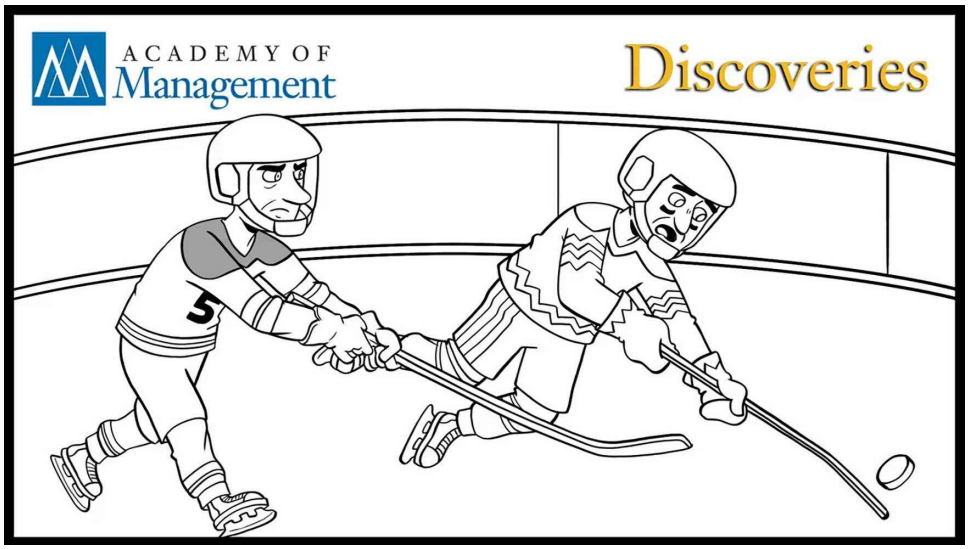
© *Academy of Management Discoveries*
2025, Vol. 11, No. 2, 180–203.
https://doi.org/10.5465/amd.2022.0280

A (BOUNDED) PREFERENCE FOR RULE BREAKERS

S. WILEY WAKEMAN
Stockholm School of Economics

PHILIP YANG
Paderborn University

CELIA MOORE
Imperial College London



This paper explores whether, under what conditions, and why supervisors treat rule breakers favorably. Using field data from six seasons in the National Hockey League, we find that coaches select rule breakers for play more often than rule-abiding players, even if rule breaking may ultimately hurt a team's chance of winning. However, this preference is bounded in several ways. It reverses when a player's rule breaking becomes extreme, holds for minor (but not major) rule breaking, is absent for both players and teams in the lowest quintile of rule breaking, is observed in regular season (but not playoff) games, and is amplified when teams are on losing streaks but disappears after repeated wins. An experiment replicates this bounded preference for rule breakers and identifies that those in a position to reward rule breakers do so in part because they perceive them as more committed to team success. However, at extreme levels of rule breaking, concerns about the liability rule breakers represent eclipse these positive perceptions, reversing the preferential treatment they enjoyed when their rule breaking was more moderate. Together, these findings illuminate one unexpected reason why rule breaking is so rampant and unethical behavior remains so pernicious in organizations.

The authors would like to thank Mark Conley, Katleen De Stobbeleir, Thorsten Grohsjean, Martin Kilduff, Holmer Kok, Julia Lee, Maddy Ong, and Colleen Stuart for their comments on a draft version of this manuscript; as well as Simona Botti, David Faro, Randall Peterson, and Shaul Shalvi for their thoughts and suggestions on earlier drafts of this manuscript as part of the first author's dissertation committee. Special thanks should be given to Jasper Cushing, Juliane Schittek, and Michael Davies for their help with data coding regarding offset penalties. Participants at the International Association of Conflict Management annual meeting, the Academy of Management annual meeting, the Interdisciplinary Network for Groups Research annual conference, members of the Organizational Behavior Lab at Bocconi University and Centre for Responsible Leadership at Imperial Business School, the House of Innovation Brownbag at Stockholm School of Economics, UCL Organisations and Innovation group, as well as colleagues at London Business School and Stockholm School of Economics also provided thoughtful suggestions. Finally, the authors wish to thank the Leadership Institute at London Business School, Stockholm School of Economics, University of Paderborn, and Centre for Responsible Leadership at Imperial College for resource support. Correspondence should be addressed to Wiley Wakeman, Stockholm School of Economics, wiley.wakeman@hhs.se.

Accepted by Kevin Rockmann

Rules are enacted to constrain people from behaving in ways that legislators, regulators, industry bodies, or organizations have deemed to be unfair, harmful, or otherwise inappropriate (Posner, 1997). By directing and constraining employees' efforts, they are a necessary part of organizational life (March & Simon, 1958; Okhuysen & Bechky, 2009; Scott & Davis, 2007). Yet, rules can only forbid the behavior they endeavor to regulate. In practice, rule breaking remains tempting and is often perceived to be beneficial (Hollinger, 1986; Mars, 1982), which is part of the reason why it remains rife in organizations (MacLean, 2001; Martin, Lopez, Roscigno & Hodson, 2013). What is missing in this disconnect between proscriptive and descriptive outcomes of rule breaking is a better understanding of the routes through which rule breaking might evolve from being prohibited behavior to a normalized aspect of organization culture (Ashforth & Anand, 2003; Brief, Buttram & Dukerich, 2001).

Scholarship on organizational rule breaking has tended to focus on antecedents rather than consequences (MacLean, 2001; Vadera, Pratt & Mishra, 2013). When consequences of rule breaking have been explored, negative outcomes have dominated (Butterfield, Treviño & Ball, 1996; Dahling, Chau, Mayer & Gregory, 2012; Podsakoff, Bommer, Podsakoff & MacKenzie, 2006; Tyler & Boeckmann, 1997). However, for rule violations to be as common as they are, there must be alternative processes beyond decisions to punish (or not punish) rule breakers. Rule breaking must be rewarded and reinforced in ways that explain its normalization as part of an organization's or industry's culture (Ashforth & Anand, 2003). To date, existing scholarship has broadly overlooked the possibility that rule breaking persists because those in positions of power—individuals whose decisions to select, reward, and advance employees shape organizational culture (Brief et al., 2001)—routinely reward it.

This paper explores whether, under what conditions, and why supervisors might treat rule breakers preferentially. Our curiosity about this prospect was motivated by stories in the media about individuals who, before being indicted for serious misconduct, had regularly received preferential treatment from their supervisors. For example, Kweku Adoboli, the "rogue" trader from Swiss Bank UBS who ultimately cost the bank $2.3 billion (FINMA, 2012), had broken rules successfully for years. For much of this time, others at the bank, including those with supervisory authority over him, were well aware of Adoboli's actions. Yet, they had done nothing to stop him (Financial Conduct Authority, 2014). Indeed, it became clear during his trial that, instead of punishing his rule breaking, his supervisors had praised and promoted him regularly (Croft, 2012), tripling his compensation in the year before his arrest (Walker, 2012). Though this is only one example of supervisors treating rule breakers favorably, it remains an open question as to whether this example is unique or perhaps indicative of a broader preference for rule breakers.

Using an archival dataset of players' rule breaking in the National Hockey League (NHL) across six seasons, we explore whether players who violate codified and widely agreed upon league rules receive favorable treatment from coaches. This context provides several advantages for studying the consequences of rule breaking. First, rule breaking is observable and judged by independent third parties. This active monitoring by experts who use a clear classification system to determine which rules have been broken and what the penalties are for having done so is less subject to local understandings or political processes that might bias the reporting of such behaviors (Roulet, 2019). Second, favorable treatment (playing time) is measured accurately and precisely, providing a behavioral (vs. attitudinal) measure of preferential treatment. We pair these archival data with an experiment that addresses endogeneity concerns by manipulating rule breaking, which allows us to explore what might motivate supervisors' preferential treatment of rule breakers.

RULE BREAKING IN ORGANIZATIONS

Rule breaking is typically considered as one form of deviant behavior—a broader category constituted by actions that violate or contravene norms, expectations, or mandates of one's managers, organization, or other form of authority (Dahling & Gutworth, 2017; Hollinger, 1986; Jetten & Hornsey, 2014; Mishra, Ghosh & Sharma, 2021; Vadera et al., 2013). Existing scholarship on rule breaking has focused largely on *why* individuals break rules. Some rule breaking is selfishly motivated, to increase personal gain (Balasubramanian, Bennett & Pierce, 2017; Mars, 1982), to protect oneself from perceived threats (Belmi, Barragan, Neale & Cohen, 2015; Wakeman, Moore & Gino, 2019), or to "fit" into an organizational culture that appears to require it (Apel & Paternoster, 2009; Fleming, 2020; Roulet, 2019). Other rule breaking is prosocially motivated (Malik & Mishra, 2023; Morrison, 2006). Individuals break rules to help customers (Ambrose, Taylor & Hess, 2015; Gazzoli, Chaker, Zablah & Brown, 2022; Gong, Sun & Kang, 2022), peers (Ghosh & Shum, 2019), or their organization (Dahling et al., 2012; Morrison, 2006), particularly when they perceive the rules as unfair (Gong et al., 2022; Veiga, Golden & Dechant, 2004), or as an obstacle to delivering on their objectives

(MacLean, 2001; Martin et al., 2013). Though this extant research provides great insight into why actors break rules, it is silent about the consequences they face *after* breaking them.

Work on the consequences of rule breaking concentrates almost exclusively on negative outcomes (Malik & Mishra, 2023; Mo, Lupoli, Newman & Umphress, 2023). The most dominant perspective focuses on how rule breakers are punished (Butterfield et al., 1996; Mooijman, van Dijk, Ellemers & van Dijk, 2015; Tyler & Boeckmann, 1997; Zhang, Ma, Chen, Cheng, Tian & Tian, 2019). Punishing rule breakers reconfirms a group's existing values (Okimoto & Wenzel, 2009) and discourages subsequent misbehavior (Apel, 2013; Carlsmith, Darley & Robinson, 2002). A second perspective focuses on how, in certain contexts, groups tolerate those rule breakers (Coser, 1962; Dentler & Erikson, 1959; Jetten & Hornsey, 2014; Jetten, Hornsey, Spears, Haslam & Cowell, 2010), particularly if doing so is perceived to contribute to group goals (Fielding, Hogg & Annandale, 2006; Morrison, 2006).

Remarkably few studies have investigated the potential positive consequences of rule breaking. However, those that have done so have reported inconsistent and paradoxical findings. For example, rule breaking has been associated with feeling pride at the same time as guilt (Tang, Yam & Koopman, 2020), and morally deficient at the same time as psychologically entitled (Liao, Yam & Johnson, 2019). Rule breaking has been associated with higher levels of job satisfaction (Kim & Zhan, 2023), but also higher levels of work-life conflict (Liu, Lu, Zhang & Cai, 2023), emotional exhaustion (Kim & Zhan, 2023), and anxiety (Liu et al., 2021). The range and inconsistency of these intrapersonal reactions to rule breaking hint at the complexity of its effects for rule breakers themselves.

When we look at how rule breakers are evaluated and treated by others—particularly those in positions of authority over the rule breaker—we find inconclusive results. As mentioned above, the dominant perspective on the consequences of rule breaking focuses on punishment, and in particular how those in positions of power punish rule violations (Mooijman et al., 2015; Tyler, 2006; Wiltermuth & Flynn, 2013). This body of work takes the view that those with authority are motivated to regulate those under their influence. From this perspective, when subordinates break rules, the relevant question is not whether to punish or reward the rule breaker, but whether the punishment for the rule breaker should be severe or lenient (Zhang et al., 2019; Zipay, Mitchell, Baer, Sessions & Bies, 2021).

The conversation shifts a little when it turns to rule breaking that is perceived to be prosocially motivated. Morrison's (2006: 24) original theorizing about prosocial rule breaking was explicitly undecided about whether supervisors would respond positively or negatively to it, and the empirical evidence that has emerged since has not helped clarify that indecision. One study finds that employees who engage in unethical pro-organizational behavior are evaluated positively by their supervisors (Zhan & Liu, 2022), but others find that it is associated with more negative supervisor evaluations, particularly of employee task performance (Dahling et al., 2012; Rhee, 2019). Still others find that supervisor responses to rule breaking depend on the type of rule breaking (Ghosh & Shum, 2019), or whether they are able to decouple its "bad" aspects from its contributions to performance (Fehr, Welsh, Yam, Baer, Wei & Vaulont, 2019). These radically divergent findings indicate the need to explore how supervisors treat rule breakers in more depth.

Complicating matters is the fact that, to date, studies of supervisors' responses to rule breaking have relied on self- or other-reported scale-based measures that tap a mix of behavioral tendencies (e.g., Dahling et al., 2012), motivations for (e.g., Borry, 2017), and intentions to engage in rule breaking (e.g., Kim & Zhan, 2023), rather than employees' *actual* rule breaking. The assumption underlying these measures is that they represent a solid enough proxy for actual behavior, but it is difficult to ignore the bias that self-presentational concerns likely introduce in this domain (Hewitt, Poole & Regoli, 1984; Webley, Cole & Eidjar, 2001). To address this bias, scholars have turned to reports from organizational peers (e.g., Irshad, Bartels, Majeed & Bashir, 2022; Tang, Yam, Koopman & Ilies, 2022), though the fact that these peers often benefit from the target's rule breaking activates a different self-serving bias that may influence the accuracy of these measures as well (Bocian & Wojciszke, 2014).

