

# ATTRIBUTION MATTERS

## How Causal Explanations Influence Perceptions of Dangerousness and Racial Classification

DVIR YOGEV   
UC Berkeley

---

---

Understanding the factors influencing perceptions of dangerousness and their relationship with racial classification is crucial within the criminal legal system. These subjective perceptions affect discretionary parole, bail, and plea-bargaining decisions, shaping the lives of individuals and the safety of society. Existing psychological research highlights the Fundamental Attribution Error's role in shaping perceptions of morality, primarily by attributing others' actions to their intrinsic personality traits or character. However, its influence on perceptions of dangerousness remains less explored. Through two online experiments ( $N = 1005$  and  $N = 276$ ), this study investigates how attributing crime to internal or external factors shapes perceptions of dangerousness and racial classification. The results reveal that external attribution increases perceptions of dangerousness and predisposes individuals to classify a wrongdoer as a person of color. These findings have immediate and significant implications for the criminal legal system, suggesting urgent areas for policy reform.

**Keywords:** attribution theory; parole; perceived dangerousness; racial classification

---

---

### INTRODUCTION

Assessing a person's potential danger is not just a routine task but a crucial skill that permeates every aspect of the criminal legal system (Kraemer et al., 1997; Skeem & Monahan, 2011). From bail hearings to parole decisions, the criminal legal system is consistently making judgments to protect society, often keeping those deemed "dangerous" under close scrutiny (Simon, 2005). Despite the growing use of algorithms to assist in risk assessments (Yogev & Mehozay, 2022), the process remains inherently subjective, rooted in human perception (Skeem et al., 2020; Tonry, 2019).

During legal hearings such as parole, bail, or sentencing, individuals are evaluated based on explicit and implicit factors. Explicit factors include age, gender, and criminal record (Young et al., 2016), while implicit factors encompass the individual's behavior, statements during the assessment process (Shammas, 2019; Young & Chimowitz, 2022), the ideologies

---

**AUTHOR'S NOTE:** *This study was registered with OSF Registries. I have no conflict of interest to disclose. Correspondence concerning this article should be addressed to Dvir Yogev, Criminal Law and Justice Center, UC Berkeley, Berkeley, CA, 94720; e-mail: dyo@berkeley.edu.*

CRIMINAL JUSTICE AND BEHAVIOR, 202X, Vol. XX, No. X, Month 2024, 1–21.

DOI: 10.1177/00938548241307234

Article reuse guidelines: [sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)



© 2024 International Association for Correctional and Forensic Psychology

of the assessors (Werth, 2019), and prevailing socio-political norms (Aviram, 2020). Notably, in the American criminal legal context, race can influence perceptions of dangerousness (Harcourt, 2015; Monahan, 2006).

This article uses online experiments to build on existing qualitative and ethnographic insights to empirically examine the effect of causal attribution in a statement about past crime on perceptions of risk and racial classification. Specifically, it explores the implications of different causal attributions—emphasizing either external (environmental pressures, situations, or stressors) or internal factors (disposition, traits, and intentions)—on suitability for prison release and perceptions of future dangerousness.

How does prioritizing one causal attribution over the other affect perceptions of future dangerousness? Drawing from the existing research on discretionary parole hearings (Aviram, 2020; Shammass, 2019; Werth, 2019; Yelderman et al., 2021; Young & Chimowitz, 2022; Young & Pearlman, 2021), this article theorizes that attributing past wrongdoing to societal factors—an external attribution—may heighten perceptions of dangerousness. This is of significant normative concern, considering the well-documented influences of systemic social disadvantages on patterns of wrongdoing. Decades of sociological and criminological research have highlighted the myriad external causes of crime, particularly within the contemporary U.S., where institutional racism, poverty, restricted access to education, and patterns of abuse are key drivers of criminal behavior. Yet, the American criminal justice system upholds free will as the central tenet of criminal responsibility (e.g., Lacey et al., 2021; Wacquant, 2009).

In addition, this article contributes to the literature on the intersection of criminal justice contact and racial perceptions in the United States. Prior research has demonstrated that incarceration significantly skews racial perceptions, often leading to a higher likelihood of people being classified as Black and a lower likelihood of being classified as White (Penner & Saperstein, 2008; Saperstein & Penner, 2010; Saperstein et al., 2014). Thus, this study also investigates whether judgments of dangerousness are linked with racial perceptions: Does perceiving a person as dangerous coincide with perceiving them as a member of a racial minority?

Perceptions of dangerousness are particularly relevant to criminal justice decision-making. This article argues that the cognitive tendency to explain the actions of others through internal factors (choice, free will, and effort, for instance) and our own actions through external factors (such as situational and environmental aspects) affects perceptions of dangerousness. Discretionary parole is utilized to illustrate this argument, as it is a compelling case of how perceptions of dangerousness can be influenced through the subjective interpretation of a parole candidate's causal attribution style.

## CAUSAL ATTRIBUTION THEORY

When constructing the meaning of an event, people differentiate between internal and external causes. While internal attribution ascribes causes to individual ability, effort, choice, and will, external attribution is contextual. External attribution emphasizes the role of luck, situational aspects, and environmental differences as the driving causes of realized events (Weiner, 1985; Weiner & Graham, 1990). One aspect of Causal Attribution Theory is the Fundamental Attribution Error (FAE)—a tendency to prefer dispositional causal attribution for the actions of others, which is not grounded in factual reality but intuition

(Gilbert & Malone, 1995; Ross & Nisbett, 2011). Research shows that people prefer to explain their actions using external terms and the actions of others through internal causal terms (Ross, 1977). Accordingly, in the context of criminal justice, it has been argued that decision-makers like judges, juries, and prosecutors “are likely to overestimate the causal significance of personal choice, and to correspondingly underestimate the causal significance of situational factors in the behavior of others” (Dripps, 2003; Headworth, 2021).

Moral judgments often make people think in causal terms (Heider, 1958; Weiner, 2006). In criminal justice contexts, consider determining appropriate punishments or assessing an individual’s readiness for release from prison. Attribution theories suggest that when making moral judgments, people consider both dispositional factors (internal to the individual, e.g., free will) and situational factors (external influences, e.g., discrimination; Tetlock et al., 2007).

This dichotomy is reflected in criminological and psychological theories. Dispositional theories (also known as “rational” or “neo-classical” theories) view personal choice as the primary cause of crime. In contrast, situational theories (also called “positivist” theories) characterize crime as the result of social forces (Cullen et al., 1985; Cushman, 2008; Garland, 2001).

Existing literature, although sparse, generally shows that people assign less blame when external factors cause crime. In 1981, a survey on perceptions of the causes and the appropriate prescribed punishment for crimes suggested that the public generally supported less severe punishments for crimes caused by external factors compared with those caused by internal factors (Hawkins, 1981). Another survey from 1986 showed that presenting a crime caused by internal factors caused participants to judge the person as more dangerous and less “treatable” compared with situations with external causes of crime (Quinsey & Cyr, 1986). In 2008, a study by psychologist Fiery Cushman reached similar conclusions, arguing that when people ascribe blame and judge deserved punishment, they rely heavily on their analysis of causal responsibility (Cushman, 2008).

The FAE predicts that people will overestimate the effect of personality and character on others’ actions (internal attribution) and underestimate the influence of situational context (external attribution; Harman, 1999; Perez & Salter, 2019; Pickett & Baker, 2014; Tetlock, 1985). More specifically, research showed that conservatives are expected to favor internal causal attribution to crime and disapprove of external attribution more often than their liberal counterparts (Clarkson et al., 2015). However, Americans generally tend to endorse an individualistic culture that corresponds with a dispositional viewpoint toward the actions of others (Morris & Peng, 1994; Triandis, 2001). Research shows that the FAE also impacts professional decision-makers. For example, a study of probation officers found evidence of interaction between the FAE and racial bias, as probation officers described Black youths with negative personality traits and White youths using negative environmental influences (Bridges & Steen, 1998).

### **Attribution Congruence**

This article introduces a novel theory of attribution congruence. It is an extension of the FAE that applies the FAE’s prediction to the setting of judgments on dangerousness. The Attribution Congruence theory suggested here posits two main points. First, observers are inclined to attribute the misconduct of others to internal factors (as predicted by the FAE,

which would also predict a person who committed an offense's attribution of their own misconduct to external factors). Second, and a main contribution of this article, this inclination leads observers to perceive wrongdoers as *less* dangerous when those wrongdoers attribute their actions to internal causes. In sum, when both the observer and the wrongdoer attribute the crime to internal causes, the perceived dangerousness would be lower. Conversely, when the wrongdoer's explanation diverges from the observer's internal attribution, typically by citing external factors (as the FAE would predict), the wrongdoer is perceived as more dangerous.

