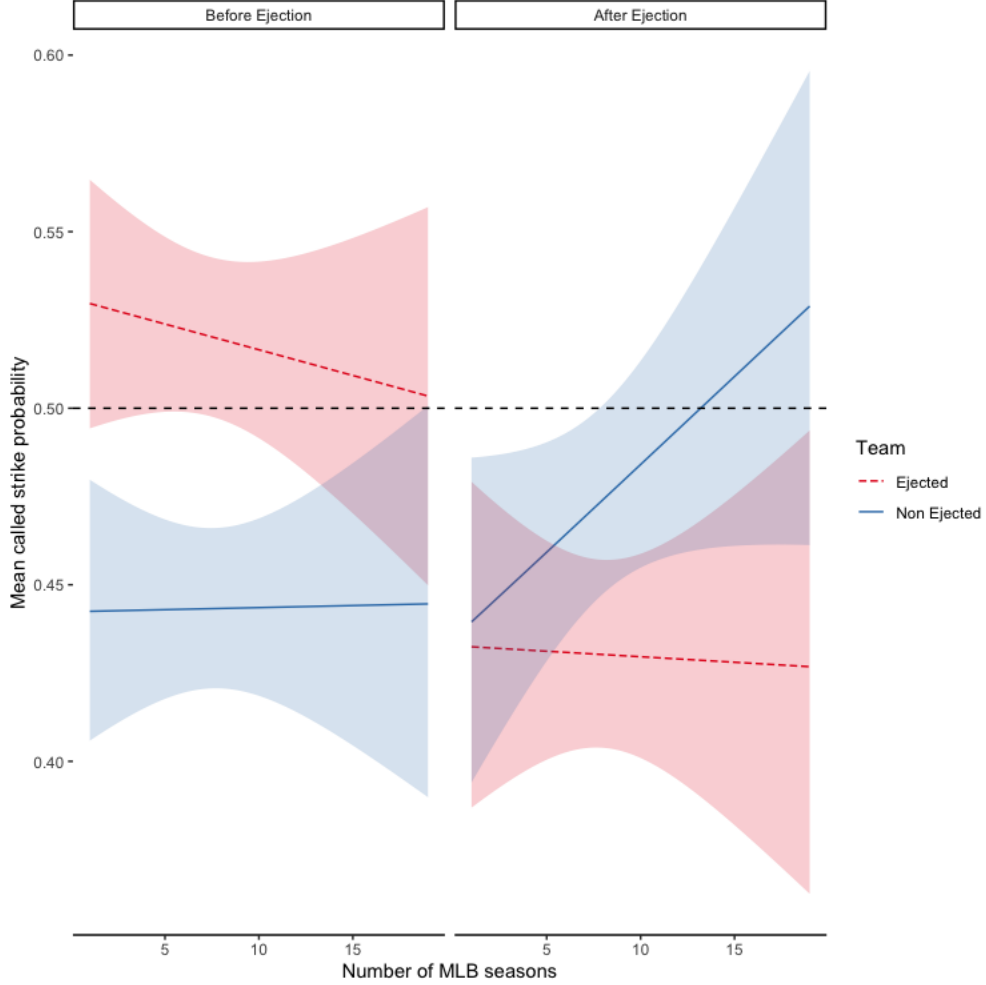


Fig. S21. Mean called strike probability by ejected player WAR (model 7). Fig. S21 is generated from model 7, with the variable BCSprob set to 0.5 (represented by the dotted line). Triple interaction is not significant.