Despite the limitations of self-reported measures, the use of real, behavioral measures of rule breaking in research is disappointingly rare. To avoid the penalties that rule breaking elicits, the behavior itself is usually hidden, and companies rarely report instances of internal rule breaking willingly. When organizations are caught violating external regulatory bodies' rules, they often seek ways to avoid disclosing anything about their infractions. One study reported that less than 5% of U.S. Securities and Exchange Commission (SEC) settlements involve admitting any information about the violation that elicited the enforcement action (Winship & Robbennolt, 2018). Yet, for scholars to truly understand rule breaking in organizations, it remains important to measure its various forms accurately and objectively (Hewitt et al., 1984), as well as capture data rich enough to explore its contextual effects.

METHODS

It is difficult to study rule breaking empirically. Even when rule breaking has been discovered and recorded accurately, organizations often work diligently to ensure it remains buried from public view. For example, the research team approached both the Commodity Futures Trading Commission (CFTC) in the United States and the European Commission (EC) for text data gathered during the prosecution of several multinational banks for rigging interbank lending rates London Inter-Bank Offered Rate (LIBOR). Both the CFTC[1] and EC[2] responded that such data were effectively unavailable, as the organizations had successfully lobbied that public disclosure of client information, even if redacted, would violate customer privacy and potentially inhibit future whistleblowing.

To find a context where we could study how supervisors treat their team members as a function of their rule-breaking behavior, we turned to professional sports. Professional sports offer a useful context to study rule breaking because most sports have codified rules that are policed actively. In addition, in many sports, rule breaking is easily observed and well recorded, making data on player-level and team-level rule breaking extensive and accessible.

Research Context

We considered various sports, but ultimately chose the NHL for three primary reasons. First, rule breaking in the NHL—engaging in behavior that elicits formal penalties from the referees overseeing the game—is discrete, observable, and easily measured. As third-party professionals who watch every minute of play, referees observe, independently judge, and penalize players when they violate the standardized and codified rules of the professional association that regulates the industry (NHL, 2023). In other sports, rule breaking is observed less easily and can go undetected until long after it has occurred. Efforts to determine how coaches treat members of their teams as a function of their rule breaking would be complicated in contexts where players can accrue multiple rule violations before being removed from play (e.g., football or soccer), or in sports with long delays between rule breaking, its discovery, and the penalties imposed for it (e.g., cycling).

Second, we wanted a context where rule breaking is both common and includes enough variance across players, games, and teams to detect its effects and determine potential moderators. Though rule breaking is common in many sports, it is particularly so in ice hockey. On average, NHL teams incur between two and five penalties per game, which vary both across players and within players throughout a season, as well as within and across teams. This variance provided the empirical traction necessary to test how coaches treat players as a function of their rule breaking, while also enabling identification of contextual boundaries.

Finally, we wanted a context in which penalties for rule breaking have meaningful negative implications for the team, as they do in organizational settings. Though rule breaking is normative in the sport (as it is in many organizations; Martin et al., 2013; Veiga et al., 2004), it nevertheless imposes an immediate cost on NHL teams. In ice hockey, referees remove penalized players from the game for fixed periods of time, meaning that the opposing team has the advantage of an extra player on the ice for that period. Thus, when a player breaks a rule, it represents an immediate liability for the team.

Data

We collected our data from hockeyreference.com, a comprehensive database on NHL players, coaches, teams, and games. Our sample consists of complete game-level data on all NHL players for six seasons (2012–2013 to 2017–2018). We merge this player- and game-level dataset with a dataset of coaches to create player–coach dyads. The coaching file provides information on the exact dates that coaches worked for any team. Finally, we merge these files with data on team performance for each game. Though all data are from the same site, the different datasets are not equally complete, thus we restrict our sample to observation years with information on all three datasets. The starting date also comes after the NHL began to police rule breaking (particularly that with the potential to cause injury) more rigorously, a change that occurred following a series of head injuries to important players in the 2009–2010 season (Wyshynski, 2021).

Measures

Dependent variable: Preferential treatment. We operationalize preferential treatment in terms of the total minutes of ice time a player receives in a focal game. Decisions about who among a pool of players gets time on ice—a scarce and valuable resource—are among the most important coaches make (Staw & Hoang, 1995; Zhang, 2017). Unlike sports where players can play for the entire game (soccer, baseball), coaches regularly substitute in new offensive or defensive player groups in hockey. Coaches select players for given shifts based on their evaluation of players' current value to the game, expectations of the future benefits and costs of

[1] Personal communication, April 2, 2021.
[2] Personal communication, May 20, 2019.

assigning that playing time to one player over another, and static and dynamic team strategy. Importantly, playing time determines the extent to which team members can showcase their skills and value to those in a position to reward them (Zhang, 2017, 2019). Every player a coach chooses to put on the ice means other players remain off the ice without that opportunity. Research shows positive associations between playing time and contract extensions, individual reputation, and monetary awards (Staw & Hoang, 1995; Zhang, 2017). These meaningful organizational and career rewards means that time on ice is a clear form of preferential treatment.

Independent variable: Penalties. We operationalize rule breaking using penalties accrued. These are codified in the *NHL Rulebook*, a 200-page manual detailing how the game should be played, the rules players are required to follow, and the penalties that are imposed for breaking them (NHL, 2023). Penalties are enforced by referees, who remove players from the game for a set number of minutes for each violation they observe, leaving the focal team shorthanded for that amount of time. Our primary independent variable is *penalty minutes (PIM)*, a rolling total of penalty minutes accrued by a given player in the five games prior to the focal game (in which we measure time on ice). This is a continuous measure, with higher numbers indicating more extreme rule breaking.

The NHL also differentiates penalties in terms of the severity of the infraction. *Minor* penalties are assessed for violating rules intended to ensure fair play (e.g., interfering with the goalkeeper, throwing one's stick, tripping an opponent) or to maintain standards of behavior (e.g., using obscene language). They rule players off the ice for two minutes each. *Major* penalties are assessed for more extreme violations; most actions that elicit major penalties could cause serious injury to other players (e.g., fighting, headbutting). They rule players off the ice for five minutes each. In some cases the same behavior can incur either a minor or a major penalty depending on the referee's evaluation regarding the egregiousness of the behavior. For example, high-sticking (a player's stick touching another player above their shoulders) typically elicits a minor penalty, but if a player engages in high-sticking with excessive force or causes injury, it elicits a major penalty. In addition to the time difference between minor and major penalties, minor penalties are cancelled if the opposing team scores during the penalty, while major penalties are not, making major penalties more costly to the team in multiple ways. We calculate penalty minutes arising exclusively from *minor penalties* in the previous five games *(MIP)* and penalty minutes arising exclusively from *major penalties* in the previous five games *(MAP)* separately, as they impose different costs to the team.

As in many sports, several nuances in the rules can affect how penalties are served. Some penalties that occur at the same time between two players of opposing teams "offset" each other, such that penalized players are replaced while serving their penalty time. "Offset" penalties leave neither team shorthanded, and so are arguably less costly than penalties that are not offset. However, offset penalties both tend to be relatively infrequent (roughly 11.5% of penalties in our sample) and nonetheless lead to strategic and logistical costs (i.e., other team members need to replace them, causing issues with on-ice team composition, as well as team and opposition-focused strategies).

Control variables. We control for two main factors that may also influence player ice time. First, as players' historical performance is relevant to the amount they are chosen to play, we hold two primary individual-level performance indicators constant: *plusminus* and *Corsi percentage*. *Plusminus* measures a player's contribution to team-level outcomes by calculating the player's team's total scoring minus the opposing team's total scoring while that player was on ice over the previous five games. Higher values indicate the player's team outscored the opponent's team when the player was on the ice. *Corsi percentage* is a measure of performance that examines non-goal-related performance outcomes. Specifically, it represents the percentage of time that a player was on ice when their team controlled the puck. It is calculated as the sum of the player's team's shots on goal, missed shots, and blocked shots, over the sum of the opposing team's shots on goal, missed shots, and blocked shots, at even strength. If an individual player has a 50% Corsi percentage, it means that their team had as many shots on the opposing team's goal as the opposing team had on their own goal, at even strength, when that player was on the ice. We sum the player's Corsi percentage for the previous five games.

Second, we also control for *player–coach familiarity*, measured in terms of the number of games that a player has played with a given coach (Zhang, 2017). The more games a coach has observed a player for, the more information they have about that player's underlying quality, allowing them to rely less on rule breaking as a factor motivating selection decisions.

Estimation Strategy

In line with recent literature, we estimate linear panel models with fixed effects for player, coach, and season (Zhang, 2017, 2019). By including these fixed effects, our estimates hold constant everything time-invariant for players and coaches during a given season, such as a player's innate ability or a coach's skill (Halaby, 2004; Zhang, 2017). The only variables

we can include beyond these fixed effects are those that do change within a season for a player, such as game-to-game performance or game type (e.g., playoff vs. regular season). All regressions also include overtime fixed effects to account for variance in the length (in minutes) of a given game. Our baseline estimation strategy tests how penalties accrued by an individual player in the previous five games affect the number of minutes they are chosen to play in a focal game (Zhang, 2017). As in Zhang (2019), we cluster our standard errors at the player and coach level.

Our main analyses restrict the sample in two ways. In line with previous literature, we exclude goalkeepers, who tend to play for entire games (and thus have essentially no variation in ice time per game), and accrue penalty minutes very rarely—factors that contribute to removing them in similar analyses (Kakkar, Sivanathan & Gobel, 2020). We also exclude observations for players who accrued 20 or more penalty minutes in the prior five games (<1% of the sample), as there are certain players ("enforcers") who occupy a unique role in ice hockey that specializes in activity that accrues penalties (see Stuart & Moore, 2017), which this set of observations would include. We conduct a robustness test for these observations. Our final sample after applying these restrictions consists of 213,850 player–game observations.

Results

Primary results. Our primary results use our continuous measure of rule breaking. Summary statistics and correlations are included in Table 1. We report the effect of rule breaking on preferential treatment in Table 2. Model 1 includes the control variables alone. Model 2 tests the relationship between the total number of penalty minutes accrued in the prior five games and playing time in a focal game, controlling for a player's performance (*plusminus*, *Corsi percentage*) and coaches' knowledge of players (*player–coach familiarity*). Our continuous measure of rule breaking is associated with more ice time in the focal game ($\beta = .12$, $p < .001$).

However, it seems unlikely that coaches would reward rule breaking irrespective of its extremity. To understand whether the favorable treatment of rule breakers is bounded, we tested for a curvilinear effect following the guidelines provided by Haans, Pieters, and He (2016). Model 3 adds the squared term for penalty minutes to the model. The main coefficient remains significant and positive ($\beta = .18$, $p < .001$) and the squared term is significant and negative ($\beta = -.04$, $p < .001$), indicating that coaches' preferential treatment of rule breaking reverses at higher levels. Model 4 uses Simonsohn's (2018) two-lines procedure to test the significance of the simple slopes before and after the inflection point (about nine minutes of penalties in the prior five games). This analysis shows that the slope of the line up to the inflection point (X_{low}) is significant and positive ($\beta = .05$, $p < .001$), and the slope of the line from the inflection point onward (X_{high}) is significant and negative ($\beta = -.03$, $p = .04$; see Figure 1). The significance and opposite sign of both these slopes indicate that the effect is truly curvilinear.

Robustness tests. We conducted several robustness tests to confirm the validity of these primary findings. First, we ran models without control variables to determine whether our results were artificially inflated by multicollinearity. We report these results in Models 5–7 of Table 3. The direction and significance of our primary independent variables are unchanged. A second concern might be that our time-on-ice estimate is biased. If players commit penalties in the focal game in which we estimate time on ice, it reduces the amount of potential playing time a coach can allocate to them, compared to the players who accrued no penalties in that game. Model 8 provides an estimate excluding all game observations in which the focal player accrued any penalty minutes, to equalize the potential minutes of play available for each player observation. Finally, we vary the window of previous games used to calculate prior penalty minutes to be both smaller and larger than the five games in our baseline estimation (1, 3, 7, and 10 game windows). We

TABLE 1
Study 1: Summary Statistics

	Mean	*SD*	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Time on ice	16.77	4.820	—						
(2) PIM	2.510	3.136	−.03	—					
(3) MIP	0.817	0.986	.08	.65	—				
(4) MAP	0.085	0.322	−.15	.65	.08	—			
(5) Plusminus	0.030	2.478	.08	−.01	.00	−.02	—		
(6) Corsi percentage	250.8	43.19	.18	−.04	.03	−.09	.11	—	
(7) Player–coach familiarity	83.32	74.66	.20	−.01	−.01	−.04	.04	.08	—

Note: n = 213,850 player-game observations. Significance levels are omitted as repeated within-player observations bias correlations.

