

“Internally Wicked”: Investigating How and Why Essentialism Influences  
Punitiveness and Moral Condemnation

Justin W. Martin<sup>1</sup> & Larisa Heiphetz<sup>2\*</sup>

<sup>1</sup> Department of Psychology, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA.

Email: justin.martin.3@bc.edu

<sup>2</sup> Department of Psychology, Columbia University, 1190 Amsterdam Ave., New York, NY

10027. Email: lah2201@columbia.edu

\*Corresponding author

In press, *Cognitive Science*

Acknowledgements: The authors wish to thank Fiery Cushman, Susan Gelman, Phillip Atiba Goff, Liane Young, and the SHAME writing group for their comments on this manuscript. This work was funded by NSF Grant #DGE1144152 and TREAT grant from the Department of Psychology at Boston College to JWM. This work was also funded by NSF Grant #1408989, John Templeton Foundation Grant #61080, a Visiting Scholar Award from the Russell Sage Foundation, and Columbia University’s Provost’s Grants Program for Junior Faculty who Contribute to the Diversity Goals of the University to LH. Any opinions expressed are those of the authors alone and should not be construed as representing the opinions of the John Templeton Foundation or the Russell Sage Foundation.

### Abstract

Kant argued that individuals should be punished “proportional to their internal wickedness,” and recent work has demonstrated that essentialism—the notion that observable characteristics reflect internal, biological, unchanging “essences”—influences moral judgment. However, these efforts have yielded conflicting results: essentialism sometimes increases and sometimes decreases moral condemnation. To resolve these discrepancies, we investigated the mechanisms by which essentialism influences moral judgment, focusing on perceptions of actors’ control over their behavior, the target of essentialism (particular behaviors versus actors’ character), and the component of essentialism (biology versus immutability). Participants punished people described as having a criminal essence more than those with a non-criminal essence or no essence. Probing potential mechanisms underlying this effect, we found a mediating role for perceptions of control and weak influences of essentialism focus (behavior versus character) and component of essentialism (biology versus immutability). These results extend prior work on essentialism and moral cognition, demonstrating a causal link between perceptions of “internal wickedness” and moral judgment. Our findings also resolve discrepancies in past work on the influence of essentialism on moral judgment, highlighting the role that perceptions of actors’ control over their behavior play in moral condemnation.

Keywords: essentialism; moral cognition; punishment; social cognition; law; criminal justice

## 1. Introduction

Americans love stories of change, except for when they don't. They can't get enough of stories describing people who overcame difficulties and went on to succeed, eagerly retelling stories of millionaires who grew up in poverty, divorcees who became better partners the second time around, and addicts who "got clean" and went on to lead fulfilling lives (McAdams, 2006). But when Michelle Jones applied to Harvard's doctoral history program, she learned that redemption was not for her. Ms. Jones had served more than two decades in prison, a time during which she also published scholarship on American history and used teleconferencing to present her work to academics and politicians. Harvard's history program admitted her, but Harvard's administration later overturned this decision. A formerly incarcerated person could not redeem herself enough to earn a place at Harvard's table.<sup>1</sup>

Although it is impossible to know all of the reasons why administrators overturned Jones' admission (for example, some of their thought process may not have become public), an intuitive possibility is the stigma associated with her incarceration. The idea that past behavior has an enduring impact on a person's "moral essence" has profound consequences. Like Ms. Jones, many formerly incarcerated individuals experience difficulty when seeking to re-enter life beyond the prison walls, including struggles with finding employment, securing housing, and maintaining (or establishing) family ties—struggles that stem in part from the common perception that formerly incarcerated people are bad people who will never change (Alexander, 2012; Pager, 2007; Travis & Maul, 2003). Here, we ask: how does such psychological essentialism influence moral judgment, especially moral condemnation of people who have committed crimes, and what mechanisms underlie this judgment?

---

<sup>1</sup> Details taken from <https://nyti.ms/2x2kQHP>

### **1.1 Connections between psychological essentialism and moral cognition.**

Psychological essentialism refers to the lay notion that internal, invisible “essences” separate members of different categories (Cimpian & Salomon, 2014; Haslam et al., 2002; Medin & Ortony, 1989). Essentialized characteristics are often perceived to be biological and unchanging (Gelman, 2003; Haslam et al., 2000). Thus, a person who holds an essentialist view of criminality may posit that some people have an internal criminal essence that distinguishes them from non-criminals, that this essence is rooted in biology, and that it cannot change over time.

Previous research has found evidence of essentialism in domains such as race, ethnicity, gender, mental illness, religion, and social class (Berryessa, 2019, 2020; Cohen, 2004; Dar-Nimrod et al., 2011; Dar-Nimrod & Heine, 2011; Harris & Socia, 2016; Haslam & Levy, 2006; Heiphetz et al., 2017; Imhoff & Jahnke, 2018; Jayaratne et al., 2006; Kraus & Keltner, 2013; Prentice & Miller, 2006; Roberts et al., 2017), and essentialism in these domains has a number of social consequences (e.g. Dar-Nimrod et al., 2011; Keller, 2005). In particular, psychological essentialism negatively influences perceptions of and behaviors toward individuals with these characteristics. In one line of work, inducing an essentialist understanding of gender increased participants’ endorsement of gender stereotypes (Brescoll & LaFrance, 2004). Other work has found that individual differences in essentialism predict support for legislation restricting the rights of transgender people (Roberts et al., 2017). In another experiment, White and Asian men who read an ostensible scientific article priming an essentialist, biological understanding of race exhibited less interest in befriending a Black individual than did participants who read an article priming a social understanding of race (Williams & Eberhardt, 2008). Indeed, essentialism appears to increase anti-Black bias even among Black participants (Mandalaywala et al., 2017).

Essentialism also appears influential early in development; in one line of work, 4- to 6-year-olds shared fewer resources with a novel group when they were led to perceive that group in essentialist, versus non-essentialist, terms (Rhodes et al., 2018).

Past work on the relation between essentialism and moral judgment has revealed two distinct patterns. On the one hand, some work demonstrates that essentialism reduces moral condemnation. Men who read an explanation of rape that highlighted men's biological predispositions to commit sexual crimes expressed more leniency toward rapists than did men who read an explanation highlighting external, social factors (Dar-Nimrod et al., 2011). In another study, participants assigned less punishment to a man who committed murder due to a biological predisposition than to a man who committed an identical crime due to childhood abuse (Monterosso et al., 2005). Related work focusing on legal decision-making has found that providing information about a genetic versus an environmental basis for a predisposition toward violence led to reduced recommended sentences and that this relationship was mediated both by perceptions of conscious control (which reduced sentences) and internal causal attributions (which increased sentences; Cheung & Heine, 2015). Finally, trial judges in the United States assigned a reduced sentence to a hypothetical individual when evidence of a biomechanical basis for the individual's psychopathy was included relative to when this biomechanical information was not included (Aspinwall et al., 2012).

On the other hand, some work demonstrates that essentialism increases punitiveness. In one series of studies, viewing social class in essentialist terms led individuals to endorse harsh retributive punishment (e.g. expulsion for cheating in a course), whereas viewing social class as socially constructed led individuals to endorse restorative punishment (e.g. community service; Kraus & Keltner, 2013). Another line of work found that the degree to which participants

essentialized sexual crimes and robbery predicted increased support for restrictive societal policies, including offender registries, occupational restrictions, and the confiscation of assets (de Vel-Palumbo et al., 2019). Other work has found a similar relation between individual differences in tendency to essentialize criminality and punishment using a range of harm severities (Martin et al., forthcoming). Similarly, labeling people who commit sexual crimes using noun labels suggesting that the crime reflects the person's essence (e.g. "sex offender", "pedophile") compared to generic language (e.g. "people who have committed crimes of a sexual nature") leads to greater support for restrictive policies (Harris & Socia, 2016) and increased punitive attitudes (Imhoff & Jahnke, 2018). In other research, when a person deliberately performed a bad action (e.g. stealing money or smashing a car window), those who viewed that bad action as reflecting the person's essential "true self" blamed them more (Newman et al., 2015). Finally, qualitative work on US judges found that those who expressed essentialist views regarding mental disorders with known genetic influences that could have led to someone's crime, were more likely to support restrictive sentencing (Berryessa, 2019).

Although much prior work has shown some relationship between essentialism and punishment, a smaller group of studies also exists that has not found such a link (Fuss et al., 2015; Scurich & Appelbaum, 2016). Furthermore, some research has found that some aspects of essentialism lead to increased support for punishment (including informativeness and inherence), whereas other aspects of essentialism lead to reduced support for punishment (including discreteness and continuity) (Berryessa, 2020). Given these varying patterns in past work, additional research is needed to clarify how and why essentialism and punitiveness may be related.

## **1.2 Novel contributions of the current work**

The current research sought to reconcile these differing results. We focused on three differences between the studies finding an accentuating versus attenuating influence of essentialism on moral condemnation that may serve as potential explanations: first, perceptions of the actor's control over their behavior; second, the target of essentialism (a particular behavior vs. the entirety of the actor's character); and third, the component of essentialism targeted (biology versus immutability).

First, when essentialism reduced moral condemnation, essentialism also reduced perceptions of the perpetrator's volition. For instance, Dar-Nimrod and colleagues (2011) found that explanations of rape that highlighted men's biological predispositions to commit sexual crimes, versus social causes, led participants both to express more leniency toward rapists and to view men as having less control over their sexual urges. Monterosso and colleagues (2005) manipulated essentialism by describing the actor as having "five times the amount of a particular chemical" in the actor's brain, which reduced both punishment and perceptions that the behavior in question was voluntary. Similarly, Cheung & Heine (2015) found that providing information about a genetic basis for a predisposition to violence led to reduced perceptions of a perpetrator's conscious control over his actions and thereby reduced punishment.

Most studies in which essentialism increased condemnation (de Vel-Palumbo et al., 2019; Harris & Socia, 2016; Imhoff & Jahnke, 2018; Kraus & Keltner, 2013; Martin et al., forthcoming; Newman et al., 2015) did not measure perceptions of control. Therefore, it is possible that essentialism altered perceptions of control in the former studies but not the latter studies, yielding different influences of essentialism on moral condemnation. Consistent with this possibility is the fact that in the one study that did include measures about volition and in which essentialism led to increased punitive attitudes (Imhoff & Jahnke, 2018), essentialist

labels did not impact perceptions of control. More generally, this possibility suggests that essentialism may reduce punitiveness when it reduces perceptions of individuals' control over the outcomes they produce (Ryazanov & Christenfeld, 2018). Furthermore, this idea is consistent with past work on the relation between control and moral judgment (Alicke, 2000; Darley et al., 2000; Martin & Cushman, 2016b; Robinson & Darley, 1995). This relation would also be consistent with past work on explanations for punishment, which has shown that people punish in part to change others' behavior (Clutton-Brock & Parker, 1995; Martin & Cushman, 2016a; Morris et al., 2017; Trivers, 1971). If people cannot control the behavior in question now, they are likely not to have control over it in the future either. Therefore, punishment cannot serve its function of changing behavior and should be withheld if one's goal is to change behavior (Martin & Cushman, 2016a). Thus, manipulations that reduce perceptions of control over behavior could decrease punitiveness.