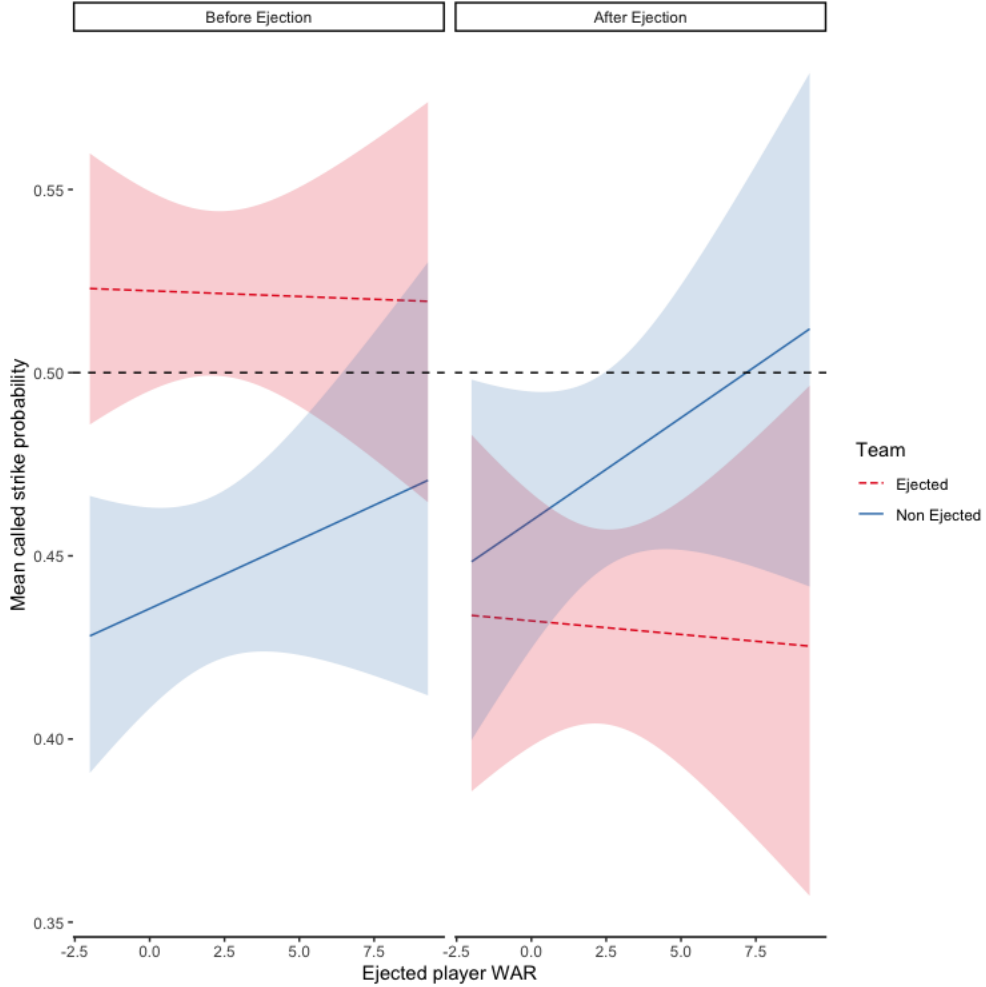


Table S1. Reasons reported for the ejections included in the final sample.

Reasons	Number
<i>Strike zone related excessive criticisms</i>	
Arguing strike one call	1
Balls and strikes	538
Called third strike	403
<i>Non-strike zone related excessive criticisms</i>	
Arguing balk call	4
Arguing bunt attempt	1
Arguing HBP was intentional	1
Arguing hit by pitch	2
Arguing intentional HBP	1
Arguing pitches were intentionally high and inside	1
Arguing warnings after HBP	1
Balk call	16
Balk call (Reversed)	1
Balk non-call	7
Call at HP	23
Call at HP (Reversed)	1
Check swing	46
Foul tip call	2
Foul tip call (overturned)	1
HBP call	7
HBP call (Reversed)	1
HBP non-call	9
Hit by pitch non-call	1
Intentional HBP non-call	2
Mount visit denied	1
Protesting HBP warning	1
Quick pitch	2
Quick pitch call	1
Refusal of appeal on check swing	2
<i>Aggressive behaviour towards the opposing team</i>	
Automatic ejection for HBP after warning	4
Automatic ejection for throwing at batter after warning	2
Automatic ejection on HBP after warning	2
Hit batter after warning	2
Hit by pitch after previous warnings	2
Hit by pitch after warning	2
Intentional HBP	79
Intentional hit by pitch	1
Intentionally hitting batter	1
Throwing at batter	30
Throwing behind batter	1
Total	1,204

Table S2. Regular season games exported to calculate baseline called strike probability (BCSprob).