TABLE 2
Study 1: Main and Curvilinear Effects of Rule Breaking on Preferential Treatment

	Model 1		Model 2		Model 3		Model 4	
DV = Time on ice	**β**	***SE***	**β**	***SE***	**β**	***SE***	**β**	***SE***
Controls								
Plusminus	.13**	(.01)	.13**	(.01)	.13**	(.01)	.13**	(.01)
Corsi percentage	.20**	(.02)	.20**	(.02)	.20**	(.02)	.20**	(.02)
Player–coach familiarity	.44**	(.04)	.44**	(.04)	.44**	(.04)	.44**	(.04)
Main independent variable								
PIM			.12**	(.01)	.18**	(.02)		
PIM × PIM					−.04**	(.01)		
Two-lines test								
X_{low} (PIM)							.05**	(.00)
X_{high} (PIM)							−.03*	(.01)
High (PIM)							.07	(.07)
Player, coach, and season fixed effects and overtime controls	Yes		Yes		Yes		Yes	
Observations	213,850		213,850		213,850		213,850	

Notes: Robust standard errors clustered on player and coach level in parentheses. All non-dummy variables are standardized with a mean of zero and standard deviation of one. For the two-lines test only, we used the non-standardized values to correctly estimate slopes based on the turning point.
*$p < .05$
**$p < .01$.

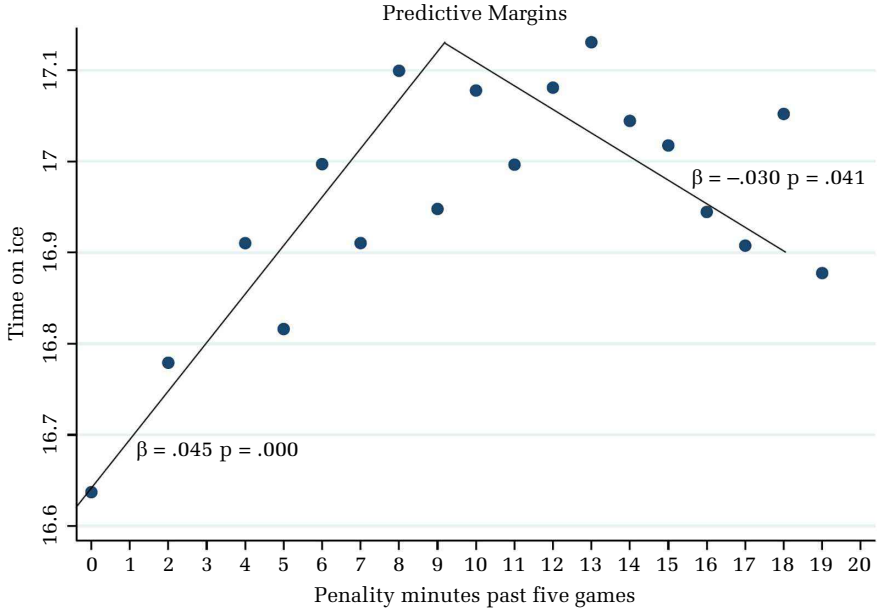
report these in Models 9–12 (Table 4), providing reassurance that our effects are not driven by the five-game window used in our primary models. In all these robustness checks (Models 5–12), the main effect and curvilinear term remain significant, indicating that our effects are robust after adjusting for a variety of potential sources of bias.

Boundary Conditions

We conducted several additional analyses to explore potential contextual, individual, or team-level boundaries of these effects.

Type of rule breaking. It is reasonable to assume that coaches will not treat all types of rule breaking equally. More extreme rule breaking imposes a

FIGURE 1
Study 1: Effects of Rule Breaking (Penalty Minutes, Previous Five Games) on Preferential Treatment (Playing Time, Focal Game)



Note: As there are neither one- nor three-minute penalties, there are no point estimates for these values.

TABLE 3
Study 1: Robustness Tests

	Model 5		Model 6		Model 7		Model 8	
	PIM → TOI		PIM × PIM → TOI		Two-Lines Test			
	Without Controls		Without Controls		Without Controls		Games Without Player Penalties	
DV = TOI	**β**	***SE***	**β**	***SE***	**β**	***SE***	**β**	***SE***
Controls								
Plusminus							.13**	(.01)
Corsi percentage							.21**	(.02)
Player–coach familiarity							.45**	(.04)
Main independent variable								
PIM	.12**	(.01)	.18**	(.02)			.18**	(.02)
PIM × PIM			−.04**	(.01)			−.04**	(.01)
Two-lines test								
X_{low} (PIM)					.05**	(.00)		
X_{high} (PIM)					−.03*	(.01)		
High (PIM)					.08	(.07)		
Player, coach, and season fixed effects, overtime controls	Yes		Yes		Yes		Yes	
Observations	213,850		213,850		213,850		174,337	

Notes: Robust standard errors clustered on player and coach level in parentheses. All non-dummy variables are standardized with a mean of zero and standard deviation of one. TOI = time on ice.
*$p < .05$
**$p < .01$

higher cost on teams in terms of how long it commits them to shorthanded play. Our primary models suggest that coaches do treat more extreme rule breakers less favorably than moderate ones, in terms of the *number* of penalty minutes a player has accrued in the prior five games. However, another way to differentiate moderate and extreme rule breaking is by rule breaking *severity*. We explore whether coaches treat players' penalties differently as a function of whether the *NHL Official Rules* designates those penalties as minor (less serious) or major (more egregious).

In Table 5, we find a positive and significant relationship between a player's minor penalties and playing time ($\beta_{MIP} = .08$, $p < .001$), and the squared term, while smaller, does not reverse in sign ($\beta_{MIP}{}^2 = .02$, $p < .001$, Model 13). This suggests that the preferential treatment a player receives for incurring minor penalties levels out but does not decline. On the other hand, penalty minutes accrued for major penalties are unrelated to playing time (Model 14). This contrast indicates in a different way that the preference for rule breakers appears to be restricted to moderate forms of

TABLE 4
Study 1: Using Different Windows (Number of Previous Games) to Calculate Penalty Minutes

	Model 9		Model 10		Model 11		Model 12	
	Ten Games		Six Games		Three Games		One Game	
DV = Time on ice	**β**	***SE***	**β**	***SE***	**β**	***SE***	**β**	***SE***
Controls								
Plusminus	.12**	(.01)	.13**	(.01)	.13**	(.01)	.13**	(.01)
Corsi percentage	.20**	(.02)	.20**	(.02)	.20**	(.02)	.20**	(.02)
Player–coach familiarity	.44**	(.04)	.44**	(.04)	.44**	(.04)	.44**	(.04)
Main independent variable								
PIM	.18**	(.02)	.16**	(.02)	.13**	(.01)	.08**	(.01)
PIM × PIM	−.02**	(.00)	−.02**	(.00)	−.02**	(.00)	−.01**	(.00)
Player, coach, and season fixed effects and overtime controls	Yes		Yes		Yes		Yes	
Observations	190,821		209,074		213,850		213,850	

Notes: Robust standard errors clustered on player and coach level in parentheses. All non-dummy variables are standardized with a mean of zero and standard deviation of one. The observations of Model 16 and Model 17 are the same as in the main models because the control variables are held constant at the level of five games prior.
**$p < .01$

TABLE 5
Study 1: Differentiating Moderate and Severe Rule Breaking

	Model 13		Model 14	
DV = Time on ice	**β**	***SE***	**β**	***SE***
Controls				
Plusminus	.13**	(.01)	.13**	(.01)
Corsi percentage	.20**	(.02)	.20**	(.02)
Player–coach familiarity	.45**	(.04)	.44**	(.04)
Main independent variable				
MIP	.08**	(.01)		
MIP × MIP	.02**	(.01)		
MAP			.02	(.02)
MAP × MAP			–.01	(.00)
Player, coach, and season fixed effects and overtime controls	Yes		Yes	
Observations	213,850		213,850	

Notes: Robust standard errors clustered on player and coach level in parentheses. All non-dummy variables are standardized with a mean of zero and standard deviation of one.
**$p < .01$

rule breaking—a preference that attenuates but does not reverse—and that this preference does not exist when rule breaking is instead severe. We note, however, that the average major penalty minutes for a player in the past five games is substantially lower and less varied ($M_{MAP} = .09$, $SD = .32$) than their average minor penalty minutes ($M_{MIP} = .82$, $SD = .99$), which may have contributed to the null findings here.

Type of player. Players have various styles of play, including differences in how often they break the rules. Some players play more "honorably" and elicit few penalties, while others make a career of breaking the rules. In Table 6, we explore how the frequency of players' rule breaking (compared to their teammates) affects coaches' preferential treatment. For those players whose total penalty minutes over that season is in the lowest quintile for their team ($PIM_{LowAvg} = 0.98$, $SD = 1.49$), there is no relationship between penalty minutes and time on ice ($\beta_{PIM} = .06$, $p = .18$; $\beta_{PIM}^2 = .03$, $p = .54$; Model 15). However, there is a positive linear association between rule breaking and time on ice for every other quintile of player (Models 16–19). Only for players at the top quintile of penalty minutes on their team ($PIM_{HighAvg} = 5.00$, $SD = 4.45$) is the linear and squared term significant ($\beta_{PIM} = .11$, $p < .001$; $\beta_{PIM}^2 = -.03$, $p = .02$; Model 20).

These results suggest that preferential treatment of rule breakers operates at every part of the distribution of players' rule breaking except among the most rule-abiding players. Indeed, more than half of the players receive only benefits for accruing more penalty

TABLE 6
Study 1: Player's Level of Rule Breaking (Relative to Their Team) as a Boundary Condition

	Model 15 Lowest RB Players		Model 16 Quintile Two		Model 17 Quintile Three		Model 18 Quintile Four		Model 19 Highest RB Players		Model 20 Enforcers Only	
DV = Time on Ice	**β**	***SE***	**β**	***SE***	**β**	***SE***	**β**	***SE***	**β**	***SE***	**β**	***SE***
Controls												
Plusminus	.14**	(.02)	.12**	(.02)	.12**	(.02)	.12**	(.02)	.13**	(.02)	.14	(.13)
Corsi percentage	.21**	(.03)	.21**	(.04)	.12**	(.04)	.13**	(.03)	.15**	(.04)	.17	(.13)
Player–coach familiarity	.43**	(.09)	.40**	(.10)	.47**	(.08)	.39**	(.08)	.14	(.09)	1.29**	(.25)
Independent variable												
PIM	.06	(.05)	.11**	(.03)	.08*	(.03)	.08*	(.03)	.11**	(.03)	–.21	(.28)
PIM × PIM	.03	(.06)	–.02	(.03)	–.01	(.02)	–.00	(.01)	–.03*	(.01)	.02	(.02)
Player, coach, and season fixed effects and overtime controls	Yes		Yes		Yes		Yes		Yes		Yes	
Observations	48,083		43,096		42,959		42,323		37,389		1,567	

Notes: Robust standard errors clustered on player and coach level in parentheses. All non-dummy variables are standardized with a mean of zero and standard deviation of one. RB = rule-breaking.
**$p < .01$

minutes. This preferential treatment only reverses for the highest quintile of rule breakers on any team. The final model (Model 20) of Table 6 reports results for the most extreme <1% of the sample (observations with 20 or more penalty minutes in the prior five games). Here, the pattern we find across the rest of the distribution does not hold at all. To the extent that this group represents "enforcers"—players who occupy a role that specializes in physical intimidation of opponents—this result is perhaps not surprising. Since their role is to engage in activity that elicits penalties, they spend little time on ice generally (Keneski, 2011), and coaches may be disinclined to reward them for eliciting penalties because doing so is the purpose of their ice time to begin with.

Type of team. Teams also differ in terms of how normative rule breaking is within them. These different cultures for rule breaking within teams likely influence preferential treatment (Borry, 2017). Thus, we explored whether the preference for rule breakers might be conditional on a team's rule breaking culture (Table 7, Models 21–25). We find that our effect is absent for teams in the lowest quintile of penalty minutes across teams for that season ($PIM_{LowTeamAvg} = 2.01$, $SD = 2.59$: $\beta_{PIM} = .05$, $p = .15$, $\beta_{PIM}^2 = -.01$, $p = .61$; Model 21), but holds for teams where rule breaking is more rampant ($PIM_{HighTeamAvg} = 3.08$, $SD = 3.66$: $\beta_{PIM} = .14$, $p < .001$, $\beta_{PIM}^2 = -.03$, $p = .01$; Model 25), highlighting that team culture remains important in moderating these effects.

Recent team performance. We also explore whether a team's recent performance moderates coaches' preference for rule breakers. If a team is on a winning streak, a coach may judge a player's rule breaking—and its cost—differently than if they are on a losing streak. In Table 8 (Models 26–31) we present results based on all combinations of teams' wins and losses in the previous five games, ranging from having lost all of their past five games to having won all of their past five games. We note that these models attend only to the aggregated number of wins and losses across games (i.e., combinations), and so do not capture order effects (i.e., permutations). For teams that have lost their past five games, there is a linear (but not curvilinear) effect of rule breaking, suggesting that coaches play rule breakers more after losing five games, which does not attenuate at the high end of players' rule breaking ($\beta_{PIM} = .20$, $p = .05$, $\beta_{PIM}^2 = .00$, $p = .99$; Model 26). However, if a team has won one (or more) of the past five games, the results are consistent with our primary linear and curvilinear effects, where coaches select players up to a certain level of rule breaking and then reverse this preference at high levels of rule breaking. Uniquely, when teams have been on a winning streak (winning five games in a row), players' rule breaking no longer has a significant influence on coaches' decisions to play them ($\beta_{PIM} = .11$, $p = .13$, $\beta_{PIM}^2 = -.00$, $p = .93$; Model 32).