Second, prior studies finding that essentialism decreases condemnation have tended to focus on essentialism of specific criminal behaviors, such as rape, murder or battery (e.g., Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005). Critically, what people *do* and who they *are* are distinct. When essentialism targets behavior, it may impact perceptions of the actor's control over this behavior. To the extent that someone's behavior is unchanging or biologically induced, others may view their behavior as outside their control and therefore not an appropriate target for moral condemnation. For instance, observers may judge offensive speech as less blameworthy if the speaker has Tourette's syndrome than if she chose to say hurtful things.

However, when essentialism targets the entirety of a person's moral character, participants may be unwilling to reduce perceptions of *all* of a person's behaviors as under their

control, consistent with general resistance to determinist claims (Nahmias et al., 2005; Nichols & Knobe, 2007). Instead, participants may focus on the badness of the characteristic itself, emphasizing the notion that those who harm others are bad people who should receive severe punishment. Indeed, Kraus and Keltner (2013) found that essentializing social class (a characteristic and not a behavior) led individuals to endorse harsh retributive punishment (e.g. expulsion for cheating in a course), whereas the view that social class is socially constructed led individuals to endorse restorative punishment (e.g. community service). Newman and colleagues (2015) found that participants who viewed behavior as reflective of an actor's "true self"—a general characteristic—blamed them more for bad behavior. Imhoff and Jahnke (2018) and Harris and Socia (2016) both focused on general classes related to sexual offenses—"pedophilia" and "sexual offenders," respectively—and found that these labels led to harsher punitive attitudes relative to generic language (e.g. "people who have committed crimes of a sexual nature"). These labels may have encouraged participants to focus on the more general characteristic that these labels connote rather than any specific behavior that might be performed by these individuals. Finally, while de Val-Palumbo and colleagues (2019) focused on essentialism regarding specific crimes, they did so in very general terms. For instance, they asked about essentialism regarding sex offenses as a broad category rather than asking about specific instances of rape, as Dar-Nimrod and colleagues (2011) did. Thus, discrepancies among past results may reflect differing targets of essentialism. It is possible that essentializing *behaviors* diminishes perceptions of control and therefore increases leniency, whereas essentializing *characteristics* (such as social class or one's true self) is associated with negative attitudes and greater punitiveness toward individuals possessing those characteristics.

Third, prior studies demonstrating that essentialism reduces moral condemnation have tended to focus on the biological component of essentialism. For instance, Aspinwell and colleagues (2012) focused on a specific neurobiological explanation of psychopathy involving enzymes and neurotransmitters, Cheung and Heine (2015) described a genetic predisposition toward violence, Dar-Nimrod and colleagues (2011) highlighted men's biological predispositions to commit sexual crimes, and Monterosso and colleagues (2005) described the actor as having "five times the amount of a particular chemical" in his brain. Studies finding that essentialism increases moral condemnation (de Vel-Palumbo et al., 2019; Kraus & Keltner, 2013; Martin et al., forthcoming; Newman et al., 2015) have tended not to highlight the biology component of essentialism. We also note that some qualitative work has found that, on the whole, biases about the genetic nature of mental illnesses experienced by people who have committed crimes lead to harsher punitive attitudes (Berryessa, 2020). Importantly, essentialism contains multiple components (Cimpian & Salomon, 2014; Haslam et al., 2000; Medin & Ortony, 1989) and while an important one is *biology* (e.g. the view that essential traits are biological in nature), *immutability* (e.g. the view that essential traits are unchangeable; Gelman et al., 2007; Heiphetz, Gelman, et al., 2017) plays a large role as well. However, the relation between essentialist descriptions that principally focus on immutability and moral condemnation is unclear. Some work has found that neither the immutability nor biology component of essentialism predicts recommended years in prison (Berryessa, 2020). Relatedly, a manipulation describing pedophilia as either controllable or not had no influence on punishment (Imhoff & Jahnke, 2018), though the authors suggest caution in interpreting this lack of effect due to potential issues with their manipulation. Thus, essentialist descriptions that principally focus on biology may

reduce moral condemnation, at least in part through a reduction in perceptions of actors' control over their behavior, whereas those focusing on immutability may have little to no impact.

In addition to helping to resolve discrepancies in prior research, examining the relation between essentialism and moral judgment contributes to the scientific literature in two additional ways. First, it clarifies how individuals perceive people who are involved in the justice system—a particularly stigmatized, yet under-studied, social group. The United States incarcerates more than 700 out of every 100,000 residents, the highest incarceration rate in the world (Mears et al., 2015). Incarceration imbues individuals with stigma that persists after their confinement, creating numerous difficulties for these individuals: employers do not want to hire them, landlords do not want to rent to them, and their families sometimes want nothing to do with them (Alexander, 2012; Clear, 2007; Pager, 2007; Travis & Maul, 2003; Western, 2018).

Incarceration may have such persistent consequences partially because individuals view people who have been imprisoned through an essentialist lens. That is, they may view these individuals as essentially criminal, unlikely to change and be rehabilitated. Our research tests this possibility by investigating the consequences of viewing people who have committed crimes in essentialist terms and probing the mechanisms that underlie the influence of essentialism on these outcomes.

Second, examining how essentialism shapes morality—particularly decisions regarding punishment—sheds light on the cognitive processes underlying moral judgments. Research on moral judgment has highlighted the prominent role of intent (Hauser et al., 2006; Hebble, 1971; Heiphetz & Young, 2014; Zelazo et al., 1996). People condemn those who intend to cause harm, even if they fail, and forgive those who do not intend harm, but cause it anyway (Cushman, 2008; Cushman et al., 2009; Martin & Cushman, 2015; Young et al., 2007). This focus on intent could potentially reflect a broader tendency to base moral judgment on internal characteristics.

Kant famously wrote that individuals should be punished “proportional to their internal wickedness” (1790/1952), and laypeople appear to agree, weighing information about an individual’s character as well as the specific behavior performed when making a moral judgment (Goodwin, 2015; Uhlmann et al., 2015; Uhlmann & Zhu, 2014). For instance, while individuals view beating one’s girlfriend as more immoral than beating a cat, they view the person who abused animals as having worse moral character than the person who abused a fellow human being (Tannenbaum et al., 2011). This person-centered approach to morality suggests that viewing individuals as having a central moral tendency—an “essential moral self” (Strohming & Nichols, 2014)—may change how they are evaluated. We investigated the consequences of viewing individuals as possessing an inherently criminal “essence” and mechanisms that may underlie the influence of essentialist views of criminality on social perceptions.

### **1.3 Overview of current studies**

The main goal of the present work was to examine the mechanisms underlying the influence of essentialism on punishment and, in doing so, to potentially resolve discrepancies among prior findings. We began by creating a novel set of vignettes to investigate the influence of essentialism on punishment. In each, an actor committed a criminal act. Whereas much past work has focused on only one or two types of criminal behavior (Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; de Vel-Palumbo et al., 2019; Harris & Socia, 2016; Imhoff & Jahnke, 2018; Kraus & Keltner, 2013), we included a range of harm severities: property damage, theft, assault, maiming, and murder. In addition, in Study 1, participants rated punishment in response to vignettes describing an actor in essentialist language highlighting his criminality, essentialist language highlighting another aspect of his character, or non-essentialist language. In this way, we tested how manipulating essentialism might alter punitiveness.

Having investigated the influence of essentialism on punishment, we then tested two potential mechanisms underlying the effect of essentialism on moral judgment—one emphasizing perceptions of control (Study 2) and the other highlighting potential differences between essentialist perceptions of behaviors versus characteristics (Study 3). Taken together, these studies provide insight into the processes underlying moral judgment—especially the influence of perceptions of essentialism, control and punishment—and uncover the cognitive underpinnings of perceptions regarding people involved in the justice system.

## **2. Study 1**

In Study 1, we developed a set of vignettes with which we could investigate how essentialist views of criminality influenced punishment. We created a novel set of vignettes varying in severity from theft to murder and asked participants how much punishment the actor in each vignette should receive. Participants read vignettes in one of three conditions which described the actor's behavior in essentialist language highlighting his criminality, essentialist language highlighting another aspect of his character, or non-essentialist language.

### **2.1 Methods**

#### **2.1.1 Participants**

Participants ( $n = 360$ ) were recruited using Amazon Mechanical Turk and received \$2.11. Sample size was determined a priori. As we were unsure of what effect size to expect, we set a goal of retaining data from at least 100 participants per condition, for a total of 300 participants, with the intention of using this study to inform power analyses for subsequent studies. We excluded participants' data from analysis based on responses to attention check questions (see below), reaction time (less than three standard deviations below the overall mean after log transformation), and non-native English speaker status. The final sample included 328

participants (8.8% excluded; final sample 59% female and 41% male). We did not collect information on participant's ethnicity and race. All participants were United States residents.

### 2.1.2 Stimuli

We developed a novel set of 25 vignettes in which an agent acted intentionally to harm another individual. Five vignettes were included in each of five harm severity levels (property damage, theft, assault, maiming, and murder). We constructed three versions of each vignette. Each version was almost identical, with only one or two sentences manipulated across conditions. For example, one vignette (in the murder category) read as follows: "Chace knows the address of a woman who has highly offended him. As he had planned the day before, he waits there for the woman to return from work and, when she appears, Chace shoots her to death." The *Criminal Essentialism* condition—the main condition of interest—attributed the actor's behavior to a criminal essence. For example, the vignette above began as follows: "Part of Chace's biological make-up is a low degree of empathy for others and a predisposition to be a criminal." We compared this condition with two controls. First, the *General Essentialism* condition attributed the actor's behavior to an essence that was not criminal in nature. For example, the vignette above began as follows: "Part of Chace's biological make-up is a good memory, especially for phone numbers and addresses." The purpose of this condition was to determine whether essentialism regarding criminality specifically influences punishment, or whether essentialism regarding any characteristic relevant to the behavior would do. Second, the *No Essentialism* condition replaced the essentialist language from the previous two conditions with filler language that did not reference essentialist ideas. For example, the vignette above began as follows: "Chace has lived in the same neighborhood for a long time and is familiar with the area." The purpose of this condition was to determine how essentialist language influenced

punitiveness compared with non-essentialist language. See Supplemental Information for all vignettes.