Season	Start	End	Games with source error	Games in ejection dataset	Games in <i>BCSprob</i> final sample	Pitches in <i>BCSprob</i> final sample
2010	2010-04-04	2010-10-03	0	99	2,326	351,088
2011	2011-03-31	2011-09-28	2	92	2,328	350,135
2012	2012-03-28	2012-10-03	1	90	2,366	349,909
2013	2013-03-31	2013-09-30	0	75	2,349	349,314
2014	2014-03-22	2014-09-28	9	105	2,339	342,539
2015	2015-04-05	2015-10-04	2	112	2,315	332,663
2016	2016-04-03	2016-10-02	2	109	2,318	338,700
2017	2017-04-02	2017-10-01	0	94	2,319	345,158
2018	2018-03-29	2018-10-01	1	95	2,320	346,198
2019	2019-03-28	2019-09-29	2	130	2,268	339,640

Table S3. GAM baseline called strike probability (BCSprob) model results. Baseline: Count = 0-0; Pitch type = Changeup; Year = 2010.

<i>Predictors</i>	Est	Std. Error	z value	p
<i>Parametric coefficients</i>				
(Intercept)	-4.297	0.077	-55.508	<0.001
Count				
0-1	-0.652	0.007	-92.086	<0.001
0-2	-0.974	0.014	-71.469	<0.001
1-0	0.139	0.007	20.782	<0.001
1-1	-0.337	0.008	-43.003	<0.001
1-2	-0.694	0.011	-62.968	<0.001
2-0	0.332	0.010	31.873	<0.001
2-1	-0.133	0.010	-12.819	<0.001
2-2	-0.447	0.011	-38.997	<0.001
3-0	0.518	0.015	34.165	<0.001
3-1	-0.017	0.014	-1.189	0.235
3-2	-0.322	0.015	-22.066	<0.001
Pitch type				
Fastball (FA)	0.705	0.208	3.386	<0.001
Four-seam fastball (FF)	0.118	0.009	13.776	<0.001
Two-seam fastball (FT)	0.151	0.010	15.481	<0.001
Cutter (FC)	0.214	0.122	17.569	<0.001
Splitter (FS)	-0.136	0.023	-5.876	<0.001
Forkball (FO)	-0.091	0.148	-0.613	0.540
Screwball (SC)	0.416	0.248	1.677	0.093
Sinker (SI)	0.143	0.010	14.369	<0.001
Slider (SL)	0.108	0.010	10.927	<0.001
Slow curve (CS)	0.004	0.836	0.005	0.996
Curveball (CU)	0.080	0.010	7.410	<0.001
Knuckle-curve (KC)	0.072	0.017	4.290	<0.001
Knuckleball (KN)	0.092	0.034	2.674	0.007
Eephus (EP)	0.316	0.146	2.169	0.030
Year				
2011	0.081	0.009	8.753	<0.001
2012	0.187	0.009	20.161	<0.001
2013	0.234	0.009	25.077	<0.001
2014	0.343	0.009	36.700	<0.001
2015	0.405	0.009	42.960	<0.001
2016	0.323	0.009	34.386	<0.001
2017	0.281	0.009	29.982	<0.001
2018	0.357	0.009	38.177	<0.001
2019	0.392	0.009	41.653	<0.001
	edf	Ref.df	Chi.sq	p
<i>Smooth terms</i>				

s(pX,pZ):LHB*LHP	48.64	48.98	51701	<0.001
s(pX,pZ):RHB*LHP	47.38	48.52	120116	<0.001
s(pX,pZ):LHB*RHP	48.91	49.00	211601	<0.001
s(pX,pZ):RHB*RHP	48.42	48.94	222735	<0.001
Observations		3,601,050		
UBRE		123.88		
R² adjusted / Deviance explained		0.721 / 0.687		

Table S4. Post-hoc analysis of variables of interest in model 1. Results are averaged over the levels of Home team variable. Confidence interval level adjustment: Tukey method for comparing a family of 4 estimates. Intervals are back-transformed from the log odds ratio scale.

Groups		Odds Ratio	Standard error	95% CI	p-value
Ejected - After	Not Ejected - After	0.837	0.0330	[0.756-0.926]	<0.0001
Ejected - After	Ejected - Before	0.697	0.0245	[0.636-0.762]	<0.0001
Ejected - After	Not Ejected - Before	1.030	0.0376	[0.938-1.132]	0.8709
Not Ejected - After	Ejected - Before	0.833	0.0300	[0.759-0.913]	<0.0001
Not Ejected - After	Not Ejected - Before	1.232	0.0440	[1.124-1.350]	<0.0001
Ejected - Before	Not Ejected - Before	1.479	0.0458	[1.366-1.602]	<0.0001

Table S5. Impact of ejection reason on strike decisions (model 2).