Attribution Congruence as a theory is agnostic regarding the kind of causal attribution (internal or external). As long as the observer and wrongdoer match in terms of their preferred causal attribution, then the perceived dangerousness would be lower. However, the FAE remains a core aspect of any addition to the Causal Attribution Theory. Thus, people will more often attribute the crimes of others to internal aspects and their own to external ones. Accordingly, if both the observer and the wrongdoer follow the FAE's predictions, then the perceived dangerousness would likely be higher than if one of them diverges from the FAE's prediction.

#### ATTRIBUTION IN THE CRIMINAL LEGAL SYSTEM

##### The Case of Discretionary Parole

Professionals of various backgrounds and occupations make decisions based on how much they perceive others as a danger to society, from 911 operators to judges. A discretionary parole hearing is a notable procedure to be affected by the parole candidate's attribution style. In a discretionary parole hearing, a parole candidate would address the causes of their past crime and might use either external or internal causal attribution. The board, in return, would have to make a subjective judgment on the candidate's suitability for release based on perceived dangerousness. Parole board hearings vary, yet they all focus primarily on the parole candidates' *expected* behavior should they be released.

Discretionary parole's primary purpose is preventive, as is practiced in most US states. After fulfilling the "retributive" minimum, a person is released when some rehabilitation conditions of suitability are fulfilled. While the legal aspect of the decision-making process seeks to provide objective rules, the psychological aspect, which governs the essence of the discretionary process, is inherently subjective. Young and Chimowitz (2022) find that parole commissioners deal with "the impossible task of knowing another person's heart or mind" (p. 237). In California, for example, the decision reflects the decision-maker's interpretation of the nebulous concept of "insight" (Aviram, 2020), which, as ethnographer Shammas (2019) noted, relates directly to the parole candidate's ability to voice internal attribution. Shammas (2019) finds that parole candidates from disadvantaged backgrounds experience "a failure to speak in accordance with legitimate language . . . which necessarily results in a situation of performative disadvantage" (p. 152).

This article's main contribution is to our knowledge of perceptions of dangerousness, yet it also contributes to understanding parole decision-making despite relying on laypeople and not professional parole board members. Psychology and law scholars commonly pursue laypeople to study parole and other legal decision-making processes (ElBassiouny et al., 2022; Engel & Grgić-Hlača, 2021; Yelderman et al., 2021). Parole relies on some resemblance of parole boards to laypeople—most parole boards do not require formal

training and expertise, even though board members usually have a criminal justice-related background (Rhine et al., 2017). In addition, public opinion often influences parole boards; hence, learning public perceptions of dangerousness is compelling (Watts & Rhine, 2018).

Moreover, discretionary parole is subjective by definition and follows almost no strict objective guidelines for who should be released (Ruhland, 2020). Research on parole decision-making shows that despite adopting guidelines and actuarial tools in modern parole, decisions are rife with arbitrariness and bias (Huebner & Bynum, 2008; Medwed, 2008; Vîlcică, 2018; Young & Pearlman, 2021). For example, race impacts a person's parole hearing, as discussed next. Thus, understanding how people construct perceptions of dangerousness has implications for understanding how parole boards do the same.

### **Dangerousness Perceptions and Racial Classification**

This article contributes additionally to our understanding of how contact with the criminal justice system shapes racial perceptions, which is imperative in an era of mass imprisonment (Saperstein & Penner, 2010; Saperstein et al., 2014). Gilliam and Iyengar show that the public is socialized to associate violent crime with non-Whiteness, and more specifically, that “a brief five-second exposure to a Black perpetrator in the news is sufficient to increase the percentage of people who believe crime is caused by individual failings” (Gilliam & Iyengar, 2000, p. 571). This article expands on their research—how does using internal or external attribution affect the belief that the person who perpetrated violence is non-White?

Scholars have been increasingly finding evidence that race is fluid, and depends on social context (Hamilton, 1988). Penner and Saperstein showed that racial perceptions change over time due to social circumstances (Penner & Saperstein, 2008). Crucially, contact with the criminal legal system increases the likelihood of being perceived by others (and identifying) as a person of color by interviewers (Saperstein et al., 2014). Others found that racial stereotypes regarding behavior can trigger shifts in visual perceptions and decision-making, even for police officers (Eberhardt et al., 2004).

This is in accordance with research showing that Americans hold stereotypes about race and crime (Gilliam & Iyengar, 2000). Previous research has shown that racial resentment predicts White Americans' support of punitive policies (Green et al., 2006), in part because they assume racial minorities would be the ones targeted (Barkan & Cohn, 2005; Brown & Socia, 2017; Hurwitz & Peffley, 2005; Peffley & Hurwitz, 2002; Unnever & Cullen, 2010). Moreover, previous findings regarding the attribution of racial disparities in the criminal legal system show that Black people explain Black people's overrepresentation in prison through racial discrimination (Cooper et al., 2021).

Understanding how contact with the criminal justice system shapes racial perceptions is crucial to addressing racial disparities in life outcomes among Americans, particularly in an era of mass imprisonment. Criminal justice contact has increased racial disparities, exacerbating racial inequality in U.S. jails, courts, and prisons (Western, 2006). Racial minorities spend longer times in prison awaiting parole compared with white people who committed an offense (Huebner & Bynum, 2008) and are less likely to be granted parole (Young & Pearlman, 2021). Therefore, it is essential to recognize that race is not just an input into the criminal justice system but is shaped by it (Walker, 2016). Thus, this article contributes by addressing how racial perceptions interact with perceptions of dangerousness.

### The Present Studies

The present studies are modeled after California's system of discretionary parole because of the state's relative importance to the study of parole (Young, 2020; Young & Chimowitz, 2022; Young & Pearlman, 2021), and the settings of its discretionary parole board hearings which are conveniently adapted to a survey experiment. California holds the largest number of parole-eligible people who are incarcerated ("Lifers," sentenced for murder, mostly) and utilizes a system of hearings to decide eligibility. In terms of scope, California's discretionary parole system is one of the most significant in the country. It controls the fate of over 30,000 people (California Department of Corrections and Rehabilitation, 2020).

Moreover, attribution theory holds significance for understanding California's discretionary parole system because of its unique jurisprudence of "Insight," which posits voicing internal attribution as a release standard into the legal doctrine. The Supreme Court of California guides the decision-makers to *deny* parole to an incarcerated person that "lacks insight," that is, makes a causal connection between his "life crime" and *external* factors: childhood trauma, the influence of friends, economic hardship, or the effects of structural racism and social segregation (*In re Shaputis*, 265 P.3d 253, 275 (Cal. 2011), at 216). The court made "insight" the decisive factor for release on parole and tied it to the parole candidate's explanation of their crime (*In re Lawrence*, 190 P.3d 535 (Cal. 2008); *In re Shaputis*, 190 P.3d 573 (Cal. 2008); *In re Shaputis*, 265 P.3d 253, 275 (Cal. 2011) ("Shaputis II")). Justice Liu wrote in his Shaputis II concurring opinion that a lack of insight might suggest an insufficient understanding of the causes that led them to commit the crime (Shaputis II, 53 Cal. 4th 192, 218 (2011) at 275-76.). Thus, understanding the psychology of causal attribution theory concerning perceptions of dangerousness would contribute to understanding the legal doctrine of "insight."

## STUDY 1

Study 1 was pre-registered; all materials are available here: <https://shorturl.at/Zmx8w>. The Office for Protection of Human Subjects (OPHS) at UC Berkeley approved the methodology for this experiment (CPHS PROTOCOL NUMBER: 2021-01-13929).

### PARTICIPANTS

One thousand and five voting-age Americans recruited through the MTurk platform completed the survey. Fielded in February 2021 and restricted to American adults (Supplemental Table S3, available in the online version of this article). Participants were paid one dollar for an estimated assignment time of seven minutes. Participants reported a mean age of 40 ( $SD = 12.4$ ); 41.8% identified as Female, 57.9% as Male, and 0.3% in another way. 69.6% identified as White, 10.8% as Black, 9.3% as Asian, 5.4% as Hispanic, 3.1% as Mixed, and 1.7% as Native American.

MTurk is an online labor market used extensively in social science research to recruit samples quickly and inexpensively (Berinsky et al., 2012). This work's compensation was set at approximately \$9 per hour. In addition, research suggests that MTurk participants are not receptive to experimenter demand effects (Mummolo & Peterson, 2019) and are more attentive and cooperative than participants from other online samples (Boas et al., 2020), making MTurk a proper subject pool for this specific judgment task.

## MATERIALS

Before choosing to participate, each participant read: “This is a judgment and decision-making study that examines public opinion about justice.” In the consent form, participants were told, “All the information presented in this research study is based on real-world people but is not the information of any single real person.” Each participant was told: “[Y]ou will be asked to decide whether a particular incarcerated person should be granted a release from prison on parole (supervision in the community). This is what people who sit on parole boards do.” The following part consisted of instructions on how to make the decision, emphasizing the guiding principle of California’s discretionary parole: “The only condition for release on parole is: the prisoner will not pose a risk of danger to society if released from prison.” (See all the instructions, attention checks, and comprehension checks in the Supplemental Material, available in the online version of this article).