### **2.1.3 Procedure**

After being randomly assigned to one of the three conditions, participants viewed all 25 vignettes in a random order. After reading each vignette, they answered the question “How much should [actor] be punished?” on a scale ranging from 1 (“no punishment at all”) to 10 (“extreme punishment”). Past research on the relation between moral condemnation and essentialism has used a variety of methods for condemnation, including asking participants to assign years in prison (Aspinwall et al., 2012; Berryessa, 2020; Cheung & Heine, 2015; Dar-Nimrod et al., 2011), rate agreement with punitive policies (de Vel-Palumbo et al., 2019; Harris & Socia, 2016; Imhoff, 2015; Kraus & Keltner, 2013), rate agreement with specific punishments (Kraus & Keltner, 2013), or to use an abstract scale (e.g. “How severe should punishment be?”) (Monterosso et al., 2005; Newman et al., 2015). Here, we chose an approach in line with these latter studies and with past research on punishment (Adams & Mullen, 2015; Carlsmith et al., 2002; Cushman, 2008; Gino et al., 2009; Treadway et al., 2014) for two reasons. First, because we use crimes of varying severity, asking about specific punishments is challenging—the specific punishment appropriate for murder is not appropriate for property damage. Second, our goal with this project is principally to understand the psychology underlying the relation between essentialism and people’s punitive sentiments, whereas many of the prior studies using specific punishments or years in prison sought to understand legal decision-making or the justice system. Although judgments about how much punishment people should receive and how many years they should serve in prison are related, our interest in moral cognition makes an abstract scale more appropriate.

Following all vignettes, participants completed a scale designed to assess individual differences in essentialism of criminality. Because the main purpose of Study 1 was to assess causal effects of essentialism on punitiveness, we do not discuss data based on individual differences here in depth.<sup>2</sup> Participants then completed attention check questions used to ensure data quality (e.g. describing the content of the previous vignette, self-assessing the amount of attention they paid to the study, etc.), demographic questions, and a debriefing. Data and analysis scripts for all studies are available at [https://osf.io/6m9xf/?view\\_only=0b5d32f2d77d4cd58fca5557a8e61623](https://osf.io/6m9xf/?view_only=0b5d32f2d77d4cd58fca5557a8e61623).

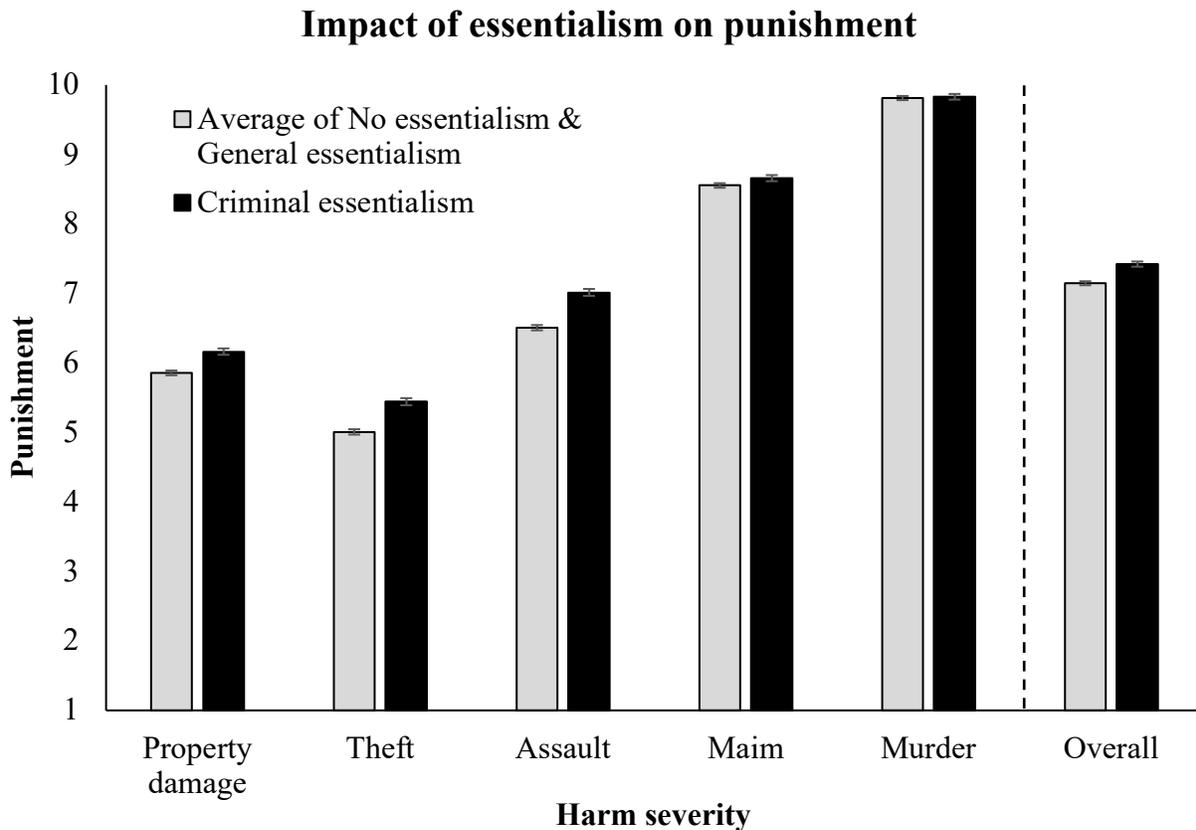
## 2.2 Results

Because participants completed a series of trials, we analyzed results using mixed-effects linear regression including random effects for participants as well as items, implemented in R using the package *lme4* (Bates et al., 2014). We obtained *p*-values for fixed effects using the Kenward-Roger approximation of degrees of freedom, implemented in *lmerTest* (Kuznetsova et al., 2015) and *pbrtest* (Halekoh & Højsgaard, 2014). Additionally, we included type of essentialism and severity as well as the interaction between them as fixed effects and a random slope for type of essentialism on vignette. Because we were primarily interested in the amount of punishment applied in the Criminal Essentialism condition relative to the other two conditions, with no prediction about relative punishment for the General Essentialism or No Essentialism conditions, we used Helmert contrasts to test our predictions. This allowed us to directly compare punishment in the Criminal Essentialism condition to punishment in the average of the General Essentialism and No Essentialism conditions.

---

<sup>2</sup> Responses on this scale (mean = 2.97, SD = 0.82, on 1-7 scale) did not correlate either with punishment across conditions ( $r_{\text{overall}} = 0.03, p = 0.55$ ) or with punishment within each of the conditions ( $r_{\text{Criminal Essentialism}} = 0.12, p = 0.22$ ;  $r_{\text{General Essentialism}} = -0.06, p = 0.54$ ;  $r_{\text{No Essentialism}} = 0.04, p = 0.64$ ).

The full model provided a better fit to the data than one without the interaction term (LRT  $X^2 [8] = 33.0, p < 0.001$ ) (see Supplemental Table 1 for full regression results). This full model revealed main effects of severity ( $F [4, 20.0] = 45.0, p < 0.001$ ) and type of essentialism ( $F [2, 307.11] = 3.41, p = 0.034$ ), which were qualified by a severity by type of essentialism interaction ( $F [8, 24.35] = 4.53, p = 0.002$ ). To interpret this interaction, we examined the fitted data from this model, looking specifically at punishment assigned for the Criminal Essentialism condition as well as the mean of the other two conditions, at each level of severity (Fig. 1). Here, we found little difference at higher severity levels but an increasingly large difference as severity decreased. Collapsing across severity, we observed more punishment for Criminal Essentialism than the average of the other conditions, consistent with the observed main effect of type of essentialism.



*Figure 1.* Results from Study 1. Participants read vignettes depicting a harm committed by a perpetrator against a victim. We manipulated whether the perpetrator’s behavior was attributed to a criminal essence (Criminal essentialism condition) or an essence that was not criminal in nature (General essentialism condition); a third condition did not describe any essence. Participants assigned higher punishment when behavior resulted from a criminal essence. Bars depict un-normalized values extracted from fitted models for ease of interpretation. Error bars are standard error of the mean.

### 2.3 Discussion

Study 1 probed the influence of essentialist views regarding criminality on punishment. Across three conditions, participants assigned more punishment when vignettes included essentialist descriptions of criminality rather than when vignettes included essentialist descriptions of non-criminal features or when vignettes did not include any essentialist

descriptions. This difference emerged for all five severity levels and was larger for less severe crimes, leading to an interaction between severity and essentialism condition.

The results of Study 1 add to past work on the relation between essentialism and punishment, which has found inconsistent results regarding whether essentialism increases (de Vel-Palumbo et al., 2019; Kraus & Keltner, 2013; Martin et al., forthcoming; Newman et al., 2015) or decreases (Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005) punitiveness. Consistent with this former group of studies, we find that essentialism of criminal behavior leads to greater punishment. We add to these findings by demonstrating this pattern across harms of varying severity, with an especially strong impact in cases of less severe harm. Studies 2-3 investigated the mechanisms underlying this influence of essentialism on punishment.

### **3. Study 2**

Study 1 found that essentialism of criminality leads to increased punishment using our novel set of vignettes. With these vignettes in hand, we next turned to the question of the mechanisms underlying the influence of essentialism on moral condemnation. In doing so, we began to resolve discrepancies between findings showing that essentialism increases punishment (de Vel-Palumbo et al., 2019; Kraus & Keltner, 2013; Martin et al., forthcoming; Newman et al., 2015), including Study 1, and other findings showing that essentialism reduces punishment (Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005).

Study 2 focused on the possibility that differences in perceptions of control explain these different patterns of results. Specifically, some prior work has essentialized behavior in

relatively strong (i.e. more deterministic or certain) terms, such as telling participants that “Joe had five times the average amount of a particular chemical in his brain” (Monterosso et al., 2005) and that an actor had a “fourfold increase in the likelihood of violence” (Cheung & Heine, 2015). These relatively more deterministic descriptions are “stronger” in the sense that they suggest that factors outside the actor’s control (e.g., biological make-up) influence the behavior in question to a relatively large extent. Indeed, conveying essentialism in strong terms has reduced both perceptions of the actor’s volition and moral condemnation (Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005). In contrast, studies that found positive associations between essentialism and punitiveness (de Vel-Palumbo et al., 2019; Kraus & Keltner, 2013; Newman et al., 2015) used more probabilistic language (i.e. “much more likely to”; “correctly guess ... significantly above chance”; Kraus & Keltner, 2015) or investigated the link between essentialism and punitiveness correlationally rather than experimentally, without measuring perceptions of control (de Vel-Palumbo et al., 2019; Martin et al., forthcoming).