<i>Predictors</i>	Ejection*Team model			Ejection*Team*Reason model		
	<i>Odds Ratios</i>	<i>95% CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>95% CI</i>	<i>p</i>
(Intercept)	0.02	[0.02-0.02]	<0.001	0.02	[0.02-0.02]	<0.001
BCSprob	1223.45	[1148.79-1302.97]	<0.001	1235.15	[1159.60-1315.62]	<0.001
Home Team	1.08	[1.03-1.12]	0.001	1.08	[1.03-1.12]	0.001
Ejection	1.20	[1.13-1.28]	<0.001	1.24	[1.15-1.33]	<0.001
Team	1.38	[1.30-1.45]	<0.001	1.47	[1.39-1.56]	<0.001
Reason [1]				1.02	[0.90-1.15]	0.812
Reason [2]				0.98	[0.85-1.13]	0.791
Ejection*Team	0.62	[0.57-0.67]	<0.001	0.56	[0.51-0.62]	<0.001
Ejection*Reason [1]				0.80	[0.66-0.98]	0.027
Ejection*Reason [2]				1.00	[0.79-1.26]	0.981
Team*Reason [1]				0.79	[0.67-0.93]	0.006
Team*Reason [2]				0.59	[0.48-0.72]	<0.001
Ejection*Team*Reason [1]				1.60	[1.22-2.09]	0.001
Ejection*Team*Reason [2]				1.57	[1.13-2.19]	0.008
<i>Random effects</i>						
σ^2		3.29			3.29	
τ_{00} umpire		0.05			0.05	
τ_{00} catcher		0.07			0.07	
τ_{00} batter		0.03			0.03	
τ_{00} pitcher		0.06			0.06	
ICC		0.06			0.06	
N _{Umpire}		110			110	
N _{Catcher}		221			221	
N _{Batter}		1,744			1,744	
N _{Pitcher}		1,468			1,468	
Observations		153,255			153,255	
Marginal R ² / Conditional R ²		0.700 / 0.718			0.701 / 0.718	

Table S6. Post-hoc analysis of variables of interest in model 2. Results are averaged over the levels of Home team variable. Confidence interval level adjustment: Tukey method for comparing a family of 4 estimates. Intervals are back-transformed from the log odds ratio scale.

Groups		Odds Ratio	Standard error	95% CI	p-value
<i>Strike zone related criticism</i>					
Ejected - After	Not Ejected - After	0.831	0.0326	[0.751-0.919]	<0.0001
Ejected - After	Ejected - Before	0.698	0.0244	[0.638-0.763]	<0.0001
Ejected - After	Not Ejected - Before	1.028	0.0372	[0.937-1.128]	0.8709
Not Ejected - After	Ejected - Before	0.840	0.0300	[0.766-0.921]	<0.0001
Not Ejected - After	Not Ejected - Before	1.238	0.0439	[1.130-1.356]	<0.0001
Ejected - Before	Not Ejected - Before	1.473	0.0452	[1.362-1.594]	<0.0001
<i>Non-strike zone related criticism</i>					
Ejected - After	Not Ejected - After	1.047	0.1071	[0.805-1.362]	0.9690
Ejected - After	Ejected - Before	0.897	0.0803	[0.713-1.129]	0.6170
Ejected - After	Not Ejected - Before	1.043	0.0964	[0.822-1.322]	0.9693
Not Ejected - After	Ejected - Before	0.856	0.0792	[0.675-1.086]	0.3349
Not Ejected - After	Not Ejected - Before	0.995	0.0920	[0.785-1.262]	1.0000
Ejected - Before	Not Ejected - Before	1.163	0.0928	[0.947-1.427]	0.2339
<i>Aggressive behaviour towards opposing team</i>					
Ejected - After	Not Ejected - After	0.771	0.1014	[0.550-1.081]	0.1976
Ejected - After	Ejected - Before	1.092	0.1248	[0.814-1.465]	0.8673
Ejected - After	Not Ejected - Before	0.952	0.1108	[0.706-1.284]	0.9747
Not Ejected - After	Ejected - Before	1.416	0.1658	[1.048-1.913]	0.0158
Not Ejected - After	Not Ejected - Before	1.234	0.1430	[0.916-1.662]	0.2659
Ejected - Before	Not Ejected - Before	0.872	0.0857	[0.677-1.122]	0.5017

Table S7. Impact of the role of the ejected individual on strike decisions (model 3).