Type of game. Finally, we explore whether our results hold for playoff games. Playoff games have higher stakes than regular season games, as they require teams to compete in a best of seven game series, which limits their ability to correct for bad performance. In addition, given the importance these games have in determining the season champion, it is possible that referees pay closer attention to infractions of league regulations. Given these differences, coaches may treat players' rule breaking differently

TABLE 7
Study 1: Team Rule-Breaking Culture as a Boundary Condition

	Model 21 Bottom Quintile Team Rule Breaking		Model 22 Quintile Two		Model 23 Quintile Three		Model 24 Quintile Four		Model 25 Top Quintile Team Rule Breaking	
DV = Time on Ice	**β**	***SE***	**β**	***SE***	**β**	***SE***	**β**	***SE***	**β**	***SE***
Controls										
Plusminus	.12**	(.02)	.12**	(.02)	.15**	(.02)	.14**	(.02)	.14**	(.02)
Corsi percentage	.20**	(.04)	.19**	(.03)	.16**	(.03)	.14**	(.03)	.19**	(.03)
Player–coach familiarity	.43**	(.11)	.30**	(.08)	.52**	(.11)	.45**	(.09)	.36**	(.11)
Independent variable										
PIM	.05	(.04)	.12**	(.03)	.13**	(.04)	.11**	(.03)	.14**	(.03)
PIM × PIM	–.01	(.02)	–.00	(.02)	–.04**	(.01)	–.00	(.01)	–.03*	(.01)
Player, coach, and season fixed effects and overtime controls	Yes		Yes		Yes		Yes		Yes	
Observations	43,068		42,780		42,715		42,780		42,507	

Notes: Robust standard errors clustered on player and coach level in parentheses. All non-dummy variables are standardized with a mean of zero and standard deviation of one.

*$p < .05$

**$p < .01$

TABLE 8
Study 1: Recent Team Performance as a Boundary Condition

	Model 26 Lost Previous Five Games		Model 27 One Win, Four Losses		Model 28 Two Wins, Three Losses		Model 29 Three Wins, Two Losses		Model 30 Four Wins, One Loss		Model 31 Won Previous Five Games	
DV = Time on ice	**β**	***SE***	**β**	***SE***	**β**	***SE***	**β**	***SE***	**β**	***SE***	**β**	***SE***
Controls												
Plusminus	–.03	(.06)	.04,†	(.02)	.09**	(.02)	.23**	(.02)	.30**	(.02)	.37**	(.05)
Corsi percentage	.31**	(.08)	.26**	(.03)	.24**	(.03)	.15**	(.03)	.14**	(.03)	.03	(.06)
Player–coach familiarity	.55**	(.11)	.41**	(.05)	.43**	(.05)	.42**	(.04)	.47**	(.05)	.41**	(.08)
Independent variable												
PIM	.20*	(.10)	.16**	(.04)	.20**	(.03)	.19**	(.03)	.15**	(.04)	.11	(.08)
PIM × PIM	.00	(.04)	–.04**	(.01)	–.03*	(.01)	–.05**	(.01)	–.04**	(.01)	–.00	(.03)
Observations	6,405		31,394		64,113		68,159		35,555		8,224	

Notes: Robust standard errors clustered on player and coach level in parentheses. All non-dummy variables are standardized with a mean of zero and standard deviation of one.
$^{\dagger}p < .10$
$^{*}p < .05$
$^{**}p < .01$

than in regular season games. Consistent with this intuition, the results show that while the effect holds for regular season games ($\beta_{PIM} = .18$, $p < .001$; $\beta_{PIM}{}^2 = -.04$, $p < .001$; Table 9, Model 32), it does not carry over into playoff games ($\beta_{PIM} = .10$, $p = .12$; $\beta_{PIM}{}^2 = -.01$, $p = .71$; Model 33).

Rule breaking and game outcomes. Up to this point we have been agnostic about whether the players' rule breaking benefits team performance. On one hand, it could be that rule breakers are treated preferentially because they help teams win games. On the other hand, our results would be more counterintuitive if coaches reward players who engage in activity that hurts their chances of winning. We therefore wanted to determine whether a team's penalties in a game are associated with whether the team wins that game. A supplementary analysis of the relationship between a team's penalty minutes and game outcomes finds that rule breaking within a game is negatively related to winning that game ($\beta = -.05$, $p < .01$; Table 10, Model 34). This suggests that penalties are truly costly for teams, heightening curiosity about why coaches treat rule breakers favorably. We address this question in our next study.

Discussion

In a sample of 213,850 player–game observations over six seasons of professional hockey, we find that players who violate official game regulations—and are penalized by game officials for doing so—are rewarded by coaches with additional playing time

TABLE 9
Study 1: Type of Game (Regular Season and Playoffs) as a Boundary Condition

	Model 32 Regular Season Games		Model 33 Playoff Games	
DV = Time on ice	**β**	***SE***	**β**	***SE***
Controls				
Plusminus	.12**	(.01)	.18**	(.04)
Corsi percentage	.20**	(.02)	.25**	(.05)
Player–coach familiarity	.41**	(.04)	.73**	(.11)
Independent variable				
PIM	.18**	(.02)	.10	(.07)
PIM × PIM	–.04**	(.01)	–.01	(.03)
Player, coach, and season fixed effects and overtime controls	Yes		Yes	
Observations	196,223		17,627	

Notes: Robust standard errors clustered on player and coach level in parentheses. All non-dummy variables are standardized with a mean of zero and standard deviation of one.
$^{**}p < .01$

TABLE 10
Study 1: Association between Game Penalty Minutes and Game Wins

	Model 34		Model 35		Model 36	
	All Teams		Bottom Quintile PIM Teams		Top Quintile PIM Teams	
DV = Game Won	β	*SE*	β	*SE*	β	*SE*
Controls						
Home game	.41**	(.01)	.43**	(.01)	.35**	(.02)
Playoff	–.03	(.01)	–.12	(.03)	.12	(.06)
Win–loss ratio	.68**	(.00)	.87**	(.01)	.85**	(.01)
Independent variable						
Penalty minutes in focal game	–.05**	(.00)	–.05	(.01)	–.09**	(.01)
Observations	12,836		2,635		2,561	

Notes: Logistic regression with robust standard errors clustered on team level in parentheses. All non-dummy variables are standardized with a mean of zero and standard deviation of one.
**$p < .01$

in future games. However, this positive association between rule breaking and playing time reverses when the number of penalty minutes becomes extreme. Importantly, we also find several boundary conditions of this effect. Our primary finding holds for minor, rather than major, rule violations; for players who break the rules often rather than rarely; in teams where rule breaking is rampant rather than restrained; when teams are suffering from a string of losses rather than on a winning streak; and for lower-stakes (regular season) rather than higher-stakes (playoff) games. We discuss the implications of these findings in the general discussion.

STUDY 2: EXPERIMENT

Though we find in Study 1 that coaches reward rule breakers with additional playing time, the nature of our data do not allow us to explore *why* coaches make these choices. The fact that rule breaking is costly to teams motivates the question as to why coaches treat rule breakers preferentially at all. Study 1 hints that the preferential treatment is not due to a (direct) performance benefit, as we find a negative relationship between rule breaking and team performance. However, there is evidence that supervisors do reward employees whom they believe are highly committed to their teams and organizations (Shore, Bommer & Shore, 2008). To fuel these beliefs, employees need to signal to their supervisors that they are willing to go *above and beyond* what is expected of them (Brown, 1996), and engage in behavior that sends these signals (Cristea & Leonardi, 2019). Nevertheless, rule breaking is costly—as we see in Study 1, as well as more generally (MacLean, 2001)—even though it may be common (Martin et al., 2013; Tyler & Blader, 2005; Veiga et al., 2004). This means that rule breakers also represent, and will likely be seen as, a liability to their teams, particularly when their rule breaking is extreme. These countervailing perceptions exist in tension: supervisors might reward rule breaking to the extent they perceive it as a signal of the rule breaker's commitment, but not to the extent that their rule breaking is perceived as a liability to the team achieving their goals.

To explore how these perceptions (of a player's commitment to and liability for their teams) work in concert to influence supervisory treatment, we conduct an experiment in which we manipulate team members' levels of rule breaking to determine (1) whether those in a position to reward potential rule breakers hold these perceptions, and (2) whether they then influence selection across levels of rule breaking. By manipulating how egregiously players in the experimental game break the rules while holding other aspects of the situation constant, we can confirm with more causal certainty that rule breaking is driving the preferential treatment we observe, as well as exploring reasons why this preference might exist.

Method

Participants. We recruited 224 participants (M_{age} = 31.08, SD = 10.54, 59.8% female) from the online survey platform Prolific Academic in exchange for £0.50. *Post hoc* analyses of the measured effect size suggest that the sample size provided more than adequate power ($1 - \beta > .95$) to detect selection decisions, and sensitivity analysis suggests that the sample could detect an odds ratio of 0.60. Following recommendations to improve data quality (Peer, Vosgerau & Acquisti, 2014), we restricted recruitment to individuals who had completed at least 50 studies and had been approved on 95% of those studies. Note that participants are paid for honest and accurate performance on the platform, conditions they agree to when signing up on the platform and in the study

itself. As a result, not following the rules can lead to not being paid for the study, and can affect their reputation on the platform and access to future studies. This loss of payment, reputation, and potential exclusion from future studies, and thus income, represent a significant cost to participants in this context (Peer et al., 2014).

Procedure. We informed participants that they would play a team game that required a six-sided die. We provided a link to an online die rolling website (www.random.org/dice/). We told participants that that they would be grouped into teams of five and assigned to one of two roles: player or advisor. Players would be individually responsible for reporting the outcome of their die rolls across three rounds of the game. Advisors would not roll dice but would report their opinions of team members and select one player for a bonus round of the game. We informed participants that each team—four players and one advisor—would compete against other five-person teams in three rounds of the game and that each player on the team with the highest score would win £2.00.

In each round of the game, players would roll their die five times and report the result of each roll. The sum of all four players' rolls on each team would then represent the team's score in each round. To inform their evaluations of each player, at the end of each round, advisors would see the rolls of each player and the resulting team score. Before beginning, participants read the rules of the game, which included a rule specifying that players needed to report their die rolls accurately. We informed them that the likelihood of rolling any number once was 16%, the same number twice in a row was 2.8%, and the same number across all 15 rolls of the three rounds of the game was 0.000000000002%, reinforcing the statistical (im)probability of various patterns of die rolls, as well as specifying the sort of performance that would almost certainly involve breaking the rules.

Rule-breaking manipulation. Unbeknownst to them, all participants were assigned to the advisor role, and the rolls of the other four members were manipulated (see Appendix A for tables of the rolls and mean scores in each condition). In all conditions, Players 1, 3, and 4 reported rolls consistently averaging somewhere between 3.4 and 3.6 in each round, performance that approximates the long-run statistical average of die rolls (i.e., 3.5). However, we assigned advisors randomly to one of three conditions that manipulated the reported rolls of Player 2 and the extent to which they broke the rules of the game.

In our control (*rule-abiding*) condition (n = 73), Player 2 behaved in a way that signaled honest play, reporting rolls that averaged 3.5 across all three rounds (range: 3.4–3.6), which represents what is statistically likely in this context. In the other two conditions, Player 2 reported rolls significantly higher than what is statistically probable, to elicit the perception that they were breaking the rules in a game that asked players to report their rolls accurately. In the *moderate* rule-breaking condition (n = 75), Player 2 reported rolls resulting in an *average* of five across all 15 rolls. We designed this condition so that participants would sense that this player was almost certainly violating the rules, but not so egregiously that the rule breaking was undeniable. In the *extreme* rule-breaking condition (n = 76), Player 2 reported *all* fives across the 15 rolls. Having explained this outcome as nearly impossible in the instructions, we intended to remove any doubt that this player was flouting the rules egregiously. Both rule breaking conditions held constant the amount Player 2 contributed to the team's performance.[3]

Perceptions of Player 2. After each "round" (every five rolls), we presented advisors with a table showing what each of the four players on their team reported rolling (Appendix A), and asked them to indicate their agreement (1 = completely disagree, 7 = completely agree) with two statements for each player: "*Player* N *is willing to do anything necessary to ensure the team's success*" and "*Player* N *is a liability to the team.*" We purposely used short measures to minimize interruptions in advisors' evaluations of the players (for the reliability of single-item measures, see Bergkvist & Rossiter, 2007). At the end of three rounds, we also asked advisors to indicate to what extent (1 = not at all, 7 = to a great extent) each player was "*breaking the rules*" and "*cheating,*" which we averaged as a measure of perceived rule breaking (α = .97).

Favorable treatment. At the end of three rounds, we asked advisors to choose one player to compete in a bonus round, the results of which would be doubled and added to the team score. We use the binary decision of whether the advisor selected Player 2 to compete in the bonus round (1 = selected, 0 = not selected) as our measure of favorable treatment.