Similarly, Study 1 did not measure perceptions of control and used relatively weaker descriptions of essentialism—describing it as a predisposition or bias, rather than a near certainty—which could have led to increased moral condemnation. In contrast, Study 2 manipulated the strength of the essentialism manipulation, including both a strong (i.e. more deterministic) and a weak (i.e. more probabilistic) condition. We asked participants to rate how much control the actor had over the outcome of the situation and to make a moral judgment of the actor. This design allowed us to test whether perceptions of control mediated the influence of this strength manipulation on moral condemnation. If strength of essentialism explains the difference between our results and some prior work, perceptions of control and moral

condemnation should both be lower in the strong essentialism condition, and differences in perceptions of control across conditions should mediate differences in moral condemnation.

We also extended the results of Study 1 by focusing on another moral judgment in addition to punitiveness: judgments of moral character. If essentialism only affects punishment judgments, we should not observe an influence on evaluations of character. In contrast, if essentialism affects moral judgment writ large, we would expect (1) an influence of our manipulation on judgments of character, and (2) a mediation of perceptions of control on ratings of character in the same manner as for punishment.

Finally, Study 2 addressed a methodological limitation of Study 1. In particular, in the Criminal Essentialism condition of Study 1, actors were sometimes described as having a history of criminal behavior, which could lead participants to judge them more harshly than those in the General Essentialism and No Essentialism conditions (even though actors in all three conditions committed the same crime presently). Therefore, in Study 2, we explicitly mentioned in both conditions for all cases that actors had previously committed three crimes so as to ensure that participants were not inferring a different number of previous criminal actions across conditions.

### **3.1 Methods**

#### **3.1.1 Participants**

Participants ( $n = 549$ ) were recruited using Amazon Mechanical Turk and received \$1.79. We determined and pre-registered sample size using a power calculation based on effect sizes from Study 1 (Cohen's  $d = 0.34$ , 80% power), which indicated that we needed 140 participants per condition. We excluded participants' data from analysis using the same criteria as in Study 1, with the additional requirement that they complete all trials. The final sample included 268 participants (49% excluded; final sample 50% female, 49% male, <1% other, and <1% did not

specify). The higher exclusion rate in Study 2 compared to Studies 1 and 3 could have arisen due to the relative length of this study and the requirement that participants complete all trials. After completing the main trials, participants completed a demographic questionnaire on which they self-identified as White or European-American (83%), Black or African-American (11%), Asian or Asian-American (4%), Native American/Pacific Islander (<1%), Multiracial (1%), and Other (<1%). Participants indicated their ethnicity separately from their race, and 5% self-identified as Hispanic or Latino/a. All participants were United States residents.

### 3.1.2 Stimuli

We constructed two modified versions of 15 of the vignettes used in Study 1, selecting three vignettes from each of the five severity levels. Specifically, we created two versions of each vignette. In the *Strong* version, we described the actors' predispositions toward criminality in more deterministic terms. As one example, the relevant portion of one vignette read, "Because of Chace's biological make-up, it is near impossible for him to not lash out violently and emotionally when offended, even to the point of committing criminal behaviors." In the *Weak* version, the influence of this predisposition was more probabilistic in nature: "Because of Chace's biological make-up, he sometimes lashes out violently and emotionally when offended, even to the point of committing criminal behaviors." See Supplemental Information for all vignettes.

### 3.1.3 Procedure

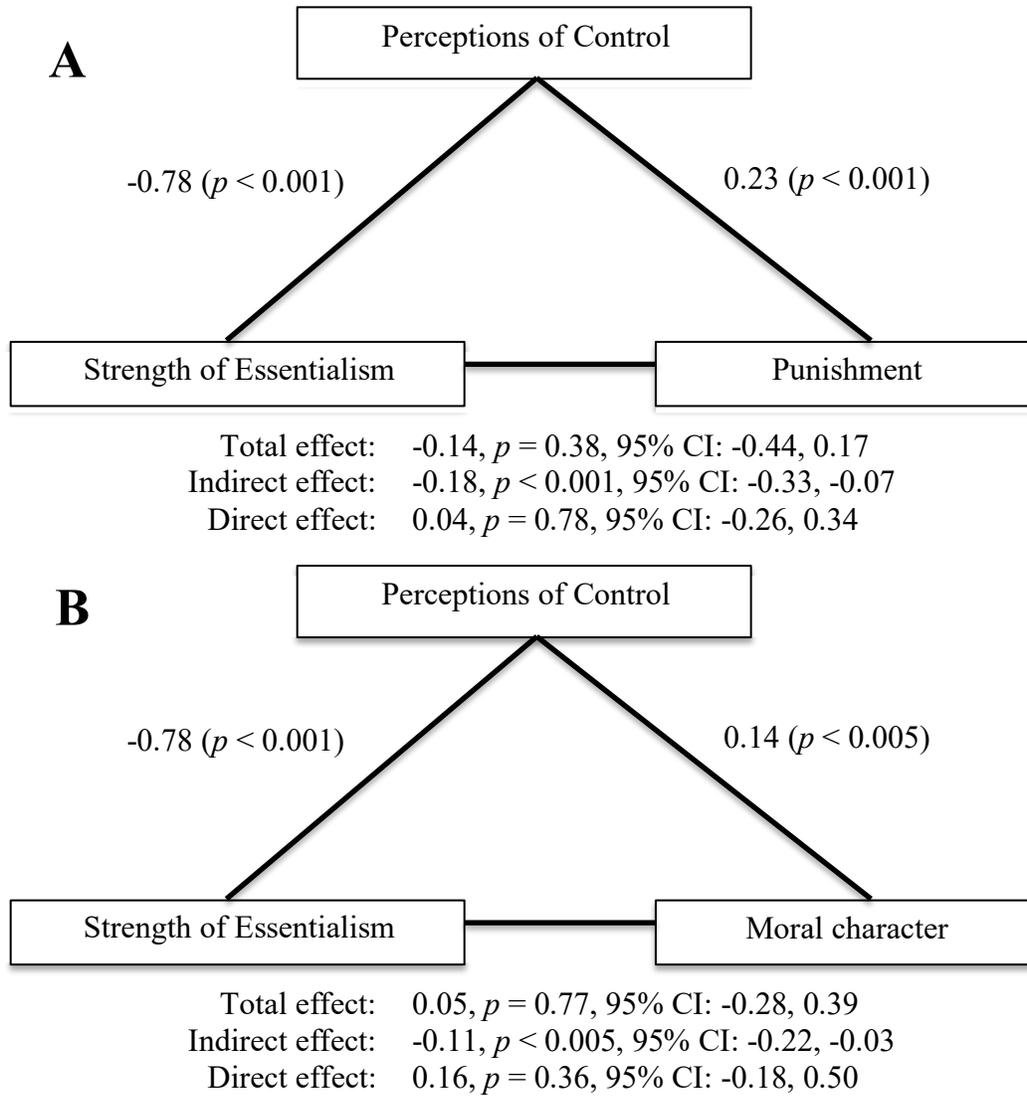
Participants were randomly assigned to either the Strong or Weak essentialism condition. They then viewed all 15 vignettes in a random order and answered the question, "How much control did [actor] have over the outcome of the situation?" Response options ranged from 1 = "No control at all" to 9 = "Complete control." Participants then read through all 15 vignettes

again, in a random order, rating punishment as described in Studies 1 or rating character with the prompt, “How would you rate [actor]’s personal moral character?” with the anchors 1 = “Horrible moral character” and 9 = “Great moral character.” Finally, participants read through all 15 vignettes another time, rating the moral judgment not yet made (either punishment or character). The order of judgments about punishment versus character was randomized across participants. After completing all of these experimental items, participants answered the same attention check and demographic questions as in Studies 1 and received a debriefing. Pre-registration of sample size and analytic approach can be found here:

<http://aspredicted.org/blind.php?x=ct3wy9>.

### **3.2 Results**

To test the hypothesis that perceptions of control mediate the relation between strength of essentialism and moral judgments, we conducted two mediation analyses—one with punishment judgments and the other with moral character judgments—using 5,000 bootstrapped samples (Hayes, 2013). Strength of essentialism served as the predictor variable, ratings of control served as the mediator, and moral judgment (either punishment or moral character) served as the dependent measure. Results are presented in Fig. 2. To summarize, we found a significant indirect effect of strength of essentialism on both punishment and character through perceptions of control, consistent with some past work on the relation between perceived control and moral condemnation (Cheung & Heine, 2015). In line with Montoya and Hayes (2017), we infer mediation from this significant effect. Using strong essentialism manipulations reduced perceptions that actors had control over the results of their behavior, which, in turn, reduced punitiveness and negative perceptions of moral character.



*Figure 2. Mediation models tested in Study 2.  
 A: Mediation of punishment judgments. B: Mediation of moral character judgments.*

### 3.3 Discussion

Study 2 investigated the influence of a stronger versus weaker manipulation of essentialism on moral condemnation. Past work has found that strong versions of essentialism reduced both perceptions of actors' volition and moral condemnation (Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005). In contrast, when essentialism was probabilistic and described a general tendency but not predetermination, moral condemnation was exacerbated—as in Study 1 and other prior work (de Vel-Palumbo et al., 2019; Kraus & Keltner, 2013; Martin et al., forthcoming; Newman et al., 2015). Thus, whether essentialism impacts perceptions of actors' volition should influence moral condemnation.

The current results support this idea. Describing essentialism in stronger terms reduced perceptions of an actor having control over the outcome of a situation. In turn, this led to reduced punishment and less harsh assessments of the actor's moral character. The results of Study 2 suggest that essentialism can lead to increased moral condemnation when manipulations of essentialism are probabilistic in nature. When essentialism is “strong” and suggests that actors' behaviors are pre-determined (reducing perceptions that they control the outcomes of their behaviors), moral condemnation is decreased—as in Study 2 and some prior work (Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005). Similar results emerged when participants made judgments regarding punishment and moral character, suggesting that the relation between essentialism of criminality and punishment shown in Study 1 generalizes beyond punishment to other moral judgments.

#### 4. Study 3

Study 2 found that perceptions of control mediate the influence of essentialism on moral condemnation, providing one possible explanation for observed differences in the relation between essentialism and moral condemnation. Of course, the identification of one explanation does not preclude there being additional explanations. Study 3 probed just such a complementary explanation, motivated by the observation that in past work, essentialism tended to decrease moral condemnation when it targeted particular behaviors (Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005). Aspinwall and colleagues (2012) focused on battery, Cheung and Heine (2015) included a case of murder, Dar-Nimrod and colleagues (2011) focused on specific sexual crimes, and Monterosso and colleagues (2005) described actors who were likely to commit “extreme *acts* of violence” (emphasis ours). In contrast, essentialist descriptions that increase moral condemnation have targeted broader, trait-level characteristics, like social class (Kraus & Keltner, 2013) or the entirety of a person’s character (i.e. their “true self”; Newman et al., 2015). Similarly, de Vel-Palumbo and colleagues (2019) asked participants to rate endorsement of restrictive policies for sexual offenses in general, targeting the broad category of offenses rather than a specific instance of behavior. Finally, Imhoff and Jahnke (2018) and Harris and Socia (2016) both focused on general terms related to sexual offenses and found that these labels (i.e. “pedophilia”, “sexual offenders”) led to harsher punitive attitudes relative to generic language.