To account for other factors that can influence perceptions of dangerousness, all participants read the same description of a single parole candidate based on a profile of the “median lifer.” As is the case in most hearings in California, the putative lifer was convicted of murder. Murder is the most egregious of offenses, allowing the study to make assumptions under the lowest a priori likelihood of perceived suitability (Hritz, 2021). The presented profile had random variation in the person who committed an offense’s age at the time of the offense (17-30), time in prison at the time of the hearing (25-40), original sentence minimum length (15/25), and whether this is the person who committed an offense’s first or second conviction. This prevents the possibility of a direct influence of one of these variables on the participants’ decisions, and they show no effect on the results.

To eliminate variance in the effect of perceived remorse and accountability, all participants read the same response to the question: “When did you start feeling remorse?” Regardless of treatment condition, all respondents read: “I’m very sorry about my actions and take full responsibility.” The response is reported in full in the Supplemental Material.

### Independent Variable

Next, participants were randomly assigned to one of three groups: control, internal attribution, or external attribution. In all three treatment conditions, the parole candidate answers the question: “explain your violent crime.” Participants in the external attribution condition read explanations focused on economic inequality, excessive policing and marginalization harms, systematic discrimination, and lack of opportunities. Participants in the internal attribution treatment group read explanations focused on making bad decisions and knowingly choosing to harm while being in control (see Supplemental Material A, available in the online version of this article for the full text). Finally, a control group read no explanation for the crime; instead, they read the following text before being asked to decide: “When asked whether he needs disability-related accommodations for participating in the parole hearing he responds: ‘No, I’m healthy.’”

The conditions manipulate the attribution while keeping the information as equivalent as possible between treatment groups (Busby et al., 2019; Esberg & Mummolo, 2018; Hetey & Eberhardt, 2018). As a placebo test, participants were also randomized into one of four treatment realizations to ensure that no idiosyncrasies of a specific vignette drive the effect of the experiment. Each realization is another version of the treatment condition, such that there are four different versions of each condition to ensure that any effects are not the result

of peculiar or distinctive wording. Multiple realizations prevent construct validity issues by assuring that the construct of interest is independent of any realization of the treatment vignettes. No statistical differences in effect were found between the four treatment realizations (See Supplemental Material C, available in the online version of this article).

### Dependent Variables

Following the instructional and informational sections, participants answered the first dependent variable: to release or to deny parole. Afterward, the second dependent variable—dangerousness—was measured by requiring a rating on a scale of 1 (“*Don’t agree at all*”) to 5 (“*Agree completely*”) the following statements: This person is likely to reoffend; This person is dangerous; This person should stay in prison. The statements regarding reoffending, dangerousness, and staying in prison are used as secondary dependent variables and combined on a scale of perceived dangerousness using sum scores (McNeish & Wolf, 2020), verifying that the decision to release is indeed correlated with perceptions about dangerousness. Reports on the effect of treatment conditions on the respondent’s confidence are also reported.

The third dependent variable is racial classification. It is measured using the following question: “What was the person’s race? Don’t worry about it if you aren’t sure.” The respondents can choose between Black, White, Asian, Hispanic, and Don’t know. In the analysis, Black, Asian, and Hispanic were coded as People of Color (“PoC”). This study relies on the concept of Information Equivalence (Dafoe et al., 2018). Some aspects of the treatment might trigger racial attitudes, accounting for the estimated effect. To verify whether the independent variable manipulates people’s perception of causal attribution independently of perceptions of racial identity, participants are asked to report the hypothetical person’s race, even though it is not disclosed in the information they receive.

### ANALYTICAL STRATEGY

This study is a between-subjects design. It first estimates the relationship between the treatment condition groups and the outcomes using the Chi-square test of independence. Second, it estimates the following Analysis of Variance model (using the *aov* function from the Stats package in R) by performing linear regression for each stratum (Chambers et al., 2017), where  $X_i$  is a vector of pre-registered, pre-treatment covariates included to improve the precision of treatment effect estimates. Results are reported both with and without the covariate controls:

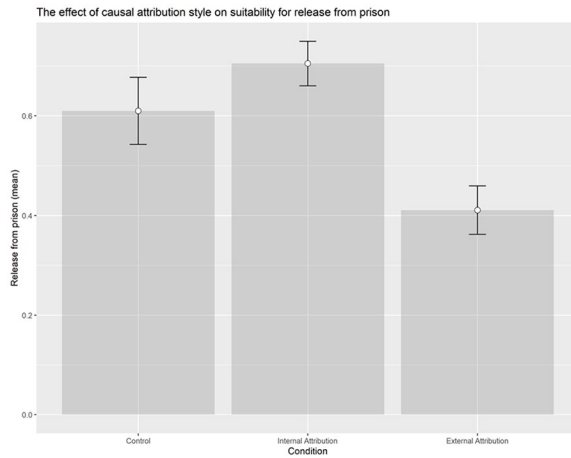
$$DV_i = \beta_0 + \beta_1 Attribution_i + \beta_2 Racialization + \beta' X_i + \varepsilon_i$$

The differences between the groups are tested using Tukey’s Honestly Significant Difference (Tukey’s HSD) post hoc test for pairwise comparisons.

### RESULTS

Study 1 shows that participants reward internal causal attribution and punish the use of external causal attribution. When a putative parole candidate assigns the cause of their behavior to their characteristics and free will, they are perceived as less dangerous; when assigning the cause to outside forces, the candidate is perceived as more dangerous.





**Figure 1 Cell Means for Supporting Release in Each Condition, With Error Bars Based on Cell Standard Errors**

*Note.* Figure 1 depicts the mean for the decisions to release from prison (1, deny = 0) across the various treatment conditions. Each bar represents the mean value, the point markers are superimposed on the bars to indicate the means, and error bars extend from each point to represent the 95% confidence intervals (CI) around those means.

Figure 1 shows that, in the control group, the parole candidate release rate was 0.609, in the internal treatment, the release rate was 0.704, and in the external treatment, it was 0.410. The Chi-square test of independence shows that the treatment significantly affected the participant's decision to release and the assessment of dangerousness,  $\chi^2(2, N = 1,005) = 72.29, p < .0001$ . An analysis of variance (ANOVA) test shows that the difference in means for a decision to release is statistically significant,  $F(2, 1,002) = 38.83, p < .001, \hat{\eta}_G^2 = .072, 90\% \text{ CI } [.048, .098]$ . Results remained significant when controlling for reported gender, party affiliation, political ideology, and reported race. The Eta Squared effect size is medium  $\eta^2 = 0.07, 95\% \text{ CI } [.05, 1]$ .

#### Situational Punishment and Dispositional Reward

According to Table 1, respondents who read an external causal explanation were 30 percentage points less likely to release the parole candidate compared with respondents who read the internal attribution explanation (the Cohen's D effect size is relatively large,  $d = 0.61, 95\% \text{ CI } [.47, .75]$ ).

#### Perceptions of Dangerousness

Participants perceived the likelihood of reoffending, dangerousness, and whether the person should stay in prison were higher in the external treatment. When combined into a "dangerousness scale" by averaging the three items, the relationship between the scale and the treatment condition is significant,  $\chi^2(24, N = 1005) = 99.59, p < .0001$ , and shows that the causal attribution used by the candidate affects how much they are perceived as dangerous (Figure 2). Figure 2 shows that the median score on a "dangerousness scale" was 0.75 in the external condition and 0.33 in the internal condition.

**TABLE 1 Post Hoc Tests—Differences Between Conditions on the Decision to Release**

$\Delta M$	$t$	$df$	$P_{adj}$	Comparison
0.09	2.32	1002	.053	Internal—Control
-0.20	-4.85	1002	<.001	External—Control
-0.29	-8.71	1002	<.001	External—Internal

Note. Multiple testing correction with a Tukey test.

Supplemental Table S5 (available in the online version of this article) shows that respondents in the situational condition perceived the parole candidate as more dangerous than in the external condition (the effect size is medium,  $d = 0.55$ , 95% CI [.41, .69]).