<i>Predictors</i>	Ejection*Team model			Ejection*Team*Job model		
	<i>Odds Ratios</i>	<i>95% CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>95% CI</i>	<i>p</i>
(Intercept)	0.02	[0.02-0.02]	<0.001	0.02	[0.02-0.02]	<0.001
BCSprob	1238.67	[1147.32-1337.30]	<0.001	1241.96	[1150.30-1340.92]	<0.001
Home Team	1.06	[1.01-1.12]	0.020	1.06	[1.01-1.12]	0.022
Ejection	1.25	[1.16-1.35]	<0.001	1.45	[1.31-1.61]	<0.001
Team	1.48	[1.39-1.59]	<0.001	1.20	[1.06-1.35]	0.004
Job [C]				1.07	[0.95-1.21]	0.264
Ejection*Team	0.55	[0.49-0.61]	<0.001	0.61	[0.53-0.71]	<0.001
Ejection*Job [C]				1.04	[0.89-1.20]	0.651
Team*Job [C]				1.05	[0.89-1.25]	0.547
Ejection*Team*Job [C]				0.78	[0.63-0.96]	0.021
<i>Random effects</i>						
σ^2		3.29			3.29	
τ_{00} umpire		0.05			0.05	
τ_{00} catcher		0.06			0.07	
τ_{00} batter		0.03			0.03	
τ_{00} pitcher		0.06			0.06	
ICC		0.06			0.06	
N _{Umpire}		107			107	
N _{Catcher}		204			204	
N _{Batter}		1,577			1,577	
N _{Pitcher}		1,287			1,287	
Observations		103,913			103,913	
Marginal R ² / Conditional R ²		0.701 / 0.718			0.701 / 0.719	

Table S8. Post-hoc analysis of variables of interest in model 3. Results are averaged over the levels of Home team variable. Confidence interval level adjustment: Tukey method for comparing a family of 4 estimates. Intervals are back-transformed from the log odds ratio scale.

Groups		Odds Ratio	Standard error	95% CI	p-value
<i>Players</i>					
Ejected - After	Not Ejected - After	0.835	0.0515	[0.713-0.979]	0.0185
Ejected - After	Ejected - Before	0.689	0.0365	[0.601-0.790]	<0.0001
Ejected - After	Not Ejected - Before	0.942	0.0522	[0.817-1.086]	0.7032
Not Ejected - After	Ejected - Before	0.825	0.0441	[0.719-0.946]	0.0018
Not Ejected - After	Not Ejected - Before	1.128	0.0606	[0.982-1.295]	0.1130
Ejected - Before	Not Ejected - Before	1.367	0.0602	[1.221-1.531]	<0.0001
<i>Coaches</i>					
Ejected - After	Not Ejected - After	0.793	0.0479	[0.679-0.926]	0.0007
Ejected - After	Ejected - Before	0.666	0.0371	[0.577-0.768]	<0.0001
Ejected - After	Not Ejected - Before	1.109	0.0632	[0.958-1.283]	0.2690
Not Ejected - After	Ejected - Before	0.839	0.0482	[0.724-0.973]	0.0123
Not Ejected - After	Not Ejected - Before	1.398	0.0780	[1.211-1.613]	<0.0001
Ejected - Before	Not Ejected - Before	1.665	0.0873	[1.455-1.905]	<0.0001

Table S9. Impact of players characteristics on strike decisions with home-field advantage variable included (models 4 to 7).