Essentialism targeted at behaviors may reduce perceptions of control, as described in Study 2, and therefore reduce moral condemnation. However, essentialism targeted at general characteristics may not have this effect because participants may not generalize inferences about an individual’s character to behaviors they perform. That is, when told someone has an

unchangingly bad character, participants may not take the additional step of then viewing this individual's behaviors as outside their control. Indeed, such a reticence is consistent with studies demonstrating that people are resistant to the idea that individuals lack control over their behaviors (Nahmias et al., 2005; Nichols & Knobe, 2007). Instead, participants may focus on the badness of the characteristic (e.g. "This person has a deeply bad character") and morally condemn those whom they view as bad individuals. Thus, essentialism regarding specific *behaviors* may have a different influence on punishment than essentialism regarding morally relevant *characteristics*.

Simultaneously, we investigated another potential explanation that relies on the way essentialism is conveyed. In particular, we were motivated by the idea that essentialism contains multiple components (Cimpian & Salomon, 2014; Medin & Ortony, 1989). Two components of chief importance are *biology* (e.g. essential traits are biological in nature) and *immutability* (e.g. essential traits are unchangeable; Gelman et al., 2007; Heiphetz, Gelman, et al., 2017). Past work showing that essentialism reduces moral condemnation has tended to focus on the biology component. Aspinwell and colleagues (2012) focused on a specific neurobiological explanation of psychopathy involving enzymes and neurotransmitters, Cheung and Heine (2015) described a genetic predisposition toward violence, Dar-Nimrod and colleagues (2011) highlighted men's biological predispositions to commit sexual crimes, and Monterosso and colleagues (2005) described the actor as having "five times the amount of a particular chemical" in his brain. In contrast, studies finding that essentialism increases moral condemnation (de Vel-Palumbo et al., 2019; Kraus & Keltner, 2013; Martin et al., forthcoming; Newman et al., 2015) did not highlight the biology component of essentialism. In contrast, the relation between essentialist descriptions that principally focus on immutability and moral condemnation is unclear. Past work has found

that neither the immutability nor biology components of essentialism predicts recommended years in prison (Berryessa, 2020) and that a manipulation of whether pedophilia is described as controllable (i.e. mutable) or did not find an influence on punishment (Imhoff & Jahnke, 2018). From this perspective, essentialist descriptions that principally focus on biology may reduce moral condemnation at least in part through a reduction in perceptions of actors' control over their behavior, whereas those focusing on immutability may have little to no impact. Thus, Study 3 contrasted the biological descriptions used in some prior work with descriptions of immutability.

## **4.1 Methods**

### **4.1.1 Participants**

Participants ( $n = 800$ ) were recruited using Amazon Mechanical Turk and received \$1.21. We determined sample size using the method outlined in Study 2 and excluded data from analysis using the same criteria as in Studies 1 and 2. The final sample included 732 participants (8.5% excluded; final sample 53% female and 47% male). At the end of the experimental session, participants completed a demographic questionnaire on which they self-identified as White or European-American (76%), Black or African-American (9%), Asian or Asian-American (5%), Native American/Pacific Islander (<1%), Multiracial (6%), and Other (2%). Participants indicated their ethnicity separately from their race, and 5% self-identified as Hispanic or Latino/a. All participants were United States residents.

### **4.1.2 Stimuli**

We constructed four modified versions of 15 of the vignettes used in Studies 1 and 2, selecting three vignettes from each of the five severity levels. Each vignette was almost identical, with only one or two sentences manipulated across conditions. We manipulated two

factors: The *component* of essentialism invoked (biology versus immutability) and the *target* of essentialism (a particular behavior versus the entirety of the actor's character). For example, in the *immutable component, character target* condition, an example vignette began: "For his whole life, Chace has had a tendency toward violent and emotional reactions and a predisposition toward criminality in his character." In the *immutable component, behavior target* condition, the same vignette began: "For his whole life, Chace has lashed out violently and emotionally when offended, even to the point of committing criminal behaviors." In the *biological component, character target* condition, the beginning of this vignette read as follows: "Part of Chace's biological make-up is a tendency toward violent and emotional reactions and a predisposition in his character toward criminality." Finally, in the *biological component, behavior target* condition, this vignette started as follows: "Because of Chace's biological make-up, he often lashes out violently and emotionally when offended, even to the point of committing criminal behaviors." In addition, due to the fact that vignettes highlighting the immutable component sometimes specified a history of criminal behavior, we explicitly mentioned in all conditions for all cases that the actor had previously committed three crimes, so as to ensure that participants were not inferring a different number of previous criminal actions across conditions. See Supplemental Information for all vignettes.

### **4.1.3 Procedure**

Participants viewed all 15 vignettes in a random order. We employed a 2 x 2 between-subjects design. Participants were randomly assigned to read all vignettes in one of the four combinations of our two factors (component of essentialism crossed with target of essentialism). After reading each vignette, they rated punishment using the same scale as in Studies 1 and 2. Participants then completed the same attention check and demographic questions as in Studies 1

and 2 and received a debriefing. Pre-registration of sample size and analytic approach can be found here: <http://aspredicted.org/blind.php?x=cy2n8v>.

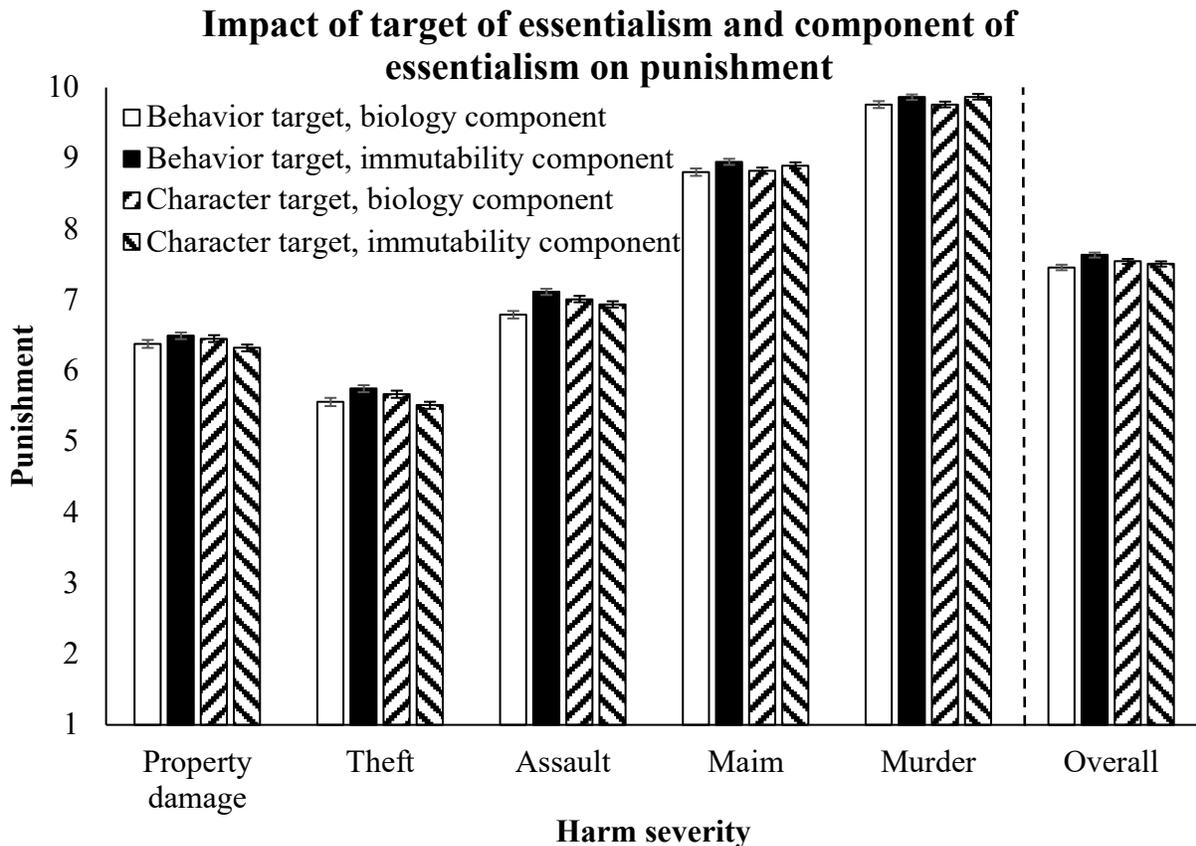
## 4.2 Results

As in Study 1, we analyzed data using mixed-effects linear regression including random effects for participants as well as items, implemented in R using the package lme4 (Bates et al., 2014). Additionally, we obtained *p*-values for fixed effects using the Kenward-Roger approximation of degrees of freedom, implemented in lmerTest (Kuznetsova et al., 2015) and pbkrtest (Halekoh & Højsgaard, 2014). We included random intercepts for participants as well as items.

Taking a model comparison approach, we began by testing the influence of our two factors of interest, comparing a model with an interaction between component of essentialism and target of essentialism and main effects of both factors, to a model with just main effects of both factors. There was no significant interaction (LRT  $\chi^2 [1] = 1.78, p = 0.18$ ) and no significant main effect of either component of essentialism (LRT  $\chi^2 [1] = 0.74, p = 0.41$ ) or target of essentialism (LRT  $\chi^2 [1] = 0.08, p = 0.79$ ) (see Supplemental Table 2 for full regression results).

Next, given the importance of harm severity in Study 1, we ran exploratory analyses testing for an interaction between harm severity and our two main factors. Specifically, we tested a model including fixed effect predictors for component of essentialism, target of essentialism, and harm severity, as well as all possible interactions between these variables, against one without the three-way interaction. We found a significant three-way interaction (LRT  $\chi^2 [4] = 11.4, p = 0.02$ ). To unpack this interaction further, we looked at the interaction of component of essentialism and harm severity at each level of target of essentialism individually

(Fig. 3). First, looking at cases in which essentialism emphasized the actor's moral character, we found a significant interaction between component of essentialism and harm severity (LRT  $\chi^2$  [4] = 10.59,  $p = 0.03$ ). We found no such interaction when looking at cases in which essentialism emphasized the immutability of the actor's behavior (LRT  $\chi^2$  [4] = 6.31,  $p = 0.16$ ). Looking at mean values for cases in which essentialism emphasized the actor's character, punishment was higher when the vignettes emphasized the immutability component of essentialism for more severe harms and the biology component for less severe harms. In other words, when essentialist descriptions emphasized an actor's character, descriptions of immutability accentuated the role of harm severity—with harsher punishment for more severe crimes and weaker punishment for less severe crimes—relative to the amount of punishment assigned when essentialist descriptions emphasized biology. In contrast, when essentialist descriptions emphasized an actor's behavior, we did not find an influence of biological descriptions versus immutable descriptions on punishment.