### Text Analysis

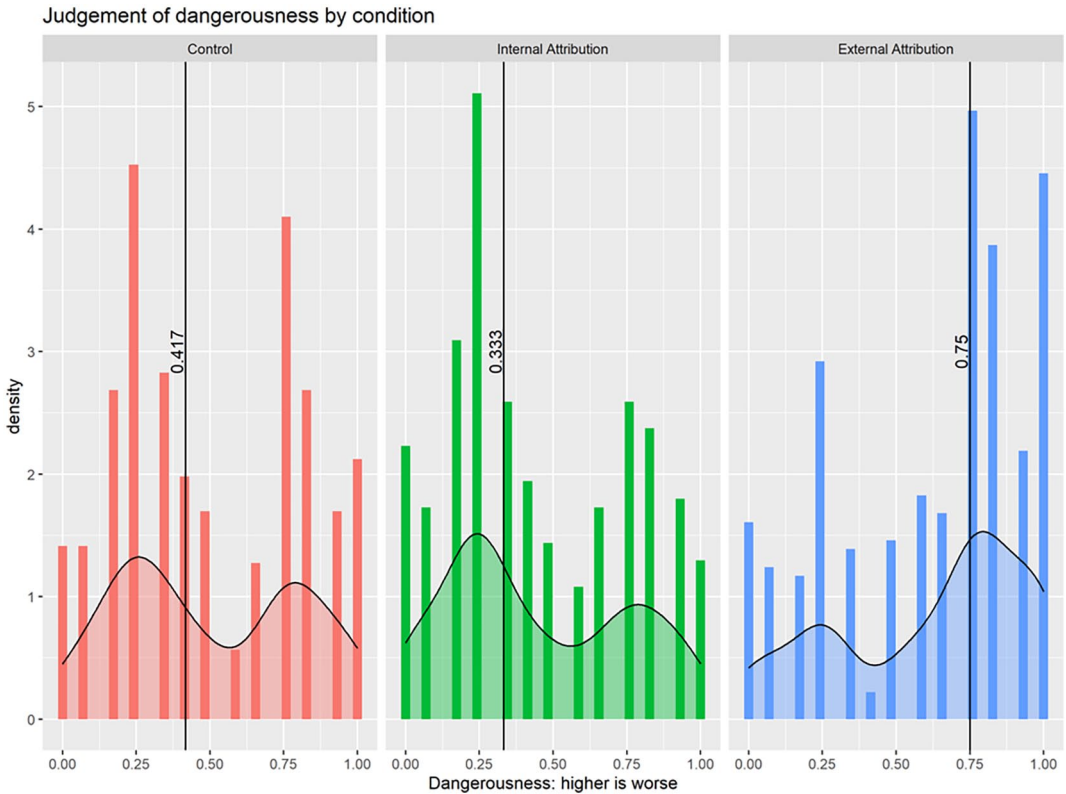
After indicating their response regarding release from prison and dangerousness, respondents were asked: “If you can, please write why you decided to release or not to release the person from prison.” I employed keyness statistics (See Supplemental Material for more information) to compare word frequency differences between respondents who opposed and supported the release of the putative incarcerated person (Zollinger, 2022). This method allows a comparison of the difference between justifying one’s decision within an experimental condition and across conditions. Figure 3 shows that respondents receiving the Internal treatment overwhelmingly justified their decision not to release based on “risk” compared to decisions to release.

Respondents in the External treatment (Figure 4) overwhelmingly relied on the concepts of responsibility, action, remorse, and blame when deciding not to release compared to those deciding to release.

This analysis affirms that people’s decisions were indeed influenced by their experimental treatment. Emphasizing free will and character can result in a person perceived as inherently “risky” (Figure 3). Conversely, describing the circumstances as the source of the criminal behavior is perceived as shifting blame, no remorse, and a lack of responsibility for one’s actions (Figure 4).

### Racial Classification

A sizable portion assumed the parole candidate was a person of color. This is demonstrated in the top part of Figure 5 (20% in control; 27.2% in the Internal treatment; 29.9% in the External treatment). Overall, 248 participants perceived the putative person who committed an offense as Black, only five as Asian, and 18 as Hispanic. The difference between the three groups is statistically significant,  $\chi^2(2, N = 1005) = 6.2156, p = .044$ . The treatments also significantly affected the racial classification of the parole candidate as White, as demonstrated in the bottom plot of Figure 5. While racial classification as a minority increased from the control to the internal condition and once again to the external condition, racial classification as White demonstrated the opposite trend (16.5% in the control condition, 12.1% in the internal condition, and 8.8% in the external condition,  $\chi^2(2, N = 1005) = 7.9878, p = .018$ ).



**Figure 2** Distribution of Dangerousness Index, Higher Means More Dangerous

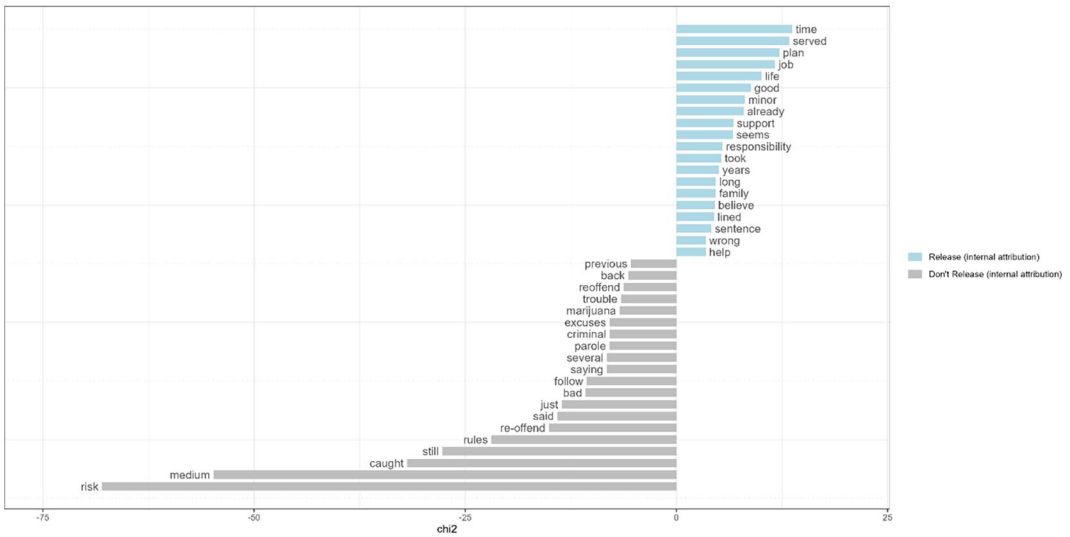
*Note.* Vertical lines are the medians. This figure illustrates the distribution of perceptions of dangerousness under three distinct treatment conditions. The x-axis represents an index of Dangerousness, where higher values indicate worse perceptions. This index is computed by averaging three items: the likelihood of the person reoffending, their perceived dangerousness, and whether they should stay in prison. Each facet represents a treatment condition, and within each, the histogram and overlaid density plot show the distribution of dangerousness scores. The vertical line in each facet represents the median score for that group, and the numerical value of the median is displayed alongside.

## STUDY 2

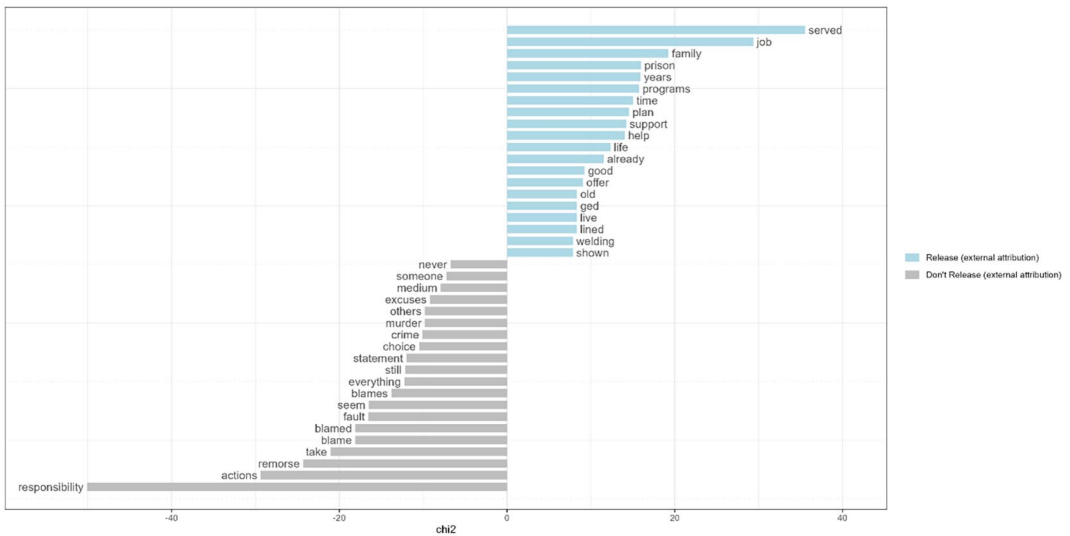
Study 2 was conducted after Study 1 to explore a hypothesis as to why the treatment had the effect it did. Study 2 kept the materials and procedures of Study 1 and added a question asking participants to rate the extent to which the crime committed by the putative person who committed an offense depicted in the vignette resulted from his disposition, traits, and intentions versus environmental pressures, situations, or stressors. The purpose is to understand the interaction between the respondent's valuation of the putative person who committed an offense's described causal attribution and the judgment of dangerousness.