Reduced model				Model 4			Model 5			Model 6			Model 7		
Ejection*Team model				Ejection*Team*All-Star			Ejection*Team*All-StarAppearances			Ejection*Team*Seasons			Ejection*Team*WAR		
Predictors	Odds Ratios	95% CI	p	Odds Ratios	95% CI	p	Odds Ratios	95% CI	p	Odds Ratios	95% CI	p	Odds Ratios	95% CI	p
(Intercept)	0.02	[0.02-0.02]	<0.001	0.02	[0.02-0.02]	<0.001	0.02	[0.02-0.02]	<0.001	0.02	[0.02-0.02]	<0.001	0.02	[0.02-0.02]	<0.001
BCSprob	1257.14	[1133.75-1393.97]	<0.001	1257.41	[1133.97-1394.28]	<0.001	1256.93	[1133.55-1393.75]	<0.001	1258.68	[1135.09-1395.72]	<0.001	1257.17	[1133.75-1394.02]	<0.001
Home Team	1.07	[0.99-1.15]	0.071	1.07	[0.99-1.15]	0.070	1.07	[1.00-1.15]	0.062	1.07	[1.00-1.15]	0.058	1.07	[1.00-1.15]	0.067
Ejection	1.13	[1.02-1.26]	0.021	1.11	[0.95-1.30]	0.204	1.13	[1.02-1.26]	0.025	1.13	[1.01-1.25]	0.025	1.12	[1.01-1.25]	0.030
Team	1.37	[1.26-1.50]	<0.001	1.42	[1.25-1.61]	<0.001	1.37	[1.26-1.50]	<0.001	1.37	[1.26-1.50]	<0.001	1.37	[1.26-1.50]	<0.001
Characteristic				1.06	[0.93-1.21]	0.380	1.02	[0.98-1.05]	0.344	1.00	[0.98-1.02]	0.913	1.02	[0.98-1.05]	0.337
Ejection*Team	0.61	[0.53-0.71]	<0.001	0.63	[0.50-0.78]	<0.001	0.61	[0.53-0.71]	<0.001	0.61	[0.53-0.71]	<0.001	0.62	[0.53-0.71]	<0.001
Ejection*Characteristic				1.03	[0.83-1.27]	0.804	0.99	[0.94-1.04]	0.644	1.02	[0.99-1.05]	0.174	1.01	[0.96-1.06]	0.721
Team*Characteristic				0.94	[0.79-1.12]	0.462	0.97	[0.93-1.02]	0.237	0.99	[0.97-1.02]	0.543	0.98	[0.94-1.02]	0.436
Ejection*Team*Characteristic				0.97	[0.72-1.30]	0.835	1.02	[0.95-1.08]	0.656	0.99	[0.95-1.02]	0.440	0.99	[0.93-1.06]	0.745
Random effects															
σ^2		3.29			3.29			3.29			3.29			3.29	
τ_{00} umpire		0.05			0.05			0.05			0.05			0.05	
τ_{00} catcher		0.05			0.05			0.05			0.05			0.05	
τ_{00} batter		0.04			0.04			0.04			0.04			0.04	
τ_{00} pitcher		0.04			0.04			0.04			0.04			0.04	
ICC		0.05			0.05			0.05			0.05			0.05	
N _{umpire}		97			97			97			97			97	
N _{catcher}		175			175			175			175			175	
N _{batter}		1357			1357			1357			1357			1,357	
N _{pitcher}		1083			1083			1083			1083			1,083	
Observations		57,223			57,223			57,223			57,223			57,223	
Marginal R ² / Conditional R ²		0.702 / 0.718			0.702 / 0.718			0.702 / 0.718			0.702 / 0.718			0.702 / 0.718	
Model comparison															
Chi-sq		3.2927			3.3080			3.5150			3.6157			3.3802	
p-value		0.0696			0.0689			0.0608			0.0572			0.0660	

Table S10. Impact of players characteristics on strike decisions (models 4 to 7).