*Figure 3.* Results from Study 3. Participants read vignettes depicting a harm committed by a perpetrator against a victim. We manipulated the descriptions of essentialism in two ways: By changing the target of essentialism (a particular behavior versus the protagonist’s character) and the component of essentialism (biology versus immutability). We found a 3-way interaction between these two manipulations and harm severity. When essentialism targeted the protagonist’s character, punishment was higher when vignettes highlighted the immutability component of essentialism for more severe harms and the biology component for less severe harms. In contrast, when essentialist descriptions emphasized a protagonist’s behavior, an emphasis on biology versus immutability did not influence punishment. Bars depict un-normalized values extracted from fitted models for ease of interpretation. Error bars are standard error of the mean.

### 4.3 Discussion

The main goal of Study 3 was to investigate whether the target of essentialism (behavior versus character) or the component of essentialism (biology versus immutability) influenced moral condemnation. Past work using biological explanations of essentialism has largely found

an attenuation of moral condemnation (Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005). However, essentialism can also be conveyed using descriptions of immutability (Gelman et al., 2007; Heiphetz et al., 2017). In addition, past work demonstrating an attenuation of moral condemnation in response to essentialism has tended to target essentialism of behaviors (Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005). In cases where essentialist descriptions increased moral condemnation, these descriptions targeted trait-level characteristics, like social class (Kraus & Keltner, 2013), the entirety of a person's character (Newman et al., 2015) or a general category of crime rather than a specific instance of it (de Vel-Palumbo et al., 2019).

We did not find a significant overall effect of whether essentialist descriptions emphasized biology or immutability, consistent with some past work (Berryessa, 2020), nor did we find a main effect of whether such descriptions emphasized behaviors or character. Combined with the results of Study 2, these findings suggest that the main factor influencing whether essentialism attenuates or accentuates moral condemnation is the degree to which essentialism impacts perceptions of an actors' volition, not the extent to which essentialist descriptions emphasize biology versus immutability or the extent to which they emphasize behaviors versus character.

Given the importance of harm severity in Study 1, we ran exploratory analyses including harm severity as a predictor. We did find a weak three-way interaction between whether essentialist descriptions emphasized biology or immutability, whether such descriptions emphasized behaviors or character, and harm severity. This manifested mainly for cases in which essentialist descriptions emphasized the actor's moral character. For these cases, descriptions of immutability (relative to biology-based descriptions) increased punitiveness for

more severe crimes and decreased punitiveness for less severe crimes. In other words, when essentialism targeted the actor's moral character, descriptions of immutability enhanced the influence of harm severity, leading to harsher punishment for more severe crimes and weaker punishment for less severe crimes. We did not find the same pattern when looking at cases in which essentialism targeted specific behaviors. These results potentially suggest that when someone's character is described as unlikely to change (versus rooted in a biological predisposition), participants take the harmful behaviors they perform as evidence of their moral culpability—punishing more when the harm is more severe and less when the harm is less severe. However, we note that these results should be interpreted with caution given their exploratory and weak nature.

## 5. General Discussion

Philosophers (Kant, 1790/1952) and psychologists (Tannenbaum et al., 2011; Uhlmann et al., 2015; Uhlmann & Zhu, 2014) have foregrounded the link between moral condemnation and “internal wickedness.” Recent work has demonstrated the potency of this link. However, this work has provided conflicting results about whether focusing on internal characteristics increases punitiveness (de Vel-Palumbo et al., 2019; Harris & Socia, 2016; Imhoff & Jahnke, 2018; Kraus & Keltner, 2013; Martin et al., forthcoming; Newman et al., 2015) or reduces it (Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005). The main goal of the current work was to help resolve this discrepancy.

We began by asking *how* essentialism influences punitiveness. To answer this question, Study 1 presented participants with information about people who had committed crimes and manipulated whether descriptions of the actor highlighted his criminal essence, his non-criminal

essence, or no essence. Relative to the latter two conditions, those described as having a criminal essence received more punishment.

Studies 2-3 investigated why essentialism might have increased punitiveness in Study 1. In doing so, these studies also clarified factors that may explain divergent patterns of results in prior work. Study 2 showed that stronger manipulations of essentialism reduced perceptions of actors' control over the outcome of their behavior, which in turn reduced moral condemnation. These results suggest that when essentialism is "strong" and implies that actors' behaviors are pre-determined (reducing perceptions that they control the outcomes of their behaviors), moral condemnation is reduced, consistent with some prior work (Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005). In contrast, when manipulations of essentialism are probabilistic in nature, they can lead to increased moral condemnation. This was true for judgments of both punishment and moral character, generalizing the results beyond punishment to other types of moral evaluation.

We next investigated whether perceptions of control were the only explanation for these past discrepancies or whether complementary explanations existed. Specifically, Study 3 examined the influence of the component of essentialism emphasized (biology or immutability) and the target of essentialism (behavior or character). We did not find a significant effect of either variable, though an interaction did emerge between these variables and harm severity. This analysis revealed a complex relation: when essentialism targeted the actor's moral character, descriptions of immutability enhanced the influence of harm severity, leading to harsher punishment for more severe crimes but weaker punishment for less severe crimes. We did not find the same pattern when looking at cases in which essentialism targeted specific behaviors. However, these analyses were exploratory and should be interpreted with caution.

The overarching pattern across Studies 2 and 3 indicates that perceptions of control serve as a mechanism underlying the relation between essentialism and punitiveness, whereas the evidence suggesting that component or target of essentialism may exert an influence is far weaker. Thus, we view differences in the influence of essentialism on perceptions of control as a better explanation for past discrepancies than differences in the component of essentialism targeted.

While our results demonstrate a link between essentialism and punitiveness, essentialism is of course not the only thing that is related to punishment. Other factors include the place that the person who has transgressed occupies in the social hierarchy—for instance, the United States legal system often punishes Black people and poor people more severely than White people and rich people for the same crimes (e.g., Alexander, 2012; Pager, 2007; Western, 2018). Additional individual difference measures include a belief in pure evil, which also predicts punitiveness (Webster & Saucier, 2013). The work presented here illuminates the psychology underlying punishment by investigating another influence on punitiveness—namely, essentialism. Doing so increases theoretical understanding of punishment by highlighting additional factors that influence punitiveness and also increases theoretical understanding of essentialism by adding to the growing literature on how it connects to punitiveness (Berryessa, 2019, 2020; Cohen, 2004; Willis, 2018). These results also have translational importance by suggesting that reducing essentialism may reduce punitiveness. In this way, the current results contribute in a unique way to the broader literature on why people punish.

Additionally, these findings have implications for the psychology underlying intergroup conflict. In particular, the present studies demonstrate that essentialist views of criminality can undergird perceptions of people who have committed crimes. Such perceptions may contribute to the stigma people face post-incarceration—stigma that makes it difficult for people returning

from prison to secure employment, find housing, and re-establish family ties (Alexander, 2012; Clear, 2007; Pager, 2007; Travis & Maul, 2003; Western, 2018). Viewing these individuals as having an unchangingly bad essence may mean that bias in this domain is particularly unyielding; it is difficult to trust someone enough to offer them a job or form intimate bonds with them if you view being a criminal as one of their central characteristics. These processes, in turn, can increase re-incarceration. The lack of secure housing or employment can be considered a parole violation, and even if people leave prison at the end of their sentence (i.e. not as parolees), it may be tempting to turn to illegal methods of securing income if legal means are not available (Alexander, 2012; Clear, 2007; Travis, 2005; Western, 2018). The current results suggest that reducing essentialism could serve as one potential intervention, a hypothesis that could be tested in future work.

The present work also informs the broader literature on punishment. Past work has demonstrated that laypeople punish individuals less for behaviors they cannot control than for behaviors that are under their control (Alicke, 2000; Darley et al., 2000; Martin & Cushman, 2016b; Robinson & Darley, 1995). These results are consistent with a role for punishment in changing behavior: If the behavior cannot be controlled now, it likely cannot be controlled in the future, and therefore punishment cannot serve its function of changing behavior (Martin & Cushman, 2016a). Indeed, consistent with some past work (Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005) and as others have also hypothesized (Ryazanov & Christenfeld, 2018), we find results that support this idea: when essentialism reduces perceptions of an actor's volition, people are less likely to endorse harsh punishments.

Beyond punishment, some have suggested that the relation between essentialism and control can have strategic value (Ryazanov & Christenfeld, 2018). According to this logic, members of marginalized groups may strategically use or avoid essentialist language or framing depending upon present goals. In one study, women who received more messages of acceptance about their weight were paradoxically more likely to lose weight (Logel et al., 2014), potentially indicating that essentializing weight (and the role of genetics in weight) can be beneficial. In another example, female engineering students were exposed to either a value-affirming intervention that emphasized social-belonging or an affirmation-training intervention that focused on self-identity (Walton et al., 2015). While both interventions increased GPA, the affirmation-training also increased identification with gender, which can be interpreted as increased essentialization of gender to the extent that such identification reflects greater belief in the meaningfulness or entitativity of gender (Ryazanov & Christenfeld, 2018). Thus, increasing essentialism in this case may have played a role in improving these students' GPA. Focusing on the criminal justice context, our results show that essentialism has a nuanced relationship with moral condemnation: in Study 2 we find that "stronger" essentialism reduces punishment relative to "weaker" essentialism. More broadly, this argument suggests that being strategic about accepting or rejecting essentialism depending upon present goals may aid marginalized or stigmatized groups (Ryazanov & Christenfeld, 2018), including those who have had contact with the legal system.

Relatedly, our results have implications for future research on essentialism in the context of the legal system. To the extent that essentialism accounts for some of the stigma faced by those involved in the legal system (Alexander, 2012; Pager, 2007; Travis & Maul, 2003), our results provide insight into potential origins of that stigma. Consistent with some past work

(Aspinwall et al., 2012; Cheung & Heine, 2015; Dar-Nimrod et al., 2011; Monterosso et al., 2005), we show that only when essentialism of criminality is strong does it reduce punishment. Essentialism leading to stigma faced by those involved in the legal system may be relatively weak (i.e. people may not believe that it is “near impossible for [these individuals] to not lash out violently and emotionally when offended”) and therefore may increase moral condemnation. Indeed, some prior qualitative work has found that essentialist biases toward mentally ill people who have committed crimes leads some judges toward harsher sentencing (Berryessa, 2019). To the extent that others in the legal system—whether correlational officers, probation officers, or police officers—similarly hold essentialist views, this may impact their moral assessment and treatment of those who have been involved in the legal system. Consistent with recommendations to be strategic about choosing when to challenge essentialist views (Ryazanov & Christenfeld, 2018), our results may suggest that society challenging such essentialist thinking would be beneficial. Of course, these implications should be interpreted with caution: our principal focus is on moral cognition and not legal decision-making. Indeed, participants in our studies made an abstract punishment decision rather than a judgment about guilt or prison time. Nevertheless, our results may provide some insight into the relation between essentialism and punishment, especially when such essentialism is likely to increase versus reduce punishment.