### PARTICIPANTS

Two-hundred and seventy-six American adults were recruited through the Lucid Theorem. Fielded in March 2022 and targeted U.S. nationally representative respondents (Supplemental Table S4, available in the online version of this article). The researcher pays

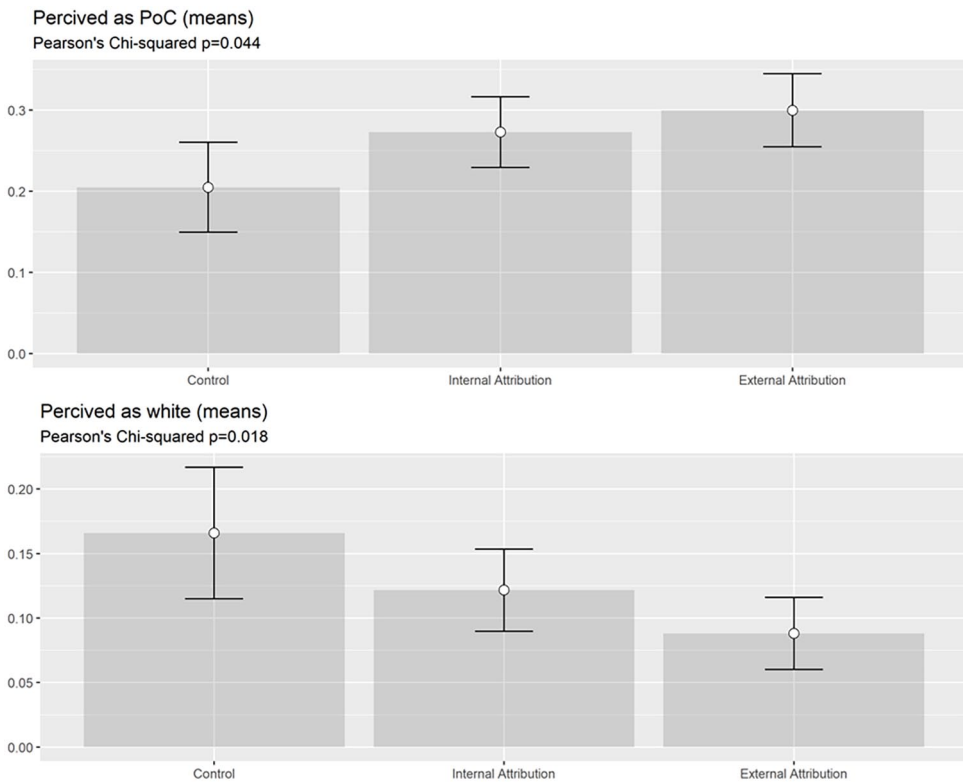


**Figure 3 Keyness Statistics by Decision to Release or Deny Release—Internal Treatment**  
*Note.* Shows terms mentioned with greatest relative frequency by respondents who decided to release a parole candidate, relative to respondents who opposed release.



**Figure 4 Keyness Statistics by Decision to Release or Deny Release—External Treatment**  
*Note.* Shows terms mentioned with greatest relative frequency by respondents who decided to release a parole candidate, relative to respondents who opposed release.

each respondent a fixed \$1 price, a tariff Lucid set. Lucid Theorem employs quota sampling to produce samples matched to the US population on age, gender, ethnicity, and geographic region; recent research demonstrates the suitability of the Lucid platform for evaluating social scientific theories (Coppock & McClellan, 2019). Using two different services



**Figure 5 The Effect of Treatment Condition on Perceptions of the Person's Race**

Note. Figure 5 displays the mean perception of a person's race as a person of color (PoC) or as White across different treatment groups. The y-axis represents the mean within each group for perceiving a person as a PoC or White. A point marker signifies the mean, along with 95% CI error bars. Pearson's Chi-square test indicates that the differences among the groups are statistically significant.

(MTurk and Lucid) increases the robustness of the studies to systematic differences between services.

Respondents reported a mean age of 45.8 ( $SD = 16.5$ ). 51.8% reported as Female gender and 47.5% as Male. 71.8% of the respondents reported they identify as White, 10.5% as Black, 9.1% as Hispanic, 6.5% as Asian, and 2.2% as Mixed and Other.

## MATERIALS

Study 2 adds the following question: "Thinking about the prisoner in question, would you say his crime was the result of. . .?" The options were: "His character and personality," "His intentions and plans," "Poverty, bad education, and lack of opportunities," "Peer pressure and a lack of a stable environment at home." The answers are presented in a matrix format using the following scale: Definitely, Somewhat, No at all, and Don't know. The analysis combines "character and personality" and "intentions and plans" to measure attribution to internal factors. Similarly, to measure attribution to external factors, the analysis combines "Poverty, bad education and lack of opportunities" with "peer pressure and lack of stable environment at home." (See Supplemental Material for a correlation table).

**TABLE 2 Mean “Release” decisions by Respondent’s Attribution Belief**

Respondent’s attribution belief	Model 1: External condition	Model 2: Control condition	Model 3: Internal condition
Respondents believe an internal attribution	<b>-0.431</b> [-0.779, -0.082] <i>p</i> = .016	<b>-0.600</b> [-1.178, -0.022] <i>p</i> = .042	-0.109 [-0.457, 0.239] <i>p</i> = .536
Respondents believe an external attribution	0.223 [-0.196, 0.642] <i>p</i> = .292	-0.242 [-0.873, 0.390] <i>p</i> = .440	<b>0.437</b> [0.099, 0.776] <i>p</i> = .012
Num. Obs.	99	45	99
<i>R</i> <sup>2</sup>	.200	.432	0.271

*Note.* Table 2 presents the results of linear regression models that predict respondents’ “Release” decisions (1 = Release, 0 = Deny) under three treatment conditions. Predictors include respondents’ beliefs in internal and external attributions (between 0 = none, and 1 = completely) and demographic control variables. The coefficients represent the average change in the dependent variable for a one-unit increase in the predictor. Confidence intervals are presented in square brackets, and *p*-values are reported below the coefficients. All models use robust standard errors. *R*-squared values are also presented to indicate the proportion of variance explained by each model.

**ANALYTICAL STRATEGY**

Multiple linear regression was used to test if matching perceived causal attribution with treatment condition (respondent attributing the crime to dispositional factors after reading the internal treatment, and same for external treatment) or diverging from the parole candidate’s expressed causal attribution significantly predicted decisions to release the candidate. The fitted regression model was:

$$DV_i = \beta_0 + \beta_1 Dispositional_i + \beta_2 Situational_i + \beta' X_i + \epsilon_i$$

*X<sub>i</sub>* is a vector of pre-registered, pre-treatment covariates included to improve the precision of treatment effect estimates. The regression model was fitted separately to the three treatment conditions, to control for the causal attribution presented to the respondent. The analysis does not include a comparison of the coefficients across models.

**RESULTS**

Study 2 explains under which conditions external attribution increases others’ perceptions of dangerousness. The use of external attribution affects only respondents who hold the opposite opinion—namely, that internal factors caused the crime.

Table 2 shows that the attribution style affects the perception of dangerousness through the observer’s beliefs. Model 1 shows that holding internal casual attribution beliefs despite being presented with the external attribution vignette, predicts a decline of about 43-percentage points in the likelihood of releasing the candidate. Model 2 shows that believing an internal causal attribution despite not being presented with any explanation by the candidate (control condition) predicts a 60-percentage point decline in the likelihood of releasing the candidate. Model 3 shows that holding external casual attribution beliefs despite being presented with the internal attribution vignette, predicts an increase of about 43-percentage points in the likelihood of releasing the candidate.

## DISCUSSION

Participants in the study were significantly more likely to view someone as dangerous when presented with scenarios that used external explanations to account for a hypothetical parole candidate's past misconduct. Notably, this finding tracks the legal doctrine of "Insight," which equates the use of internal attribution by the parole candidate with reduced dangerousness. Thus, the legal doctrine might reflect a cognitive human bias. Moreover, external attributions also increased perceptions that the candidate was Black.

The second study validated the FAE's impact and the attribution congruence theory. Respondents showed a clear bias toward attributing behavior to internal factors, and when this belief matches the wrongdoer's own internal attribution of their crime, the result is lower perceived dangerousness. When this internal attribution bias conflicted with a parole candidate's own external explanations, the candidate was seen as posing the greatest risk. In sum, attribution congruence considers both the point of view of the decision-maker (parole board member, for example) and the attribution style of the wrongdoer; when they match, dangerousness is likely to be determined lower, and when they contrast, higher.

These findings contribute to Quinsey and Cyr's (1986) conclusions; in their study, external factors caused participants to judge the person as more dangerous and less "treatable." Instead of contradicting their findings, the results presented here suggest that attribution congruence between observer and observed is required. Study 2 shows that the adverse effect of external attribution is explained by whether the observer shares the same attitude toward the cause of crime.