[3] We collected data for a fourth condition in which Player 2 rolls all 6s (n = 76). We do not report the results for this condition in the main text because the "All 6" condition confounds performance and rule breaking (players in the "All 6" condition contribute more to team performance than the "Average 5" condition). In a model with all four conditions, including both the "All 5" and "All 6" conditions, there are no meaningful differences between these two conditions, and thus, for parsimony and clarity of explanation, our primary text reports only the two rule breaking conditions that hold performance constant.

Results

Perceptions of rule breaking. A one-way analysis of variance (ANOVA) confirmed that perceptions of Player 2's rule breaking significantly differed by condition, $F(2,221) = 91.33$, $p < .001$, $\eta_p^2 = .45$. They were lowest in the *rule-abiding* condition ($M = 1.66$, $SD = 1.03$), followed by the moderate rule-breaking condition ($M = 3.86$, $SD = 2.14$) and the extreme rule-breaking condition ($M = 5.70$, $SD = 2.06$). All mean differences were significant at $p < .001$.

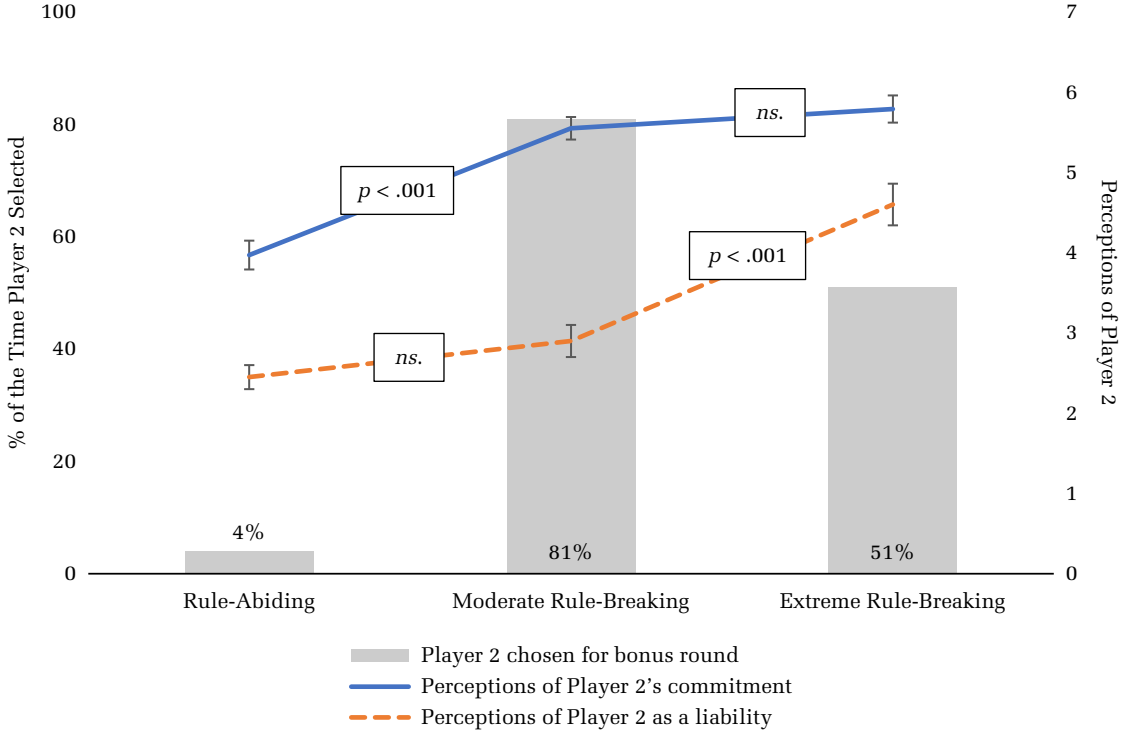
Perceived commitment. We also found a significant relationship between rule breaking and ratings of Player 2's commitment, $F(2,221) = 35.01$, $p < .001$, $\eta_p^2 = .24$. As the solid line of Figure 2 shows, advisors perceived both moderate ($M = 5.55$, $SD = 1.24$, $p < .001$) and extreme ($M = 5.79$, $SD = 1.48$, $p < .001$) rule breakers as more committed than rule-abiding players ($M = 3.97$, $SD = 1.57$), but moderate and extreme rule breakers were perceived as equally committed ($p = .90$).

Perceived liability. There was also a positive association between rule breaking and perceptions that Player 2 was a liability to the team, $F(2,221) = 29.38$, $p < .001$, $\eta_p^2 = .210$. However, unlike perceptions of commitment, rule-breaking severity significantly influenced these perceptions. When Player 2 only broke the rules moderately, their perceived liability ($M = 2.90$, $SD = 1.75$) did not differ from rule-abiding Player 2 ($M = 2.45$, $SD = 1.30$, $p = .41$: see Figure 2, dashed line). However, when Player 2 was an extreme rule breaker they *were* perceived as significantly more of a liability ($M = 4.59$, $SD = 2.23$) than rule-abiding ($p < .001$) or moderate rule breaker ($p < .001$) Player 2. In sum, liability perceptions only increased significantly at extreme levels of rule breaking.

Favorable treatment. Figure 2 also makes clear that advisors chose Player 2 for the bonus round most often when they were a moderate rule breaker (61 of 75, 81% of the time), significantly more often than rule-abiding players (3 of 73, 4% of the time: $\chi^2(1,148) = 89.89$, $p < .001$) or extreme rule breakers (39 of 76, 51% of the time: $\chi^2(1,151) = 15.21$, $p < .001$). Extreme rule breakers were still selected significantly more than rule-abiding players ($\chi^2(1,149) = 40.99$, $p < .001$). This pattern replicates the curvilinear effect of rule breaking on selection identified in Study 1.

We also tested the relationship between perceptions of Player 2's commitment and liability and the likelihood they were selected for the bonus round. In a logistic regression, perceptions of commitment ($\beta = .46$, $\exp(\beta) = 1.58$, $p > .001$) increased the odds

FIGURE 2
Study 2: Ratings of Player 2's Commitment and Liability, and Their Likelihood of Being Selected for the Bonus Round, by Experimental Condition



Note: Selection in the study was a choice between four players for the bonus round (Player 2 selected = 1, Player 2 not selected = 0).

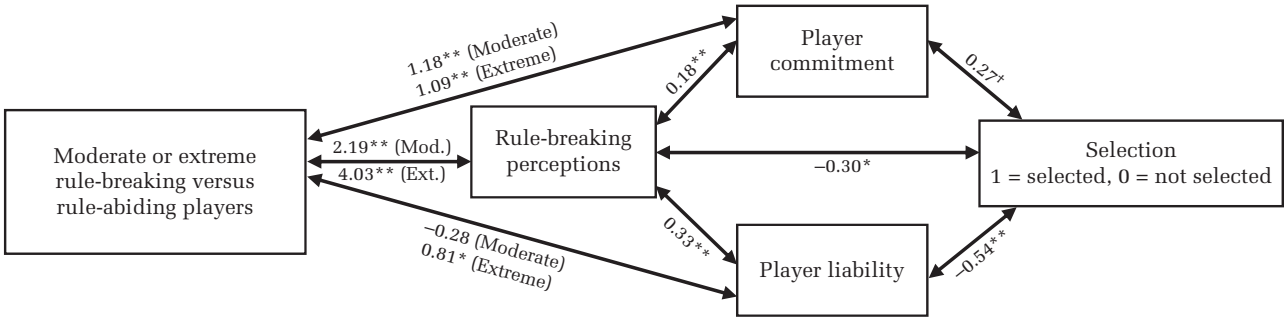
of selecting Player 2 by 58%, and perceptions of Player 2's liability ($\beta = -.39$, $\exp(\beta) = .68$, $p > .001$) decreased it by 32%, indicating clear relationships between these perceptions and the likelihood of being selected for the bonus round.

Indirect effects. Our objective in the experiment was to create unambiguous perceptions of Player 2 as a rule-abiding player, a moderate rule breaker, or an extreme rule breaker. We believed these perceptions would lead to attributions about Player 2 as committed to or a liability for their team, ultimately predicting whether they were chosen for the bonus round. Thus, we conducted a logistic mediation model with 5,000 bootstrapped samples using Model 81 of Hayes's (2018) Process model. This model (see Figure 3) tests commitment and liability as two parallel mediators stemming from perceptions of Player 2's rule breaking (which we experimentally manipulated; see Lench, Taylor & Bench, 2014). We entered our experimental condition as a multicategorical independent variable, using indicator coding with the rule-abiding condition as the reference category, rule-breaking perceptions as the intermediate step driving perceptions of Player 2's commitment and liability, which then formed parallel mediators in the relationship between rule breaking and selection. The relative direct effects indicate that moderate rule breakers were six times more likely (effect = 5.94, 95% CI = [4.34, 7.53]) and extreme rule breakers five and a half times more likely (effect = 5.61, 95% CI = [3.84, 7.39]) to be selected for the bonus round than rule-abiding Players 2s.

The analysis also found several indirect effects of rule breaking on selection. Moderate rule breakers (relative to the rule-abiding condition) were more likely to be chosen for the bonus round as a function of being perceived as more committed, both in general (effect = .32, 95% CI = [.02, .82]), and as a function of the degree to which they broke the rules (effect = .11, 95% CI = [.01, .30]). On the other hand, they were not more likely to be chosen for the bonus round because they were perceived as a liability generally (effect = .15, 95% CI = [–.13, .48]), but rather through the perception that they broke the rules increasing perceptions of them as a liability, which in turn reduced their likelihood of being selected (effect = –.39, 95% CI = [–.75, –.19]).

Extreme rule breakers (relative to the rule-abiding condition) were also more likely to be chosen for the bonus round, because they were perceived as more committed in general (effect = .30, 95% CI = [.02, .80]), and as a function of the degree to which they broke the rules (effect = .20, 95% CI = [.01, .52]). In contrast with moderate rule breakers, extreme rule breakers were consistently less likely to be chosen for the bonus round because of being perceived as a liability generally (effect = –.44, 95% CI = [–1.03, –.03]), as well as because perceptions of their rule breaking heightened these perceptions, in turn decreasing the likelihood of selection (effect = –.72, 95% CI = [–1.36, –.36]). These results suggest that breaking the rules creates both positive (commitment) and negative (liability) impressions, which influence selection decisions in the die rolling game.

FIGURE 3
Study 2: Players' Rule Breaking Severity (Experimentally Manipulated) on Selection for the Bonus Round, Mediated by Perceptions of Rule Breaking, Commitment, and Liability



Notes: Relative direct effects: moderate rule breaking (vs. rule-abiding) effect = 5.94, 95% CI = [4.34, 7.53]); extreme rule breaking (vs. rule-abiding) effect = 5.61, 95% CI = [3.84, 7.39].

Relative indirect (mediated) effects: *Via commitment*: moderate (vs. rule-abiding) effect = .32, 95% CI = [.02, .82]; extreme (vs. rule-abiding) effect = .30, 95% CI = [.02, .80]. *Via liability*: moderate (vs. rule-abiding) effect = .15, 95% CI = [–.13, .48]; extreme (vs. rule-abiding) effect = –.48, 95% CI = [–1.02, –.02]. *Via rule-breaking perceptions → commitment*: moderate (vs. rule-abiding) effect = .11, 95% CI = [.01, .30]; extreme (vs. rule-abiding) effect = .20 95% CI = [.02, .80]. *Via rule-breaking perceptions → liability*: moderate (vs. rule-abiding) effect = –.39, 95% CI = [–.75, –.19]; extreme (vs. rule-abiding): effect = –.72, 95% CI = [–1.37, –.36].

Selection was a binary choice (1 = selected for the bonus round; 0 = not selected). The figure reports regression coefficients from Model 81 of Hayes's (2018) PROCESS macro, predicting rule-breaking perceptions, commitment, liability, and selection. Relative direct and indirect effects of experimental condition on selection are reported as odds ratios.

Discussion

In Study 2, we use an experiment that manipulates levels of rule breaking to identify the causal relationships between players' level of rule breaking, their advisors' perceptions of their commitment to and liability for their team, and those advisors' selection decisions. We observe a similar curvilinear pattern of rule breaking on selection as we found in Study 1. Importantly, these results document how perceptions of commitment attenuate and perceptions of liability increase when rule breaking becomes severe. Importantly, we show that those with the power to reward "subordinates" perceive rule breakers differently depending on whether they break the rules at moderate or extreme levels. Moderate rule breakers appear to be rewarded most because they are perceived as being highly committed without provoking the perception that they are a dangerous liability for their team.

GENERAL DISCUSSION

Our research provides new insights into how rule breaking functions in teams, and deepens our understanding of how and why rule breaking can become endemic in organizations. Existing work in this area has largely focused on how rule breaking is punished (Dahling et al., 2012; Podsakoff, Bommer, Podsakoff & MacKenzie, 2006) or tolerated (Coser, 1962; Dentler & Erikson, 1959; Jetten & Hornsey, 2014; Jetten et al., 2010) by other group members. Our findings challenge this dominant perspective by focusing on how those in positions of power treat rule breakers preferentially. In both data from established NHL teams and an experiment that manipulates rule breaking directly, we find that rule breaking is rewarded by one's supervisor, though in a bounded way. This primary finding offers reasons for concern as well as optimism.