Our results also relate to prior work on the relation between belief in determinism and punishment (Clark et al., 2014; Kesberg & Keller, 2021; Shariff et al., 2014). These studies have found that greater belief in free will (and reduced belief in determinism) predicts increased punishment (Shariff et al., 2014) and that being exposed to an immoral action increases belief in free will and motivation to punish (Clark et al., 2014). Thus, there is a tight coupling between believing others have volition over their actions and punishing them. To the extent that belief in

determinism has been equated with essentialism in some past work (Kesberg & Keller, 2021), our results largely converge with this literature: when perceptions of perpetrators' control over their behavior are reduced (i.e. their free will is reduced), punishment is similarly reduced. However, belief in determinism seems conceptually most closely related to belief in genetic essentialism (Kesberg & Keller, 2021). Our manipulation of control in Study 2, through strength of essentialism, did sometimes involve a description of genetics, though this was not always the case. Furthermore, Study 3 shows that the biology component of essentialism has at best only a weak influence on punishment, consistent with some past work (Berryessa, 2020). Of course, biology and genetics are not synonymous (e.g. our biology component manipulation did not always explicitly mention genetics) and so this connection is only suggestive. In total, our results provide support for the idea that stronger manipulations of essentialism, which often but not always invoke genetics, can reduce punishment, in a way similar to how believing in determinism can undermine punishment. Whether this is accomplished most strongly by an influence through genetics or some other component of biology is a topic for future work.

Another open question is the etiology of holding these harsher views of people who have committed crimes—do they reflect natural variability in the tendency to essentialize criminality, or are there more meaningful differences between those who believe others are generally good and those who are willing to view others as essentially criminal? Further, while we have focused on the tendency of individuals who view criminality in an essentialist way to exhibit punitiveness, we have not queried their views about morally good behavior. Perhaps these individuals just hold a dimmer view of humanity in general (c.f. De Freitas, Sarkissian, et al., 2017). But they may be more likely to essentialize all forms of morally relevant behavior, whether good or bad. On this latter account, we would expect them to view good acts as

stemming from an essentially good character and to be more likely to reward or praise such individuals. Future work is needed in exploring these questions.

An additional question raised by our results is the relation between severity of harm and essentialism of criminal behavior. In Studies 1 and 3, harm severity played some role in participants' judgments. In Study 1, essentialism of criminality increased moral condemnation to a greater extent for less severe crimes, and Study 3 showed a complex interaction between harm severity, component of essentialism, and target of essentialism. Interestingly, some previous work has found that manipulations of harm severity do not influence the degree to which participants view a behavior as stemming from an essential characteristic or to which these participants make punitive policy recommendations (de Vel-Palumbo et al., 2019). One critical difference between those results and our own is that those vignettes involved sexual crimes, which were not included here. Participants' judgments may be more sensitive to severity when they evaluate crimes that they perceive to be especially deviant, whereas judgments of criminal behavior that is more instrumental (e.g., theft or property damage) might show less sensitivity to severity. Our data partially support this possibility; in our studies, essentialism played a smaller role in responses to crimes that participants could have perceived as particularly deviant, such as maiming and murdering. Future work should further probe the relation between harm severity and essentialism of criminal behavior.

Building on prior work showing evidence of essentialism in a number of domains (Dar-Nimrod & Heine, 2011), the current studies investigated the extent to which adults essentialize criminality and highlighted the predictive and causal power of essentialism in this domain. In an era of mass incarceration, these theoretical advances have far-reaching societal implications

through their ability to clarify people's perceptions of people involved in the criminal justice system.

### References

- Adams, G. S., & Mullen, E. (2015). Punishing the Perpetrator Decreases Compensation for Victims. *Social Psychological and Personality Science*, 6(1), 31–38.  
<https://doi.org/10.1177/1948550614542346>
- Alexander, M. (2012). *The new Jim Crow: Mass incarceration in the age of colorblindness*. The New Press.
- Alicke, M. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, 126(4), 556–574. <https://doi.org/10.1037/0033-2909.126.4.556>
- Aspinwall, L. G., Brown, T. R., & Tabery, J. (2012). The double-edged sword: Does biomechanism increase or decrease judges' sentencing of psychopaths? *Science*, 337(6096), 846–849. <https://doi.org/10.1126/science.1219569>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7.
- Berryessa, C. M. (2019). Judicial Stereotyping Associated with Genetic Essentialist Biases toward Mental Disorders and Potential Negative Effects on Sentencing. *Law & Society Review*, 53(1), 202–238. <https://doi.org/10.1111/lasr.12382>
- Berryessa, C. M. (2020). The effects of essentialist thinking toward biosocial risk factors for criminality and types of offending on lay punishment support. *Behavioral Sciences & The Law*, 38(4), 355–380. <https://doi.org/10.1002/bsl.2476>
- Brescoll, V., & LaFrance, M. (2004). The correlates and consequences of newspaper reports of research on sex differences. *Psychological Science*, 15(8), 515–520.  
<https://doi.org/10.1111/j.0956-7976.2004.00712.x>
- Carlsmith, K., Darley, J., & Robinson, P. (2002). Why do we punish?: Deterrence and just

deserts as motives for punishment. *Journal of Personality and Social Psychology*, 83(2), 284–299. <https://doi.org/10.1037/0022-3514.83.2.284>

Cheung, B. Y., & Heine, S. J. (2015). The Double-Edged Sword of Genetic Accounts of Criminality: Causal Attributions From Genetic Ascriptions Affect Legal Decision Making. *Personality and Social Psychology Bulletin*, 41(12), 1723–1738. <https://doi.org/10.1177/0146167215610520>

Cimpian, A., & Salomon, E. (2014). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences*, 37(05), 461–480. <https://doi.org/10.1017/S0140525X13002197>

Clark, C. J., Luguri, J. B., Ditto, P. H., Knobe, J., Shariff, A. F., & Baumeister, R. F. (2014). Free to punish: A motivated account of free will belief. *Journal of Personality and Social Psychology*, 106(4), 501–513. <https://doi.org/10.1037/a0035880>

Clear, T. R. (2007). *Imprisoning communities: How mass incarceration makes disadvantaged neighborhoods worse*. Oxford University Press.

Clutton-Brock, T., & Parker, G. (1995). Punishment in animal societies. *Nature*, 373, 209–216. [http://www.anth.ucsb.edu/faculty/gurven/anth169/punishment\\_in\\_animal\\_societies.pdf](http://www.anth.ucsb.edu/faculty/gurven/anth169/punishment_in_animal_societies.pdf)

Cohen, M. A. (2004). The Costs of Crime and Justice. In *The Costs of Crime and Justice* (Issue August). Routledge. <https://doi.org/10.4324/9780203313145>

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108(2), 353–380. <https://doi.org/10.1016/j.cognition.2008.03.006>

Cushman, F., Dreber, A., Wang, Y., & Costa, J. (2009). Accidental outcomes guide punishment in a “trembling hand” game. *PloS One*, 4(8), e6699.

<https://doi.org/10.1371/journal.pone.0006699>

Dar-Nimrod, I., & Heine, S. J. (2011). Genetic essentialism: On the deceptive determinism of DNA. *Psychological Bulletin*, *137*(5), 800–818. <https://doi.org/10.1037/a0021860>

Dar-Nimrod, I., Heine, S. J., Cheung, B. Y., & Schaller, M. (2011). Do scientific theories affect men's evaluations of sex crimes? *Aggressive Behavior*, *37*(5), 440–449.

<https://doi.org/10.1002/ab.20401>

Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2000). Incapacitation and just deserts as motives for punishment. *Law and Human Behavior*, *24*(6), 659–683.

<https://doi.org/10.1023/A:1005552203727>

De Freitas, J., Sarkissian, H., Newman, G. E., Grossmann, I., De Brigard, F., Luco, A., & Knobe, J. (2018). Consistent Belief in a Good True Self in Misanthropes and Three Interdependent Cultures. *Cognitive Science*, *42*(9), 134–160. <https://doi.org/10.1111/cogs.12505>

de Vel-Palumbo, M., Howarth, L., & Brewer, M. B. (2019). 'Once a sex offender always a sex offender'? Essentialism and attitudes towards criminal justice policy. *Psychology, Crime & Law*, *25*(5), 421–439. <https://doi.org/10.1080/1068316X.2018.1529234>

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*(6), 876–893.

<https://doi.org/10.1037/0022-3514.87.6.876>

Fuss, J., Dressing, H., & Briken, P. (2015). Neurogenetic evidence in the courtroom: A randomised controlled trial with German judges. *Journal of Medical Genetics*, *52*(11), 730–737. <https://doi.org/10.1136/jmedgenet-2015-103284>

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford University Press.