These studies contribute to the long-standing recognition that an individual's internal attribution toward crime is the most crucial attributional dimension for punishment decisions (Carroll & Payne, 1977). While previous work concluded that the belief in internal attribution is associated with punitiveness (O'Toole & Sahar, 2014; Weiner et al., 1997), the present studies find that external attribution by an observed wrongdoer increases perceived dangerousness. Importantly, this article does not explain why some prefer an internal over external attribution; as discussed before, it might be related to a person's framework of thinking about criminal behavior—positivist or rationalist (determinism vs. free will, accordingly). Yet, taken together, this article argues that we should consider both the observer's perspective and the narrative of the observed. Attribution congruence might be key to understanding perceptions of dangerousness.

Perceptions of race were affected by the presented causal attribution. The use of external attribution increased racialization as a person of color compared with internal attribution explanations. Likewise, a similar but reversed correlation is found for perceiving the putative person who committed an offense as a White person—it increases after receiving the internal condition. This contributes to understanding the effect of the "crime script" (Gilliam & Iyengar, 2000): Previous studies showed that people would racialize others who committed violent crimes as non-White, and that there is a relationship between exposure through the media to a non-White person who perpetrated violence and internal attribution (Dixon, 2008; Gilliam & Iyengar, 2000). This article shows that the relationship works in both ways—the narrative of the person who perpetrated violence causes shifts in racial perceptions. To disentangle attitudes toward crime and justice (such as perceptions of dangerousness) from racial bias we should consider nuances of the criminal legal system, such as attribution theory.

### POSSIBLE IMPLICATIONS FOR PAROLE DECISION-MAKING

Despite not using actual parole board members as respondents, research on parole decision-making consistently finds racial bias, arbitrariness, and idiosyncrasy in parole decisions, despite adopting guidelines and actuarial tools in modern parole (Hritz, 2021). Parole boards consist of individuals from various backgrounds, and variations in the subjective assessment are significant (Ruhland, 2020), which raises the concern that excessive discretion opens the door to cognitive bias. A hallmark of parole decision-making is the important role of discretion: the California Supreme Court has found that the Board's "discretion in parole matters has been described as 'great' and 'almost unlimited'" (In re Rosencrantz, 2002); this raises concerns about cognitive bias interference. Further, existing research showed that Americans, in general, tend to endorse individualistic causes for social problems, which is often attributed to the influence of individualism as a core American value (King, 2007; Kornhauser, 2015).

The potential impact of the FAE on assessing suitability is particularly concerning in the context of individuals serving life sentences ("Lifers"). Except for people sentenced to die in prison, no other group is further marginalized and excluded from society like lifers. Most of the lifers already came from class and racial marginalization backgrounds—communities rife with poverty, violence, and the notorious mix of over-and-under policing (Irwin, 2009; Nellis, 2013). This might increase the Lifers' use of external attribution and lead to issues with attribution congruence. Indeed, the California guidelines consider evidence of an unstable social background against the candidate's suitability (California Department of Corrections and Rehabilitation, n.d.).

Given the significant weight of external causes in committing crimes, especially for low-income individuals and people of color, parole boards can allow the incarcerated person to recognize the social injustices they might have experienced (Maruna & Mann, 2006). As social injustices continue to be a condition of reality, recognizing and discussing them might suggest that the incarcerated person is better prepared for the challenge they pose.

### FUTURE DIRECTIONS AND LIMITATIONS

Because the studies rely on lay people and not actual parole board members, they only hold suggestive evidence for the parole board's decision-making process. Moreover, the generalization of the findings depends on the administrative setting—these studies used a specific example (murder, California), and thus, other settings require specific theorization. Further, the studies cannot reflect the impact of external political pressures faced by the parole boards (Chen et al., 2024; Laskorunsky et al., 2023; Schwartzapfel, 2015; Vîlcică, 2016). In California, for instance, the Governor may block parole if they believe that concerns for public safety overcome approving the parole grant (Governor's Parole Review, 1988). The studies also do not emphasize the role of the victim in the hearing (Roberts, 2009; Young, 2016). Future research should examine the impact of different uses of causal attribution on recidivism beliefs in more settings. Lab studies can contribute to our understanding of the interaction between causal attribution and the identity of the parole candidate by mimicking the in-person interview setting. For instance, researchers can simulate a discretionary parole hearing in which they control and randomly manipulate candidates' demographics and their causal attribution style. Researchers can also test the effect of



informing decision-makers about the possible impact of implicit bias, and FAE in particular, on their judgment decisions.

Outside the lab, parole candidates likely voice a mix of external and internal factors. Observers, in turn, might latch onto one dominant explanation—either internal or external—even when presented with a mixed set of factors. Regarding attribution congruence, this simplification bias might lead the observer to overemphasize whichever factor category (internal or external) they deem most salient or aligned with their preexisting beliefs. This interplay between mixed causal factors and observer attribution represents an intriguing avenue for future empirical research.

Moreover, researchers can gain access to observe parole hearings and code for causal attributions, demographics and candidate information, and hearing outcomes. Observational studies can be combined with interviews conducted with parole boards to assess the degree to which the FAE might influence them. From a policymaking perspective, understanding the mechanism (*how* external attributions increase dangerousness perceptions) can help design effective policies that oppose the effects of the FAE. These findings underscore the possible need for targeted unconscious bias training to enhance parole boards' awareness of the potential influence of unconscious cognitive biases, particularly the FAE (Atewologun et al., 2018). However, implicit bias training is insufficient; organizational structural changes can also lower the effects of implicit bias (Onyeador et al., 2021). For discretionary parole, this might include implementing a more standardized decision-making process with a stronger reliance on objective metrics such as program completion rates, disciplinary records, and quantifiable measures of rehabilitation progress.

## CONCLUSION

The current research contributes to understanding what affects perceptions of dangerousness and racial classification. It shows that dangerousness is affected by others' use of causal attribution explanations, such that attributing past wrongdoing to external factors increases perceptions of dangerousness. The FAE suggests that others are likely to believe that internal factors cause others to commit crimes. This study argues that a person is better off using internal causal attribution to explain wrongdoing because congruence with the observer's belief regarding the causes of criminal behavior results in lower perceived dangerousness.

## ORCID iD

Dvir Yogev  <https://orcid.org/0000-0003-0466-9804>

## SUPPLEMENTAL MATERIAL

Supplemental Tables S3–S10 and Supplemental Figures S4–S8 are available in the online version of this article at <http://journals.sagepub.com/home/cjb>

## REFERENCES

- Atewologun, D., Cornish, T., & Tresh, F. (2018). *Unconscious bias training: An assessment of the evidence for effectiveness* [Research report 113]. Equality and Human Rights Commission.
- Aviram, H. (2020). *Yesterday's monsters: The Manson family cases and the illusion of parole*. University of California Press.
- Barkan, S. E., & Cohn, S. F. (2005). Why Whites favor spending more money to fight crime: The role of racial prejudice. *Social Problems, 52*, 300–314. <https://doi.org/10.1525/sp.2005.52.2.300>

- Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, 20, 351–368. <https://doi.org/10.1093/pan/mpr057>
- Boas, T. C., Christenson, D. P., & Glick, D. M. (2020). Recruiting large online samples in the United States and India: Facebook, Mechanical Turk, and Qualtrics. *Political Science Research and Methods*, 8, 232–250. <https://doi.org/10.1017/psrm.2018.28>
- Bridges, G. S., & Steen, S. (1998). Racial disparities in official assessments of juvenile offenders: Attributional stereotypes as mediating mechanisms. *American Sociological Review*, 63, 554–570. <https://doi.org/10.2307/2657267>
- Brown, E. K., & Socia, K. M. (2017). Twenty-first century punitiveness: Social sources of punitive American views reconsidered. *Journal of Quantitative Criminology*, 33, 935–959. <https://doi.org/10.1007/s10940-016-9320-1>
- Busby, E. C., Gubler, J. R., & Hawkins, K. A. (2019). Framing and blame attribution in populist rhetoric. *The Journal of Politics*, 81, 616–630. <https://doi.org/10.1086/701832>
- California Department of Corrections and Rehabilitation. (n.d.). *Lifer parole process*. <https://www.cdcr.ca.gov/bph/lifer-parole-process/>
- California Department of Corrections and Rehabilitation. (2020). *Overview of parole hearings and statistics*.
- Carroll, J. S., & Payne, J. W. (1977). Crime seriousness, recidivism risk, and causal attributions in judgments of prison term by students and experts. *Journal of Applied Psychology*, 62, 595–602. <https://doi.org/10.1037/0021-9010.62.5.595>
- Chambers, J. M., Freeny, A. E., & Heiberger, R. M. (2017). Analysis of variance: Designed experiments. In J. M. Chambers & T. J. Hastie (Eds.), *Statistical models in S* (pp. 145–193). Routledge.
- Chen, Y.-H., Vilciã, E. R., Ward, J., Maynard, M., Eberhardt, C., & Olaghere, A. (2024). Legitimacy of parole as a consequence of policy shock: The lived experiences of incarcerated persons during the parole moratorium in Pennsylvania. *U.S. Law & Policy*, 46(3), 246–276. <https://doi.org/10.1111/lapo.12239>
- Clarkson, J. J., Chambers, J. R., Hirt, E. R., Otto, A. S., Kardes, F. R., & Leone, C. (2015). The self-control consequences of political ideology. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 8250–8253. <https://doi.org/10.1073/pnas.1503530112>
- Cooper, M., Updegrove, A. H., Gabbidon, S. L., Higgins, G. E., & Potter, H. (2021). Examining the effects of race/ethnicity and race relations on public opinion about the overrepresentation of black people in US prisons. *Victims & Offenders*, 16, (1): 1–19. <https://doi.org/10.1080/15564886.2020.1850579>
- Coppock, A., & McClellan, O. A. (2019). Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. *Research & Politics*, 6(1), 1–14. <https://doi.org/10.1177/2053168018822174>
- Cullen, F. T., Clark, G. A., Cullen, J. B., & Mathers, R. A. (1985). Attribution, salience, and attitudes toward criminal sanctioning. *Criminal Justice and Behavior*, 12, 305–331. <https://doi.org/10.1177/0093854885012003003>
- Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108, 353–380. <https://doi.org/10.1016/j.cognition.2008.03.006>
- Dafoe, A., Zhang, B., & Caughey, D. (2018). Information equivalence in survey experiments. *Political Analysis*, 26, 399–416. <https://doi.org/10.1017/pan.2018.9>
- Dixon, T. L. (2008). Crime news and racialized beliefs: Understanding the relationship between local news viewing and perceptions of African Americans and crime. *Journal of Communication*, 58, 106–125. <https://doi.org/10.1111/j.1460-2466.2007.00376.x>
- Dripps, D. A. (2003). Fundamental retribution error: Criminal justice and the social psychology of blame. *Vanderbilt Law Review*, 56, 1383–1438.
- Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing Black: Race, crime, and visual processing. *Journal of Personality and Social Psychology*, 87, 876–893. <https://doi.org/10.1037/0022-3514.87.6.876>
- ElBassiouny, A., Ayala, R., & Sircy, K. (2022). Factors influencing mock parole board members: The impact of an incarcerated person's race/ethnicity, conversion while in prison, and mental health status. *Analyses of Social Issues and Public Policy*, 22, 286–303. <https://doi.org/10.1111/asap.12286>
- Engel, C., & Grgić-Hlača, N. (2021). Machine advice with a warning about machine limitations: Experimentally testing the solution mandated by the Wisconsin Supreme Court. *Journal of Legal Analysis*, 13, 284–340. <https://doi.org/10.1093/jla/laab001>
- Esberg, J., & Mummolo, J. (2018). Explaining misperceptions of crime [Unpublished manuscript].
- Garland, D. (2001). *The culture of control: Crime and social order in contemporary society*. Oxford University Press.
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, 117, 21–38. <https://doi.org/10.1037/0033-2909.117.1.21>
- Gilliam, F. D., & Iyengar, S. (2000). Prime suspects: The influence of local television news on the viewing public. *American Journal of Political Science*, 44, 560–573. <https://doi.org/10.2307/2669264>
- Governor's Parole Review. (1988). California Proposition 89: Amending Art. V, § 8 of the California Constitution and enacting Cal. Penal Code § 3041.2.
- Green, E. G. T., Staerklé, C., & Sears, D. O. (2006). Symbolic racism and Whites' attitudes towards punitive and preventive crime policies. *Law and Human Behavior*, 30, 435–454. <https://doi.org/10.1007/s10979-006-9020-5>

- Hamilton, C. V. (1988). [Review of the book *Racial formation in the United States from the 1960s To the 1980s*, by M. Omi & H. Winant]. *Political Science Quarterly*, 103, 158–159.
- Harcourt, B. E. (2015). Risk as a proxy for race: The dangers of risk assessment. *Federal Sentencing Reporter*, 27, 237–243. <https://doi.org/10.1525/fsr.2015.27.4.237>
- Harman, G. (1999). Moral philosophy meets social psychology: Virtue ethics and the Fundamental Attribution Error. *Proceedings of the Aristotelian Society*, 99, 315–331. <https://doi.org/10.1111/1467-9264.00062>
- Hawkins, D. F. (1981). Causal attribution and punishment for crime. *Deviant Behavior*, 2, 207–230. <https://doi.org/10.1080/01639625.1981.9967554>
- Headworth, S. (2021). Broke people, broken rules: Explaining welfare fraud investigators' attributions. *Punishment & Society*, 23(1), 24–46. <https://doi.org/10.1177/1462474520928131>
- Heider, F. (1958). *The psychology of interpersonal relations*. Wiley.
- Hetey, R. C., & Eberhardt, J. L. (2018). The numbers don't speak for themselves: Racial disparities and the persistence of inequality in the criminal justice system. *Current Directions in Psychological Science*, 27, 183–187. <https://doi.org/10.1177/0963721418763931>
- Hritz, A. C. (2021). Parole board decision-making and constitutional rights. *Annual Review of Law and Social Science*, 17(3), 335–351. <https://doi.org/10.1146/annurev-lawsocsci-051121-070621>
- Huebner, B. M., & Bynum, T. S. (2008). The role of race and ethnicity in parole decisions. *Criminology*, 46(4), 907–938. <https://doi.org/10.1111/j.1745-9125.2008.00130.x>
- Hurwitz, J., & Peffley, M. (2005). Playing the race card in the post-Willie Horton era: The impact of racialized code words on support for punitive crime policy. *Public Opinion Quarterly*, 69(1), 99–112. <https://doi.org/10.1093/poq/nfi004>
- In re Rosencrantz (2002) 29 Cal.4th 616.
- In re Shaputis, 190 P.3d 573 (Cal. 2008).
- In re Shaputis, 265 P.3d 253, 275 (Cal. 2011).
- Irwin, J. (2009). *Lifers: Seeking redemption in prison*. Routledge.
- King, A. (2007). Keeping a safe distance: Individualism and the less punitive public. *British Journal of Criminology*, 48(2), 190–208. <https://doi.org/10.1093/bjc/azm069>
- Kornhauser, R. (2015). Economic individualism and punitive attitudes: A cross-national analysis. *Punishment & Society*, 17(1), 27–53. <https://doi.org/10.1177/1462474514560393>
- Kraemer, H. C., Kazdin, A. E., Offord, D. R., Kessler, R. C., Jensen, P. S., & Kupfer, D. J. (1997). Coming to terms with the terms of risk. *Archives of General Psychiatry*, 54(4), 337–343. <https://doi.org/10.1001/archpsyc.1997.01830160065009>
- Lacey, N., Soskice, D. W., Cheliotis, L., & Xenakis, S. (2021). *Tracing the relationship between inequality, crime, and punishment: Space, time, and politics*. Oxford University Press.
- Laskorunsky, J., Mitchell, K. L., & Mullins, S. F. (2023). *Risk averse and disinclined: What COVID prison releases demonstrate about the ability of the U.S. to reduce mass incarceration*. University of Minnesota Robina Institute of Criminal Law and Criminal Justice.
- Maruna, S., & Mann, R. E. (2006). A Fundamental Attribution Error? Rethinking cognitive distortions. *Legal and Criminological Psychology*, 11(2), 155–177. <https://doi.org/10.1348/135532506X114608>
- McNeish, D., & Wolf, M. G. (2020). Corrigendum to: Thinking twice about sum scores. *Behavior Research Methods*, 52, 2674. <https://doi.org/10.3758/s13428-020-01468-3>
- Medwed, D. S. (2008). The innocent prisoner's dilemma: Consequences of failing to admit guilt at parole hearings. *Iowa Law Review*, 93, 491–558.
- Monahan, J. (2006). A jurisprudence of risk assessment: Forecasting harm among prisoners, predators, and patients. *Virginia Law Review*, 92, 391–435.
- Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology*, 67(6), 949–971. <https://doi.org/10.1037/0022-3514.67.6.949>
- Mummolo, J., & Peterson, E. (2019). Demand effects in survey experiments: An empirical assessment. *American Political Science Review*, 113, 517–5292. <https://doi.org/10.1017/S0003055418000837>
- Nellis, A. (2013, September 18). Life goes on: The historic rise in life sentences in America. *The Sentencing Project*.
- O'Toole, M. J., & Sahar, G. (2014). The effects of attributions for crime on attitudes toward prison reform. *Applied Psychology in Criminal Justice*, 10(1), 46–65.
- Onyeador, I. N., Hudson, S. T. J., & Lewis, N. A. (2021). Moving beyond implicit bias training: Policy insights for increasing organizational diversity. *Policy Insights from the Behavioral and Brain Sciences*, 8(1), 19–26. <https://doi.org/10.1177/2372732220983840>
- Peffley, M., & Hurwitz, J. (2002). The racial components of “race-neutral” crime policy attitudes. *Political Psychology*, 23(1), 59–75. <https://doi.org/10.1111/0162-895X.00270>
- Penner, A. M., & Saperstein, A. (2008). How social status shapes race. *Proceedings of the National Academy of Sciences of the United States of America*, 105(50), 19628–19630. <https://doi.org/10.1073/pnas.0805762105>
- Perez, M. J., & Salter, P. S. (2019). Trust, innocence, and individual responsibility: Neoliberal dreams of a colorblind peace. *Journal of Social Issues*, 75(1), 267–285. <https://doi.org/10.1111/josi.12317>