Rule breaking is rewarded most when it is moderate, both in amount and in form. The preferential treatment that supervisors show rule breakers declines when it becomes extreme. While it is perhaps a relief that there are limits to the preferential treatment supervisors show rule breakers, it is concerning that "moderate" rule breakers (in both studies) receive preferential treatment that is so substantive. NHL players receive the most time on ice (controlling for their legitimate performance) when they have accrued nine penalty minutes over the previous five games, which represents an average of roughly one minor penalty a game. This benefit for rule breakers also occurs in a context where rule breaking clearly and unequivocally reduces a team's chances of winning, suggesting that the NHL offers a conservative test of how positively rule breakers might be treated by their supervisors. In cases where rule breaking might enhance team performance, the benefits they receive might be even greater. This was the case in the experiment, where the moderate rule breaker dishonestly "earned" or "took" 150% of the points that a rule-abiding player would average and was chosen *20 times* more often for the bonus round than the rule-abiding player.

Constraints on Preferential Treatment of Rule Breakers

Importantly, our primary finding is bounded by several factors circumscribing *to whom*, *where*, and *when* this effect holds.

Who. Team members vary in their rule-breaking proclivities; some break them rarely, and others a great deal. We find that where a team member sits in this distribution on their team influences whether and to what extent they are rewarded for breaking the rules. In the NHL, players whose rule breaking for that season was in the bottom 20% of the team distribution receive no preferential treatment of breaking the rules. In contrast, players in every other quintile *are* rewarded for rule breaking, and *only* the highest quintile sees a significant drop in that preferential treatment. This is discouraging as it suggests that supervisors do perceive something worthy of rewarding in rule breaking behavior, just as Adoboli's supervisors appear to have done (Croft, 2012; Walker, 2012).

It is interesting to note that individuals at the most extreme ends of the rule-breaking distribution receive, in general, no preferential treatment when they do break the rules. For the most rule-abiding, this may be due to a "moral" reputation (Sperber & Baumard, 2012) that discourages supervisors from rewarding their rule breaking when they engage in it. This moral "halo" effect echoes recent work showing that individuals who signal their morality are less likely to be asked to engage in unethical behavior by their supervisors (Desai & Kouchaki, 2017). More rule-abiding subordinates likely elicit favorable treatment for other reasons, but our data do not speak to what those reasons might be. For the most extreme rule breakers, the results are less consistent. In the NHL data, we find a general trend toward continuing to treat rule breakers preferentially even at the higher end of the rule-breaking distribution, but players at the very highest end (observations in the top 1% of penalty minutes in the prior five games) receive no preferential treatment. The absence of any relationship here underscores the idea that there may be a unique role for "enforcers" in teams (Stuart & Moore, 2017), whose rewards for rule breaking may differ from those of their teammates.

Where. Individual behavior is always influenced by one's immediate context (Apel & Paternoster, 2009; Borry, 2017; Kuenzi, Mayer & Greenbaum, 2020). Thus, it is perhaps unsurprising that a team's local culture of rule breaking influences whether team members are rewarded for it. Players on the most rule-abiding teams—teams in the bottom quintile of penalty minutes among teams for that season—are also not rewarded for rule breaking, though, similar to the player-level moderation, the other 80% of teams *do* reward it. The claim of some theorists that "all organizations are inherently criminogenic" (Gross, 1978: 56) may be exaggerated. Nevertheless, when rule breaking can be perceived positively—either as prosocially motivated (Vardaman, Gondo & Allen, 2014), or as benefitting the bottom line or team performance (Apel & Paternoster, 2009)—the dominant mode appears to be to reward it.

When. Temporal dimensions also moderate this preferential treatment. Rule breaking is rewarded more after a team has been performing poorly than it is after the team has been performing well. This is consistent with research on how loss aversion motivates risk seeking (Kühberger, 1998) and unethical behavior (Cameron & Miller, 2009), as well as work showing that rule breaking is more common in highly competitive contexts (Desmet, Hoogervorst & Van Dijke, 2015), or when there is perceived scarcity in the environment (Staw & Szwajkowski, 1975). Additionally, coaches may reward rule breaking as a demonstration of commitment, particularly when their team is performing poorly, believing (for good reason) that grit is necessary to get through these slumps (Lucas, Gratch, Cheng & Marsella, 2015). In contrast, rule breaking is *not* preferred when the stakes of an individual game are higher (in our case, during the playoffs). It might be that supervisors perceive the liability of rule breakers differently in certain high-stakes contexts, motivating them to devalue rule breaking rather than reward it.

Reasons for the Favorable Treatment of Rule Breakers

We also explored reasons why supervisors might treat rule breakers preferentially. Rule breaking appears to signal a team member's commitment (Shore, Barksdale & Shore, 1995; Shore et al., 2008)—a willingness to do whatever it takes to get the job done (Brown, 1996). Thus, our findings include rule breaking as a unique addition to the list of extra-role behaviors (Van Dyne & LePine, 1998) that lead to perceptions of employee commitment—perceptions that are strongly associated with organizational rewards (Shore et al., 1995). As other work has found (Cristea & Leonardi, 2019; Werner, 1994), our results indicate that decision-makers value behaviors that signal employees' commitment even after objective measures of individual performance have been accounted for (Brown, 1996). Indeed, we control for performance that might explain why rule breakers are treated preferentially otherwise.

We show that supervisors do weigh both the costs and benefits of rewarding their team members (Hinds, Carley, Krackhardt & Wholey, 2000). Breaking the rules, even if doing so is prosocially motivated or performance-enhancing, does incur risk and can incur costly sanctions (Apel, 2013; Zhang et al., 2019). This makes rule breakers a liability to their groups, which, our results indicate, supervisors consider in their calculus of preferential treatment. However, it appears that these liability perceptions only matter when the rule breaking is truly severe and undeniable. Perhaps this is because when rule breaking is easiest to observe, it is harder to excuse in ways that might mitigate the liability it represents (Shaver, 2012).

Theories of Rule Breaking

Theories of rule breaking in organizations have focused largely on differentiating whether the rule breaker is motivated by personal gain (e.g., Mars, 1982) or by constructive (Vadera et al., 2013) or prosocial (Mo et al., 2023; Morrison, 2006) reasons. These motives are connected to the broader literature assessing decisions to punish rule breakers (Butterfield et al., 1996; Mooijman et al., 2015; Tyler & Boeckmann, 1997; Zhang et al., 2019) or to tolerate them (Coser, 1962; Dentler & Erikson, 1959; Jetten & Hornsey, 2014). We focus instead on the supervisor's point of view and remain agnostic about rule breakers' motivations. The supervisor's perspective is important. Regulatory authorities promulgate rules because they want to quash specific behavior—regardless of rule breakers' motives. Yet, if supervisors treat rule breakers positively for violating formal rules, the motivation that matters is the supervisors'—what motivated the rule breakers to violate the rules is less material.

The fact that we find a clear preference for rule breakers across two studies—but one that is bounded in several respects—indicates that there is much variance in how rule breaking functions and is treated in organizations. Our findings reaffirm that supervisors do treat subordinates differently, a key tenet in leadership theories that focus on leaders' individual relationships with their followers (Dansereau, 1995). The extent to which differential treatment of subordinates is functional in teams depends on the justifiability of the differential treatment (Chen, He & Weng, 2015). Subordinates' rule breaking does elicit differential treatment, which arguably serves a

dysfunctional purpose, raising concerns about its justifiability.

Addressing calls to move beyond linear effects (Mo et al., 2023), our study adds important nuance to our understanding of rule breaking in many ways. That our results differ as a function of the rule being violated reinforces the idea that organizational rules are not monolithic. Rather, they differ in form, severity, and implications (Lehman & Ramanujam, 2009; Martin et al., 2013). The positive implications we find for rule breaking, fortunately, seem to be limited to its "moderate" levels and forms. Past research has argued that the potential of getting caught and paying a price for breaking the rules plays an important role in decisions to break them (Allingham & Sandmo, 1972; Paternoster & Simpson, 1996). We find that the potential costs of rule breaking, as well as the positive signals it sends, inform how supervisors both perceive it and treat those who engage in it.

While more egregious rule breaking *is* treated less preferentially than moderate rule breaking in both studies, it is important to recognize that even the most egregious rule breakers are rewarded more often (for their rule breaking) than the most rule-abiding team members. This can be interpreted in two ways. An optimistic perspective focuses on the fact that what is being rewarded here is the rule breaking itself. Our findings do not imply that more rule-abiding team members are *not* treated preferentially, only that they are not treated preferentially *for rule breaking*. Our data do not allow us to make claims about what rule-abiding team members are rewarded for, but one could interpret the fact that they are simply not rewarded for rule breaking. The cynical view is that supervisors view their subordinates' rule breaking as advantageous to them. It is worth noting that in the experiment, the simplest explanation for choosing an extreme rule breaker more often than a rule-abiding player is likely the clear financial gain advisors know they will receive from doing so, without incurring any immediate personal risk, as they are collateral beneficiaries.

Theory on the Normalization of Deviance

Current scholarship on organizational wrongdoing tends to focus on how those in power often abuse it and break the rules themselves (Ashforth & Anand, 2003; Brief et al., 2001), and how this dysfunctional role modeling influences their subordinates to follow suit (Moore, Mayer, Chiang, Crossley, Karlesky & Birtch, 2019; Treviño & Brown, 2005). Here, we focus on something slightly different, where supervisors are not modeling misconduct themselves, but are providing favorable treatment to those who engage in it. We find that this preferential treatment occurs for the majority of teams and the majority of team members, indicating that—at least in this context—deviance is normalized. This supports the view that those in positions of power play a key role in advancing wrongdoing, even when they do not break the rules themselves (den Nieuwenboer, da Cunha & Treviño, 2017; Treviño & Brown, 2005). Helping rule breakers not only avoid punishment but receive rewards may facilitate "bottom up" corruption (Palmer, 2008), with the added benefit of "insulating" leaders from facing negative consequences for the morally compromised behavior of those they lead (Ashforth & Anand, 2003: 8).

Key to the normalization of deviance is managerial discretion. In both contexts we studied, supervisors had discretion over how to reward their team members and were not required to defend or explain their choices. This discretion affords them "wiggle room" (Dana, Weber & Kuang, 2007) to reward rule breaking without connecting their preferential treatment to the team members' problematic behavior. Managers tend to have wide latitude over employee treatment, particularly treatment that is not beholden to formalized organizational processes (Hambrick & Finkelstein, 1987). Discretion can lead to positive outcomes (e.g., top management teams have more gender diversity when female CEOs have greater discretion; see Corwin, Loncarich & Ridge, 2022), or negative ones (e.g., discretion has been associated with vehicle inspectors helping consumers evade emissions regulations; see Pierce & Snyder, 2012). When discretion means there is no way to connect a supervisor's preferential treatment to their subordinates' rule breaking, it will be easier to embed this behavior in the culture (Schaubroeck et al., 2012). In addition, once deviance of any kind is accepted as normative practice or the standard way of getting work done within a system (den Nieuwenboer et al., 2017; MacLean, 2001), it becomes extremely difficult to eradicate.

Limitations and Future Research

Our findings point to several potential avenues for future research. We focused on two contexts where breaking rules, and the negative consequences for breaking them, were codified clearly. However, rules are not always so explicitly defined; they are often ambiguous (Lehman & Ramanujam, 2009), fluid, and open to interpretation (Roulet, 2019). It would be interesting to explore whether breaking more ambiguous rules is rewarded even more often, or more highly, because of the variance in interpretation those rules provide. On the flip side, our contexts are limited to ones where rule breaking, while common, is formally proscribed. It would be helpful for future research to explore in more detail how contexts

might influences subsequent rewards. Do rule breakers still receive preferential treatment in contexts where rule breaking is required (e.g., the Mafia)? Conversely, we also cannot know whether rule breaking is rewarded similarly in contexts where it is truly rare. However, given evidence that rule breaking is common across individuals (Gächter & Schulz, 2016) and organizations (Martin et al., 2013; Veiga et al., 2004), these contexts may be fewer than we would like to admit.

Additionally, in the NHL context we find evidence that rule breaking is correlated negatively with winning regular season games. Though we cannot make a causal statement about this relationship, if rule breaking does undermine team performance, why are rule breakers selected to play more often? One possibility is that, if rule breaking is perceived as evidence of an employee's commitment, supervisors or organizations may spend less time monitoring them, which might let their rule-breaking behavior escalate over time. Regulators certainly noted how lax supervision of Kweku Adoboli facilitated his rogue trading (FINMA, 2012). Neglectful monitoring of rule breakers also insulates leaders from complete knowledge about it (Ashforth & Anand, 2003), which has the added advantage of shielding them from blame (Messick, 1999).

Less time monitoring rule breakers might allow teams and organizations to allocate attention to other functional areas, improving performance in *future* games while also fostering corruption. It would be interesting to understand whether, and when, perceptions that a rule breaker is highly committed allow supervisors to refocus their attention in ways that positively influence organizational performance. Alternatively, if supervisors are willing to undermine objective team performance to have committed players on teams, perhaps they trade performance for strong relationships with subordinates. Supporting the intuition behind this performance–relationship tradeoff, work shows that individuals are inclined toward those they like more in social networks rather than those who are more competent (Casciaro & Lobo, 2015). Supervisors may value committed players even if it undermines optimal team performance.