	Reduced model			Model 4			Model 5			Model 6			Model 7		
	Ejection*Team model			Ejection*Team*All-Star			Ejection*Team*All-StarAppearances			Ejection*Team*Seasons			Ejection*Team*WAR		
<i>Predictors</i>	<i>Odds Ratios</i>	<i>95% CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>95% CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>95% CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>95% CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>95% CI</i>	<i>p</i>
(Intercept)	0.02	[0.02-0.02]	<0.001	0.02	[0.02-0.02]	<0.001	0.02	[0.02-0.02]	<0.001	0.02	[0.02-0.02]	<0.001	0.02	[0.02-0.02]	<0.001
BCSprob	1255.95	[1132.70-1392.62]	<0.001	1256.28	[1132.98-1392.99]	<0.001	1255.71	[1132.48-1392.34]	<0.001	1257.37	[1133.95-1394.23]	<0.001	1255.90	[1132.64-1392.57]	<0.001
Ejection	1.13	[1.01-1.25]	0.026	1.10	[0.94-1.29]	0.234	1.12	[1.01-1.25]	0.030	1.12	[1.01-1.25]	0.031	1.12	[1.01-1.25]	0.036
Team	1.37	[1.25-1.49]	<0.001	1.41	[1.24-1.60]	<0.001	1.36	[1.25-1.49]	<0.001	1.37	[1.25-1.49]	<0.001	1.37	[1.25-1.49]	<0.001
Characteristic				1.06	[0.93-1.21]	0.400	1.01	[0.98-1.05]	0.374	1.00	[0.98-1.02]	0.976	1.02	[0.98-1.05]	0.329
Ejection*Team	0.61	[0.53-0.71]	<0.001	0.63	[0.51-0.79]	<0.001	0.62	[0.53-0.72]	<0.001	0.62	[0.53-0.72]	<0.001	0.62	[0.53-0.72]	<0.001
Ejection*Characteristic				1.03	[0.84-1.28]	0.770	0.99	[0.94-1.04]	0.657	1.02	[0.99-1.05]	0.169	1.01	[0.96-1.06]	0.759
Team*Characteristic				0.94	[0.79-1.12]	0.504	0.98	[0.93-1.02]	0.273	0.99	[0.97-1.02]	0.632	0.98	[0.94-1.02]	0.427
Ejection*Team*Characteristic				0.96	[0.71-1.29]	0.779	1.01	[0.95-1.08]	0.674	0.98	[0.95-1.02]	0.428	0.99	[0.93-1.06]	0.783
<i>Random effects</i>															
σ^2		3.29			3.29			3.29			3.29			3.29	
τ_{00} umpire		0.05			0.05			0.05			0.05			0.05	
τ_{00} catcher		0.05			0.05			0.05			0.05			0.05	
τ_{00} batter		0.04			0.04			0.04			0.04			0.04	
τ_{00} pitcher		0.04			0.04			0.04			0.04			0.04	
ICC		0.05			0.05			0.05			0.05			0.05	
N _{umpire}		97			97			97			97			97	
N _{catcher}		175			175			175			175			175	
N _{batter}		1,357			1,357			1,357			1,357			1,357	
N _{pitcher}		1,083			1,083			1,083			1,083			1,083	
Observations		57,223			57,223			57,223			57,223			57,223	
Marginal R ² /		0.702 /			0.702 /			0.702 /			0.702 /			0.702 /	
Conditional R ²		0.718			0.718			0.718			0.718			0.718	

Table S11. Overview of model comparisons. FE1: Fixed effects for baseline strike zone probability (BCSprob), and home team. FE2: Fixed effects for baseline strike zone probability (BCSprob). RE: Random effects for umpires, catchers, batters, and pitchers.

Model	Reduced model	Term added	AIC	BIC	LogLikelihood	Residual Deviance	χ^2	χ^2_{DF}	Model comparison <i>p</i> -value	Triple interaction <i>p</i> -value
Model 2	Ejection*Team + FE ₁ + RE		70232	70331	-35106	70212				
		Ejection*Team* Reason	70177	70356	-35070	70141	70.760	8	<0.001	<0.001
Model 3	Ejection*Team + FE ₁ + RE		47706	47802	-23843	47686				
		Ejection*Team* Role	47701	47834	-23836	47673	13.559	4	0.009	0.039
Model 4	Ejection*Team + FE ₂ + RE		26191	26272	-13086	26173				
		Ejection*Team* All-Star	26197	26314	-13086	26171	1.633	4	0.803	0.779
Model 5	Ejection*Team + FE ₂ + RE		26191	26272	-13086	26173				
		Ejection*Team* N _{All-Star}	26198	26314	-13086	26172	1.354	4	0.852	0.674
Model 6	Ejection*Team + FE ₂ + RE		26191	26272	-13086	26173				
		Ejection*Team* N _{Seasons}	26195	26312	-13085	26169	3.532	4	0.473	0.428
Model 7	Ejection*Team + FE ₂ + RE		26191	26272	-13086	26173				
		Ejection*Team* WAR	26197	26313	-13085	26171	2.230	4	0.694	0.783