- Pickett, J. T., & Baker, T. (2014). The pragmatic American: Empirical reality or methodological artifact. *Criminology*, 52(2), 195–222. <https://doi.org/10.1111/1745-9125.12035>
- Quinsey, V. L., & Cyr, M. (1986). Perceived dangerousness and treatability of offenders: The effects of internal versus external attributions of crime causality. *Journal of Interpersonal Violence*, 1(4), 458–471. <https://doi.org/10.1177/088626086001004005>
- Rhine, E. E., Petersilia, J., & Reitz, K. R. (2017). The future of parole release. *Crime and Justice*, 46(1), 279–338. <https://www.journals.uchicago.edu/doi/10.1086/688616>
- Roberts, J. V. (2009). Listening to the crime victim: Evaluating victim input at sentencing and parole. *Crime and Justice*, 38, 347–412. <https://doi.org/10.1086/599203>
- Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60357-3](https://doi.org/10.1016/S0065-2601(08)60357-3)
- Ross, L., & Nisbett, R. E. (2011). *The person and the situation: Perspectives of social psychology* (2nd ed.). Pinter & Martin.
- Ruhland, E. L. (2020). Philosophies and decision making in parole board members. *The Prison Journal*, 100(5), 640–661. <https://doi.org/10.1177/0032885520956566>
- Saperstein, A., & Penner, A. M. (2010). The race of a criminal record: How incarceration colors racial perceptions. *Social Problems*, 57(1), 92–113. <https://doi.org/10.1525/sp.2010.57.1.92>
- Saperstein, A., Penner, A. M., & Kizer, J. M. (2014). The criminal justice system and the racialization of perceptions. *The Annals of the American Academy of Political and Social Science*, 651(1), 104–121. <https://doi.org/10.1177/0002716213503097>
- Schwartzapfel, B. (2015). Parole boards: Problems and promise. *Federal Sentencing Reporter*, 28(2), 79–84. <https://doi.org/10.1525/fsr.2015.28.2.79>
- Shammas, V. L. (2019). The perils of parole hearings: California lifers, performative disadvantage, and the ideology of insight. *PoLAR: Political and Legal Anthropology Review*, 42(1), 142–160.
- Simon, J. (2005). Reversal of fortune: The resurgence of individual risk assessment in criminal justice. *Annual Review of Law and Social Science*, 1, 397–421. <https://doi.org/10.1146/annurev.lawsocsci.1.041604.120006>
- Skeem, J. L., & Monahan, J. (2011). Current directions in violence risk assessment. *Current Directions in Psychological Science*, 20(1), 38–42. <https://doi.org/10.1177/0963721410397271>
- Skeem, J. L., Scurich, N., & Monahan, J. (2020). Impact of risk assessment on judges' fairness in sentencing relatively poor defendants. *Law and Human Behavior*, 44(1), 51–59. <https://doi.org/10.1037/lhb0000360>
- Tetlock, P. E. (1985). Accountability: A social check on the Fundamental Attribution Error. *Social Psychology Quarterly*, 48(3), 227–236. <https://doi.org/10.2307/3033683>
- Tetlock, P. E., Visser, P. S., Singh, R., Polifroni, M., Scott, A., Elson, S. B., Mazzocco, P., & Rescober, P. (2007). People as intuitive prosecutors: The impact of social-control goals on attributions of responsibility. *Journal of Experimental Social Psychology*, 43(2), 195–209. <https://doi.org/10.1016/j.jesp.2006.02.009>
- Tonry, M. (2019). Predictions of dangerousness in sentencing: Déjà vu all over again. *Crime and Justice*, 48, 439–482. <https://doi.org/10.1086/701895>
- Triandis, H. C. (2001). Individualism-collectivism and personality. *Journal of Personality*, 69(6), 907–924. <https://doi.org/10.1111/1467-6494.696169>
- Unnever, J. D., & Cullen, F. T. (2010). The social sources of Americans' punitiveness: A test of three competing models. *Criminology*, 48(1), 99–129. <https://doi.org/10.1111/j.1745-9125.2010.00181.x>
- Vilciică, E. R. (2016). Studying parole in the spotlight: Lessons from a large American jurisdiction. *European Journal on Criminal Policy and Research*, 22, 61–88. <https://doi.org/10.1007/s10610-015-9284-8>
- Vilciică, E. R. (2018). Revisiting parole decision making: Testing for the punitive hypothesis in a large U.S. jurisdiction. *International Journal of Offender Therapy and Comparative Criminology*, 62(5), 1357–1383. <https://doi.org/10.1177/0306624X16668512>
- Wacquant, L. (2009). *Punishing the poor: The neoliberal government of social insecurity*. Duke University Press.
- Walker, M. L. (2016). Race making in a penal institution. *American Journal of Sociology*, 121(4), 1051–1078. <https://doi.org/10.1086/684033>
- Watts, A., & Rhine, E. E. (2018). *Opinion: The human cost of politicized parole consideration*. Robina Institute of Criminal Law and Criminal Justice.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92(4), 548–573. <https://doi.org/10.1037/0033-295X.92.4.548>
- Weiner, B. (2006). *Social motivation, justice, and the moral emotions: An attributional approach*. Psychology Press.
- Weiner, B., & Graham, S. (1990). Attribution in personality psychology. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 465–485). The Guilford Press.
- Weiner, B., Graham, S., & Reyna, C. (1997). An attributional examination of retributive versus utilitarian philosophies of punishment. *Social Justice Research*, 10, 431–452. <https://doi.org/10.1007/BF02683293>
- Werth, R. (2019). Theorizing the performative effects of penal risk technologies: (Re)producing the subject who must be dangerous. *Social & Legal Studies*, 28(3), 327–348. <https://doi.org/10.1177/0964663918773542>

- Western, B. (2006). *Punishment and inequality in America*. Russell Sage Foundation.
- Yelderman, L. A., Lawrence, T. I., Lyons, C. E., & DeVault, A. (2021). Actor-observer asymmetry in perceptions of parole board release decisions. *Psychiatry, Psychology, and Law*, 28(5), 623–644. <https://doi.org/10.1080/13218719.2020.1821826>
- Yogev, D., & Mehozay, Y. (2022). Comparative criminology in the time of algorithmic knowledge: The challenges of global comparison. In D. Nelken, D. Siegel, & R. van Swaaningen (Eds.), *Research handbook of comparative criminology* (pp. 365–382). Edward Elgar.
- Young, K. M. (2016). Parole hearings and victims' rights: Implementation, ambiguity, and reform. *Connecticut Law Review*, 49, 431–498.
- Young, K. M. (2020). Legal ruralism and California parole hearings: Space, place, and the carceral landscape. *Rural Sociology*, 85, 938–965. <https://ssrn.com/abstract=4198276>
- Young, K. M., & Chimowitz, H. (2022). How parole boards judge remorse: Relational legal consciousness and the reproduction of carceral logic. *Law & Society Review*, 56(2), 237–260. <https://doi.org/10.1111/lasr.12601>
- Young, K. M., Mukamal, D. A., & Favre-Bulle, T. (2016). Predicting parole grants: An analysis of suitability hearings for California's lifer inmates. *Federal Sentencing Reporter*, 28(4), 268–277. <https://doi.org/10.1525/fsr.2016.28.4.268>
- Young, K. M., & Pearlman, J. (2021). Racial disparities in lifer parole outcomes: The hidden role of professional evaluations. *Law & Social Inquiry*, 46, 625–662.
- Zollinger, D. (2022). Cleavage identities in voters' own words: Harnessing open-ended survey responses. *American Journal of Political Science*, 68, 139–159. <https://doi.org/10.1111/ajps.12743>

**Dvir Yogev** is a post-doctoral researcher at UC Berkeley's Criminal Law & Justice Center. He holds a PhD in Jurisprudence and Social Policy from UC Berkeley.