While we only explored preferential treatment in terms of supervisors' selection of certain team members over others, it would be interesting to connect rule breaking to financial rewards. Testing this in a compelling way would require a different context than the NHL, given the complexities of multiyear contracts. Nevertheless, as an opening effort, we matched the increase in playing time to salary data in our Study 1 sample. Our back-of-the-envelope (though admittedly non-causal) calculations hint that if an average player (one in the 50th percentile of playing time) who had previously been rule-abiding were to then incur nine minutes of penalties across a five-game window, it would translate to 45 seconds of additional playing time in the next game. Averaged across a season, this increase would move them into the 56th percentile of playing time. Players who play at the 56th percentile of playing time have an average salary that is $865,000 more than players at the 50th percentile of playing time. This represents more than a 30% increase in annual income. While any conclusions drawn about financial outcomes are necessarily crude, these numbers do suggest that future research on the economic consequences of rule breaking are warranted. For example, does a team member's contract, which represents their perceived value upon joining the organization, moderate the degree to which their rule breaking is rewarded?

Finally, our work hints at reasons why some groups may have a specific role for rule breakers within organizations. Extensive work has examined how individuals inhabit specific roles in groups (Biddle, 1986; Slater, 1955), and some teams include a specific informal role for individuals who specialize in breaking rules (Stuart & Moore, 2017). However, we do not know how these roles are created, and what sort of "gatekeepers" support their entry, rise, or exit. What is clear is that rule breaking becomes embedded in certain organizational cultures thanks to supervisory support (den Nieuwenboer et al., 2017; Roulet, 2019). Thus, it remains important for research to continue to explore how leaders support or inhibit rule breaking in organizations.

REFERENCES

Allingham, M. G., & Sandmo, A. 1972. Income tax evasion: A theoretical analysis. ***Journal of Public Economics***, 1: 323–338.

Ambrose, M. L., Taylor, R., & Hess, R. L., Jr, 2015. Can I help you? Employee prosocial rule breaking as a response to organizational mistreatment of customers. In P. L. Perrewé & J. R. B. Halbesleben & C. C. Rosen (Eds.), ***Mistreatment in organizations***: 1–31. Bingley, U.K.: Emerald Group.

Apel, R. 2013. Sanctions, perceptions, and crime: Implications for criminal deterrence. ***Journal of Quantitative Criminology***, 29: 67–101.

Apel, R., & Paternoster, R. 2009. Understanding "criminogenic" corporate culture: What white-collar crime researchers can learn from studies of the adolescent employment-crime relationship. In S. S. Simpson & D. Weisburd (Eds.), ***The criminology of white-collar crime***: 15–33. New York: Springer.

Ashforth, B. E., & Anand, V. 2003. The normalization of corruption in organizations. ***Research in Organizational Behavior***, 25: 1–52.

Balasubramanian, P., Bennett, V. M., & Pierce, L. 2017. The wages of dishonesty: The supply of cheating under high-powered incentives. ***Journal of Economic Behavior & Organization***, 137: 428–444.

Belmi, P., Barragan, R. C., Neale, M. A., & Cohen, G. L. 2015. Threats to social identity can trigger social deviance. ***Personality and Social Psychology Bulletin***, 41: 467–484.

Bergkvist, L., & Rossiter, J. R. 2007. The predictive validity of multiple-item versus single-item measures of the same constructs. ***Journal of Marketing Research***, 44: 175–184.

Biddle, B. J. 1986. Recent developments in role theory. ***Annual Review of Sociology***, 12: 67–92.

Bocian, K., & Wojciszke, B. 2014. Self-interest bias in moral judgments of others' actions. ***Personality and Social Psychology Bulletin***, 40: 898–909.

Borry, E. L. 2017. Ethical climate and rule bending: How organizational norms contribute to unintended rule consequences. ***Public Administration***, 95: 78–96.

Brief, A. P., Buttram, R. T., & Dukerich, J. M. 2001. Collective corruption in the corporate world: Toward a process model. In M. E. Turner (Ed.), ***Groups at work: Theory and research***: 471–499. Mahwah, NJ: Erlbaum.

Brown, R. B. 1996. Organizational commitment: Clarifying the concept and simplifying the existing construct typology. ***Journal of Vocational Behavior***, 49: 230–251.

Butterfield, K. D., Treviño, L. K., & Ball, G. A. 1996. Punishment from the manager's perspective: A grounded investigation and inductive model. ***Academy of Management Journal***, 39: 1479–1512.

Cameron, J. S., & Miller, D. T. 2009. Different ethical standards in gain versus loss frames. In D. De Cremer (Ed.), ***Psychological perspectives on ethical behavior and decision making***: 91–106. Charlotte, NC: Information Age.

Carlsmith, K. M., Darley, J. M., & Robinson, P. H. 2002. Why do we punish? Deterrence and just deserts as motives for punishment. ***Journal of Personality and Social Psychology***, 83: 284–299.

Casciaro, T., & Lobo, M. S. 2015. Affective primacy in intraorganizational task networks. ***Organization Science***, 26: 373–389.

Chen, X.-P., He, W., & Weng, L.-C. 2015. What is wrong with treating followers differently? The basis of leader–member exchange differentiation matters. ***Journal of Management***, 44: 946–971.

Corwin, E. S., Loncarich, H., & Ridge, J. W. 2022. What's it like inside the hive? Managerial discretion drives TMT gender diversity of women-led firms. ***Journal of Management***, 48: 1003–1034.

Coser, L. A. 1962. Some functions of deviant behavior and normative flexibility. ***American Journal of Sociology***, 68: 172–181.

Cristea, I. C., & Leonardi, P. M. 2019. Get noticed and die trying: Signals, sacrifice, and the production of face time in distributed work. ***Organization Science***, 30: 552–572.

Croft, J. 2012. November 20: Rise and fall of Adoboli the "family" man. ***Financial Times.***

Dahling, J. J., Chau, S. L., Mayer, D. M., & Gregory, J. B. 2012. Breaking rules for the right reasons? An investigation of pro-social rule breaking. ***Journal of Organizational Behavior***, 33: 21–42.

Dahling, J. J., & Gutworth, M. B. 2017. Loyal rebels? A test of the normative conflict model of constructive deviance. ***Journal of Organizational Behavior***, 38: 1167–1182.

Dana, J., Weber, R., & Kuang, J. 2007. Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. ***Economic Theory***, 33: 67–80.

Dansereau, F. 1995. A dyadic approach to leadership: Creating and nurturing this approach under fire. ***Leadership Quarterly***, 6: 479–490.

den Nieuwenboer, N. A., da Cunha, J. V., & Treviño, L. K. 2017. Middle managers and corruptive routine translation: The social production of deceptive performance. ***Organization Science***, 28: 781–803.

Dentler, R. A., & Erikson, K. T. 1959. The functions of deviance in groups. ***Social Problems***, 7: 98–107.

Desai, S. D., & Kouchaki, M. 2017. Moral symbols: A necklace of garlic against unethical requests. ***Academy of Management Journal***, 60: 7–28.

Desmet, P. T. M., Hoogervorst, N., & Van Dijke, M. 2015. Prophets vs. profits: How market competition influences leaders' disciplining behavior towards ethical transgressions. ***Leadership Quarterly***, 26: 1034–1050.

Fehr, R., Welsh, D., Yam, K. C., Baer, M., Wei, W., & Vaulont, M. 2019. The role of moral decoupling in the causes and consequences of unethical pro-organizational behavior. ***Organizational Behavior and Human Decision Processes***, 153: 27–40.

Fielding, K. S., Hogg, M. A., & Annandale, N. 2006. Reactions to positive deviance: Social identity and attribution dimensions. ***Group Processes and Intergroup Relations***, 9: 199–218.

Financial Conduct Authority. 2014. Final Notice: Mr. John Christopher Hughes.

FINMA. 2012. UBS trading losses in London: FINMA finds major control failures.

Fleming, C. J. 2020. Prosocial rule breaking at the street level: The roles of leaders, peers, and bureaucracy. ***Public Management Review***, 22: 1191–1216.

Gächter, S., & Schulz, J. F. 2016. Intrinsic honesty and the prevalence of rule violations across societies. ***Nature***, 531: 496–499.

Gazzoli, G., Chaker, N. N., Zablah, A. R., & Brown, T. J. 2022. Customer-focused voice and rule-breaking in the frontlines. ***Journal of the Academy of Marketing Science***, 50: 388–409.

Ghosh, A., & Shum, C. 2019. Why do employees break rules? Understanding organizational rule-breaking behaviors in hospitality. ***International Journal of Hospitality Management***, 81: 1–10.

Gong, T., Sun, P., & Kang, M. J. 2022. Customer-oriented constructive deviance as a reaction to organizational injustice toward customers. ***Cornell Hospitality Quarterly***, 63: 119–135.

Gross, E. 1978. Organizational crime: A theoretical perspective. In: N. K. Denzin (Ed.), ***Studies in symbolic interaction***, vol. 1: 55–85. Greenwich, CT: JAI Press.

Haans, R. F. J., Pieters, C., & He, Z.-L. 2016. Thinking about U: Theorizing and testing U- and inverted U-shaped relationships in strategy research. ***Strategic Management Journal***, 37: 1177–1195.

Halaby, C. N. 2004. Panel models in sociological research: Theory into practice. ***Annual Review of Sociology***, 30: 507–544.

Hambrick, D. C., & Finkelstein, S. 1987. Managerial discretion: A bridge between polar views of organizational outcomes. ***Research in Organizational Behavior***, 9: 369–406.

Hayes, A. F. 2018. ***Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*** (2nd ed.). New York: Guilford.

Hewitt, J. D., Poole, E. D., & Regoli, R. M. 1984. Self-reported and observed rule-breaking in prison: A look at disciplinary response. ***Justice Quarterly***, 1: 437–447.

Hinds, P. J., Carley, K. M., Krackhardt, D., & Wholey, D. 2000. Choosing work group members: Balancing similarity, competence, and familiarity. ***Organizational Behavior and Human Decision Processes***, 81: 226–251.

Hollinger, R. C. 1986. Acts against the workplace: Social bonding and employee deviance. ***Deviant Behavior***, 7: 53–75.

Irshad, M., Bartels, J., Majeed, M., & Bashir, S. 2022. When breaking the rule becomes necessary: The impact of leader–member exchange quality on nurses' pro-social rule-breaking. ***Nursing Open***, 9: 2289–2303.

Jetten, J., & Hornsey, M. J. 2014. Deviance and dissent in groups. ***Annual Review of Psychology***, 65: 461–485.

Jetten, J., Hornsey, M. J., Spears, R., Haslam, S. A., & Cowell, E. 2010. Rule transgressions in groups: The conditional nature of newcomers' willingness to confront deviance. ***European Journal of Social Psychology***, 40: 338–348.

Kakkar, H., Sivanathan, N., & Gobel, M. S. 2020. Fall from grace: The role of dominance and prestige in the punishment of high-status actors. ***Academy of Management Journal***, 63: 530–553.

Keneski, M. E. 2011. ***Hockey fights of yesteryear: A look back on the careers of classic NHL enforcers***. Seattle: CreateSpace.

Kim, S. K., & Zhan, Y. 2023. Breaking rules yet helpful for all: Beneficial effects of pro-customer rule breaking on employee outcomes. ***Journal of Organizational Behavior***, 44: 739–759.

Kuenzi, M., Mayer, D. M., & Greenbaum, R. L. 2020. Creating an ethical organizational environment: The relationship between ethical leadership, ethical organizational climate, and unethical behavior. ***Personnel Psychology***, 73: 43–71.

Kühberger, A. 1998. The influence of framing on risky decisions: A meta-analysis. ***Organizational Behavior and Human Decision Processes***, 75: 23–55.

Lehman, D. W., & Ramanujam, R. 2009. Selectivity in organizational rule violations. ***Academy of Management Review***, 34: 643–657.

Lench, H. C., Taylor, A. B., & Bench, S. W. 2014. An alternative approach to analysis of mental states in experimental social cognition research. ***Behavior Research Methods***, 46: 215–228.

Liao, Z., Yam, K. C., & Johnson, R. E. 2019. How does CSR reconcile employees' paradoxical responses to unethical pro-organizational behavior. ***Academy of Management Proceedings***, 2019: 15486.

Liu, X. L., Lu, J. G., Zhang, H., & Cai, Y. 2021. Helping the organization but hurting yourself: How employees' unethical pro-organizational behavior predicts work-to-life conflict. ***Organizational Behavior and Human Decision Processes***, 167: 88–100.

Lucas, G. M., Gratch, J., Cheng, L., & Marsella, S. 2015. When the going gets tough: Grit predicts costly perseverance. ***Journal of Research in Personality***, 59: 15–22.

MacLean, T. 2001. Thick as thieves: A social embeddedness model of rule breaking in organizations. ***Business and Society***, 40: 167–196.

Malik, L. R., & Mishra, M. 2023. Prosocial rule-breaking: A systematic literature review. ***Journal of Hospitality and Tourism Insights***, 7: 763–782.