- Gelman, S. A., Heyman, G. D., & Legare, C. H. (2007). Beliefs about developmental changes in the coherence of essentialist characteristics psychological. *Child Development, 78*(3), 757–774. <https://doi.org/10.1111/j.1467-8624.2007.01031.x>
- Gino, F., Moore, D., & Bazerman, M. (2009). No harm, no foul: The outcome bias in ethical judgments. *Harvard Business School NOM Working Paper No. 08-080*.
- Goodwin, G. P. (2015). Moral Character in Person Perception. *Current Directions in Psychological Science, 24*(1), 38–44. <https://doi.org/10.1177/0963721414550709>
- Halekoh, U., & Højsgaard, S. (2014). *A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed Models —The R package pbkrtest* (pp. 1–32). *Journal of Statistical Software*. <https://doi.org/10.18637/jss.v059.i09>
- Harris, A. J., & Socia, K. M. (2016). What’s in a Name? Evaluating the Effects of the “Sex Offender” Label on Public Opinions and Beliefs. *Sexual Abuse, 28*(7), 660–678. <https://doi.org/10.1177/1079063214564391>
- Haslam, N., & Levy, S. R. (2006). Essentialist beliefs about homosexuality: Structure and implications for prejudice. *Personality and Social Psychology Bulletin, 32*(4), 471–485. <https://doi.org/10.1177/0146167205276516>
- Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology, 39*(1), 113–127. <https://doi.org/10.1348/014466600164363>
- Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology, 41*(1), 87–100. <https://doi.org/10.1348/014466602165072>
- Hauser, M. D., Young, L., & Cushman, F. (2006). *Reviving Rawls’s Linguistic Analogy :*

*Operative Principles and the Causal Structure of Moral Actions.*

- Hayes, A. (2013). *Introduction to mediation, moderation, and conditional process analysis. A regression-based approach.* Guilford Press.
- Hebble, P. W. (1971). The Development of Elementary School Children's Judgment of Intent. *Child Development, 42*(4), 1203. <https://doi.org/10.2307/1127804>
- Heiphetz, L., Gelman, S. A., & Young, L. (2017). The perceived stability and biological basis of religious beliefs, factual beliefs, and opinions. *Journal of Experimental Child Psychology, 156*, 82–98. <https://doi.org/10.1016/j.jecp.2016.11.015>
- Heiphetz, L., & Young, L. (2014). A social cognitive developmental perspective on moral judgment. *Behaviour, 151*(2–3), 315–335. <https://doi.org/10.1163/1568539X-00003131>
- Imhoff, R. (2015). Punitive Attitudes Against Pedophiles or Persons With Sexual Interest in Children: Does the Label Matter? *Archives of Sexual Behavior, 44*(1), 35–44. <https://doi.org/10.1007/s10508-014-0439-3>
- Imhoff, R., & Jahnke, S. (2018). Determinants of Punitive Attitudes Toward People with Pedophilia: Dissecting Effects of the Label and Intentionality Ascriptions. *Archives of Sexual Behavior, 47*(2), 353–361. <https://doi.org/10.1007/s10508-017-1048-8>
- Jayaratne, T. E., Ybarra, O., Sheldon, J. P., Brown, T. N., Feldbaum, M., Pfeffer, C. A., & Petty, E. M. (2006). White Americans' genetic lay theories of race differences and sexual orientation: Their relationship with prejudice toward blacks, and gay men and lesbians. *Group Processes & Intergroup Relations, 9*(1), 77–94. <https://doi.org/10.1177/1368430206059863>
- Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and*

*Social Psychology*, 88(4), 686–702. <https://doi.org/10.1037/0022-3514.88.4.686>

Kesberg, R., & Keller, J. (2021). Personal values as motivational basis of psychological essentialism: An exploration of the value profile underlying essentialist beliefs. *Personality and Individual Differences*, 171(August 2020), 110458.

<https://doi.org/10.1016/j.paid.2020.110458>

Kraus, M. W., & Keltner, D. (2013). Social class rank, essentialism, and punitive judgment.

*Journal of Personality and Social Psychology*, 105(2), 247–261.

<https://doi.org/10.1037/a0032895>

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). *lmerTest: Tests in linear mixed effects models [Computer software manual]*.

Logel, C., Stinson, D. A., Gunn, G. R., Wood, J. V., Holmes, J. G., & Cameron, J. J. (2014). A little acceptance is good for your health: Interpersonal messages and weight change over time. *Personal Relationships*, 21(4), 583–598. <https://doi.org/10.1111/pere.12050>

Mandalaywala, T. M., Amodio, D. M., & Rhodes, M. (2017). Essentialism Promotes Racial Prejudice by Increasing Endorsement of Social Hierarchies. *Social Psychological and Personality Science*, 194855061770702. <https://doi.org/10.1177/1948550617707020>

Martin, J. W., Charles, S., & Heiphetz, L. (Forthcoming). Essentialist views of criminal behavior increase punishment. In J. Musolino, J. Sommer, & P. Hemmer (Eds.), *The Cognitive Science of Belief*. Cambridge University Press.

Martin, J. W., & Cushman, F. (2015). To Punish or to Leave: Distinct Cognitive Processes Underlie Partner Control and Partner Choice Behaviors. *PLOS ONE*, 10(4), e0125193.

<https://doi.org/10.1371/journal.pone.0125193>

Martin, J. W., & Cushman, F. (2016a). The Adaptive Logic of Moral Luck. *A Companion to*

- Experimental Philosophy*, 190–202. <https://doi.org/10.1002/9781118661666.ch12>
- Martin, J. W., & Cushman, F. (2016b). Why we forgive what can't be controlled. *Cognition*, 147, 133–143. <https://doi.org/10.1016/j.cognition.2015.11.008>
- McAdams, D. P. (2006). *The redemptive self: stories Americans live by* (Issue November). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195176933.001.0001>
- Mears, D. P., Cochran, J. C., & Cullen, F. T. (2015). Incarceration Heterogeneity and Its Implications for Assessing the Effectiveness of Imprisonment on Recidivism. *Criminal Justice Policy Review*, 26(7), 691–712. <https://doi.org/10.1177/0887403414528950>
- Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In V. S. & O. A. (Eds.), *Similarity and analogical reasoning* (pp. 179–195). Cambridge University Press.
- Monterosso, J., Royzman, E. B., & Schwartz, B. (2005). Explaining away responsibility: Effects of scientific explanation on perceived culpability. *Ethics & Behavior*, 15(2), 139–158. [https://doi.org/10.1207/s15327019eb1502\\_4](https://doi.org/10.1207/s15327019eb1502_4)
- Morris, A., MacGlashan, J., Littman, M. L., & Cushman, F. (2017). Evolution of flexibility and rigidity in retaliatory punishment. *Proceedings of the National Academy of Sciences*, 114(39), 10396–10401. <https://doi.org/10.1073/pnas.1704032114>
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2005). Surveying Freedom: Folk Intuitions about free will and moral responsibility. *Philosophical Psychology*, 18(5), 561–584. <https://doi.org/10.1080/09515080500264180>
- Newman, G. E., De Freitas, J., & Knobe, J. (2015). Beliefs about the true self explain asymmetries based on moral judgment. *Cognitive Science*, 39(1), 96–125. <https://doi.org/10.1111/cogs.12134>
- Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of

- folk intuitions. *Nous*, 41(4), 663–685. <https://doi.org/10.1111/j.1468-0068.2007.00666.x>
- Pager, D. (2007). *Marked: Race, crime, and finding work in an era of mass incarceration*. The University of Chicago Press.
- Prentice, D. A., & Miller, D. T. (2006). Essentializing differences between women and men. *Psychological Science*, 17(2), 129–135. <https://doi.org/10.1111/j.1467-9280.2006.01675.x>
- Rhodes, M., Leslie, S.-J., Saunders, K., Dunham, Y., & Cimpian, A. (2018). How does social essentialism affect the development of inter-group relations? *Developmental Science*, 21(1), e12509. <https://doi.org/10.1111/desc.12509>
- Roberts, S. O., Ho, A. K., Rhodes, M., & Gelman, S. A. (2017). Making Boundaries Great Again: Essentialism and Support for Boundary-Enhancing Initiatives. *Personality and Social Psychology Bulletin*, 43(12), 1643–1658. <https://doi.org/10.1177/0146167217724801>
- Robinson, P., & Darley, J. (1995). *Justice, liability, and blame: Community views and the criminal law*. Westview Press. <http://psycnet.apa.org/psycinfo/1995-97571-000>
- Ryazanov, A. A., & Christenfeld, N. J. S. (2018). The strategic value of essentialism. *Social and Personality Psychology Compass*, 12(1), e12370. <https://doi.org/10.1111/spc3.12370>
- Scurich, N., & Appelbaum, P. (2016). The blunt-edged sword: Genetic explanations of misbehavior neither mitigate nor aggravate punishment. *Journal of Law and the Biosciences*, 3(1), 140–157. <https://doi.org/10.1093/jlb/lsv053>
- Shariff, A. F., Greene, J. D., Karremans, J. C., Luguri, J. B., Clark, C. J., Schooler, J. W., Baumeister, R. F., & Vohs, K. D. (2014). Free Will and Punishment: A Mechanistic View of Human Nature Reduces Retribution. *Psychological Science*. <https://doi.org/10.1177/0956797614534693>
- Strohinger, N., & Nichols, S. (2014). The essential moral self. *Cognition*, 131(1), 159–171.

<https://doi.org/10.1016/j.cognition.2013.12.005>

Tannenbaum, D., Uhlmann, E. L., & Diermeier, D. (2011). Moral signals, public outrage, and immaterial harms. *Journal of Experimental Social Psychology, 47*(6), 1249–1254.

<https://doi.org/10.1016/j.jesp.2011.05.010>

Travis, J. (2005). *But they all come back: Facing the challenges of prisoner reentry*. The Urban Institute.

Travis, J., & Maul, M. (2003). *Prisoners once removed: The impact of incarceration and reentry on children, families, and communities* (J. Travis & M. Waul (eds.)). The Urban Institute.

Treadway, M. T., Buckholtz, J. W., Martin, J. W., Jan, K., Asplund, C. L., Ginther, M. R., Jones, O. D., & Marois, R. (2014). Corticolimbic gating of emotion-driven punishment. *Nature Neuroscience, 17*(9), 1270–1275. <https://doi.org/10.1038/nn.3781>

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology, 46*(1), 35–57. <https://doi.org/10.1086/406755>

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A Person-Centered Approach to Moral Judgment. *Perspectives on Psychological Science, 10*(1), 72–81.

<https://doi.org/10.1177/1745691614556679>

Uhlmann, E. L., & Zhu, L. [Lei]. (2014). Acts, Persons, and Intuitions. *Social Psychological and Personality Science, 5*(3), 279–285. <https://doi.org/10.1177/1948550613497238>

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two brief interventions to mitigate a “chilly climate” transform women’s experience, relationships, and achievement in engineering. *Journal of Educational Psychology, 107*(2), 468–485.

<https://doi.org/10.1037/a0037461>

Webster, R. J., & Saucier, D. A. (2013). Angels and Demons Are Among Us: Assessing

Individual Differences in Belief in Pure Evil and Belief in Pure Good. *Personality and Social Psychology Bulletin*, 39(11), 1455–1470. <https://doi.org/10.1177/0146167213496282>

Western, B. (2018). *Homeward: Life in the year after prison*. Russell Sage Foundation.

Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology*, 94(6), 1033–1047. <https://doi.org/10.1037/0022-3514.94.6.1033>

Willis, G. M. (2018). Why call someone by what we don't want them to be? The ethics of labeling in forensic/correctional psychology. *Psychology, Crime & Law*, 24(7), 727–743. <https://doi.org/10.1080/1068316X.2017.1421640>

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences*, 104(20), 8235–8240. <https://doi.org/10.1073/pnas.0701408104>

Zelazo, P. D., Helwig, C. C., & Lau, A. (1996). Intention, act, and outcome in behavioral prediction and moral judgment. *Child Development*, 67(5), 2478–2492. <https://doi.org/10.2307/1131635>