March, J. C., & Simon, H. A. 1958. ***Organizations***. New York: Wiley.

Mars, G. 1982. ***Cheats at work: An anthropology of workplace crime***. London: George Allen & Unwin.

Martin, A. W., Lopez, S. H., Roscigno, V. J., & Hodson, R. 2013. Against the rules: Synthesizing types and processes of bureaucratic rule-breaking. ***Academy of Management Review***, 38: 550–574.

Messick, D. M. 1999. Dirty secrets: Strategic uses of ignorance and uncertainty. In L. L. Thompson & J. M. Levine & D. M. Messick (Eds.), ***Shared cognition in organizations: The management of knowledge***: 71–87. Mahwah, NJ: Lawrence Erlbaum.

Mishra, M., Ghosh, K., & Sharma, D. 2021. Unethical pro-organizational behavior: A systematic review and future research agenda. ***Journal of Business Ethics***, 179: 63–87.

Mo, S., Lupoli, M. J., Newman, A., & Umphress, E. E. 2023. Good intentions, bad behavior: A review and

synthesis of the literature on unethical prosocial behavior (UPB) at work. ***Journal of Organizational Behavior***, 44: 335–354.

Mooijman, M., van Dijk, W. W., Ellemers, N., & van Dijk, E. 2015. Why leaders punish: A power perspective. ***Journal of Personality and Social Psychology***, 109: 75–89.

Moore, C., Mayer, D. M., Chiang, F. F. T., Crossley, C., Karlesky, M. J., & Birtch, T. A. 2019. Leaders matter morally: The role of ethical leadership in shaping employee moral cognition and misconduct. ***Journal of Applied Psychology***, 104: 123–145.

Morrison, E. W. 2006. Doing the job well: An investigation of pro-social rule breaking. ***Journal of Management***, 32: 5–28.

National Hockey League. 2023. ***Official rules 2023-2024***. Ontario: Raster Graphics.

Okhuysen, G. A., & Bechky, B. A. 2009. Coordination in organizations: An integrative perspective. ***Academy of Management Annals***, 3: 463–502.

Okimoto, T. G., & Wenzel, M. 2009. Punishment as restoration of group and offender values following a transgression: Value consensus through symbolic labelling and offender reform. ***European Journal of Social Psychology***, 39: 346–367.

Palmer, D. 2008. Extending the process model of collective corruption. ***Research in Organizational Behavior***, 28: 107–135.

Paternoster, R., & Simpson, S. 1996. Sanction threats and appeals to morality: Testing a rational choice model of corporate crime. ***Law & Society Review***, 30: 549–584.

Peer, E., Vosgerau, J., & Acquisti, A. 2014. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. ***Behavior Research Methods***, 46: 1023–1031.

Pierce, L., & Snyder, J. A. 2012. Discretion and manipulation by experts: Evidence from a vehicle emissions policy change. ***B.E. Journal of Economic Analysis & Policy***, 13. doi:10.2139/ssrn.1831494

Podsakoff, P. M., Bommer, W. H., Podsakoff, N. P., & MacKenzie, S. B. 2006. Relationships between leader reward and punishment behavior and subordinate attitudes, perceptions, and behaviors: A meta-analytic review of existing and new research. ***Organizational Behavior and Human Decision Processes***, 99: 113–142.

Posner, E. A. 1997. Standards, rules, and social norms. ***Harvard Journal of Law and Public Policy***, 21: 101–117.

Rhee, Y. W. 2019. Varying effects of unethical pro-group behavior on team and individual performances. Abstract. ***Academy of Management Proceedings***, 2019: 16590.

Roulet, T. J. 2019. Sins for some, virtues for others: Media coverage of investment banks' misconduct and adherence to professional norms during the financial crisis. ***Human Relations***, 72: 1436–1463.

Schaubroeck, J. M., Hannah, S. T., Avolio, B. J., Kozlowski, S. W. J., Lord, R. G., Treviño, L. K., Dimotakis, N., & Peng, A. C. 2012. Embedding ethical leadership within and across organization levels. ***Academy of Management Journal***, 55: 1053–1078.

Scott, W. R., & Davis, G. F. 2007. ***Organizations and organizing: Rational, natural, and open systems perspectives***. Upper Saddle River, NJ: Prentice Hall.

Shaver, K. G. 2012. ***The attribution of blame: Causality, responsibility, and blameworthiness***. New York: Springer.

Shore, L. M., Barksdale, K., & Shore, T. H. 1995. Managerial perceptions of employee commitment to the organization. ***Academy of Management Journal***, 38: 1593–1615.

Shore, T. H., Bommer, W. H., & Shore, L. M. 2008. An integrative model of managerial perceptions of employee commitment: Antecedents and influences on employee treatment. ***Journal of Organizational Behavior***, 29: 635–655.

Simonsohn, U. 2018. Two lines: A valid alternative to the invalid testing of u-shaped relationships with quadratic regressions. ***Advances in Methods and Practices in Psychological Science***, 1: 538–555.

Slater, P. E. 1955. Role differentiation in small groups. ***American Sociological Review***, 20: 300–310.

Sperber, D., & Baumard, N. 2012. Moral reputation: An evolutionary and cognitive perspective. ***Mind & Language***, 27: 495–518.

Staw, B. M., & Hoang, H. 1995. Sunk costs in the NBA: Why draft order affects playing time and survival in professional basketball. ***Administrative Science Quarterly***, 40: 474–494.

Staw, B. M., & Szwajkowski, E. 1975. The scarcity-munificence of organizational environments and the commission of illegal acts. ***Administrative Science Quarterly***, 20: 345–354.

Stuart, H. C., & Moore, C. 2017. Shady characters: The implications of illicit organizational roles for resilient team performance. ***Academy of Management Journal***, 60: 1963–1985.

Tang, P. M., Yam, K. C., & Koopman, J. 2020. Feeling proud but guilty? Unpacking the paradoxical nature of unethical pro-organizational behavior. ***Organizational Behavior and Human Decision Processes***, 160: 68–86.

Tang, P. M., Yam, K. C., Koopman, J., & Ilies, R. 2022. Admired and disgusted? Third parties' paradoxical emotional reactions and behavioral consequences towards others' unethical pro-organizational behavior. ***Personnel Psychology***, 75: 33–67.

Treviño, L. K., & Brown, M. E. 2005. The role of leaders in influencing unethical behavior in the workplace. In R. E. Kidwell & C. L. Martin (Eds.), ***Managing***

organizational deviance: 69–87. Thousand Oaks, CA: Sage.

Tyler, T. R. 2006. Restorative justice and procedural justice: Dealing with rule breaking. ***Journal of Social Issues***, 62: 307–326.

Tyler, T. R., & Blader, S. L. 2005. Can businesses effectively regulate employee conduct? The antecedents of rule following in work settings. ***Academy of Management Journal***, 48: 1143–1158.

Tyler, T. R., & Boeckmann, R. J. 1997. Three strikes and you are out, but why? The psychology of public support for punishing rule breakers. ***Law & Society Review***, 31: 237–265.

Vadera, A. K., Pratt, M. G., & Mishra, P. 2013. Constructive deviance in organizations: Integrating and moving forward. ***Journal of Management***, 39: 1221–1276.

Van Dyne, L., & LePine, J. A. 1998. Helping and voice extra-role behaviors: Evidence of construct and predictive validity. ***Academy of Management Journal***, 41: 108–119.

Vardaman, J. M., Gondo, M. B., & Allen, D. G. 2014. Ethical climate and pro-social rule breaking in the workplace. ***Human Resource Management Review***, 24: 108–118.

Veiga, J. F., Golden, T. D., & Dechant, K. 2004. Why managers bend company rules. ***Academy of Management Executive***, 18: 84–90.

Wakeman, S. W., Moore, C., & Gino, F. 2019. A counterfeit competence: After threat, cheating boosts one's self-image. ***Journal of Experimental Social Psychology***, 82: 253–265.

Walker, P. 2012. September 14: UBS trader "risked the very existence of the bank", court told. ***Guardian.***

Webley, P., Cole, M., & Eidjar, O. P. 2001. The prediction of self-reported and hypothetical tax-evasion: Evidence from England, France and Norway. ***Journal of Economic Psychology***, 22: 141–155.

Werner, J. M. 1994. Dimensions that make a difference: Examining the impact of in-role and extrarole behaviors on supervisory ratings. ***Journal of Applied Psychology***, 79: 98–107.

Wiltermuth, S. S., & Flynn, F. J. 2013. Power, moral clarity, and punishment in the workplace. ***Academy of Management Journal***, 56: 1002–1023.

Winship, V., & Robbennolt, J. K. 2018. An empirical study of admissions in SEC settlements. ***Arizona Law Review***, 60: 1–66.

Wyshynski, G. 2021. November 1: The first decade of the NHL's Rule 48. How it completely changed the league's trajectory. ESPN.

Zhan, X., & Liu, Y. 2022. Impact of employee pro-organizational unethical behavior on performance evaluation rated by supervisor: A moderated mediation model of supervisor bottom-line mentality. ***Chinese Management Studies***, 16: 102–118.

Zhang, L. 2017. A fair game? Racial bias and repeated interaction between NBA coaches and players. ***Administrative Science Quarterly***, 62: 603–625.

Zhang, L. 2019. Who loses when a team wins? Better performance increases racial bias. ***Organization Science***, 30: 40–50.

Zhang, X., Ma, C., Chen, Z.-X., Cheng, D. C-M., Tian, Y., & Tian, G. 2019. Leaders' disciplinary reactions to followers' unethical pro-organizational behavior. Abstract. ***Academy of Management Proceedings***, 2019: 17357.

Zipay, K. P., Mitchell, M. S., Baer, M. D., Sessions, H., & Bies, R. J. 2021. Lenient reactions to misconduct: Examining the self-conscious process of being lenient to others at work. ***Academy of Management Journal***, 64: 351–377.

S. Wiley Wakeman (wiley.wakeman@hhs.se) is an assistant professor in the department of management and organization at Stockholm School of Economics. His research explores ethics, inequality, and sustainability, particularly looking at how forms of deviance (innovation) might address individual and social goals.

Philip Yang (philip.yang@uni-paderborn.de) holds the Professorship in Sustainable Human Resource Management and Leadership at the University of Paderborn. His research examines sustainable human resource management and leadership, with an emphasis on diversity, ethics, new work forms, green HRM, and innovation.

Celia Moore (c.moore@imperial.ac.uk) is professor of organizational behavior at Imperial Business School, where she also directs the Center for Responsible Leadership. Her research focuses on how individuals enact their moral agency and navigate morally consequential decisions in their professional lives.

APPENDIX A
EXPERIMENTAL GAME

Participants were instructed they would be playing a team game that included four "players" and an "advisor" [leader]. All participants were assigned the advisor role. The game consisted of four rounds: three regular rounds and one bonus round.

In each of the three normal rounds, players would roll a die five times and report their rolls, or scores, which would be aggregated to provide a team score. In the bonus round, the advisor would

select one player who would complete an additional round of five rolls. Their score would be doubled and added to the team score.

Players' rolls were reported to the advisor after each round (see Table A1–A3 for each round). We asked advisors to report their happiness and worry about each player's performance individually after each round, to select one of the four players as the most valuable in that round, and to select one player for the bonus round.

Though the performance of all players was manipulated in the team game, the performance of Player 2 represented our manipulation of the degree and severity of rule breaking by not reporting scores in a statistically probable fashion, as was noted in the rules of the game.

TABLE A1
Player Rolls: Round 1

Player number	Roll 1	Roll 2	Roll 3	Roll 4	Roll 5
Player 1	1	3	5	5	2
Player 2	By cond.	By cond.	By cond.	By cond.	By cond.
Player 3	6	6	5	5	3
Player 4	3	3	4	2	5

Note: By cond. = by condition.

TABLE A2
Player Rolls: Round 2

Player number	Roll 6	Roll 7	Roll 8	Roll 9	Roll 10
Player 1	4	3	2	4	4
Player 2	By cond.	By cond.	By cond.	By cond.	By cond.
Player 3	2	6	4	6	1
Player 4	6	2	3	3	5

Note: By cond. = by condition.

TABLE A3
Player Rolls: Round 3

Player number	Roll 11	Roll 12	Roll 13	Roll 14	Roll 15
Player 1	3	4	1	2	4
Player 2	By cond.	By cond.	By cond.	By cond.	By cond.
Player 3	4	5	2	3	6
Player 4	4	4	2	3	5

Note: By cond. = by condition.

MANIPULATION: PLAYER 2 ROLLS IN EACH CONDITION (BY ROUND)

Avg. 3.5 (Rule-Abiding)

Round 1: 1, 3, 4, 6, 3
Round 2: 3, 1, 5, 3, 6
Round 3: 1, 4, 5, 6, 2

Avg. 5 (Moderate)

Round 1: 6, 6, 5, 4, 4
Round 2: 4, 6, 5, 5, 5
Round 3: 4, 6, 6, 4, 5

All 5 (Extreme)

Round 1: 5, 5, 5, 5, 5
Round 2: 5, 5, 5, 5, 5
Round 3: 5, 5, 5, 5, 5