

Children's and adults' views of punishment as a path to redemption

James P. Dunlea^a and Larisa Heiphetz^b

^a Department of Psychology, Columbia University. 1190 Amsterdam Ave., New York, NY 10027, United States. E-mail: james.dunlea@columbia.edu. Phone: 212-853-1406.

ORCID: 0000-0003-2499-4970 (Corresponding Author)

^b Department of Psychology, Columbia University. 1190 Amsterdam Ave., New York, NY 10027, United States. E-mail: lah2201@columbia.edu. Phone: 212-854-1348.

Word count (abstract + main text): 10,005

In press, *Child Development*

Acknowledgments: The authors wish to thank Jane Acinero, Sophie Bair, Jenny Barshay, Emma Bippart-Butler, Riley Burke, Jack Demaree, Tamás Gömöri, Kaya Hantsbarger, Rahil Kamath, Brendan Krovatin, Abby McLaughlin, Rachel Mulholland, Carolina Santiago-Robles, Andre Senecal, Samuel So, Aaliyah Triumph, Redeate Wolle, and Julia Yermash for assisting with data collection and data coding; Geraldine Downey, Andrea Fields, Mike Naft, and members of the Columbia Social and Moral Cognition Lab for providing feedback; and the Brooklyn Children's Museum and the Cognitive Development Lab at New York University for providing testing space. This project was made possible through the support of grant #61080 from the John Templeton Foundation to LH. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. This work has also been supported (in part) by a Visiting Scholar award from the Russell Sage Foundation to LH. Any opinions expressed are those of the author alone and should not be construed as representing the opinions of the Foundation. Columbia University provided additional funding to support this work.

Abstract

The current work investigated the extent to which children ($N=171$ 6- to 8-year-olds) and adults ($N=94$) view punishment as redemptive. In Study 1, children—but not adults—reported that “mean” individuals became “nicer” after one severe form of punishment (incarceration). Moreover, adults expected “nice” individuals’ moral character to worsen following punishment; however, we did not find that children expected such a change. Study 2 extended these findings by showing that children view “mean” individuals as becoming “nicer” following both severe (incarceration) and relatively minor (time-out) punishments, suggesting that the pattern of results from Study 1 generalizes across punishment types. Together, these studies indicate that children—but not adults—may view punishment as a vehicle for redemption.

Keywords: moral cognition; punishment; social cognitive development

Children's and Adults' Views of Punishment as a Path to Redemption

In the fictional novel *The 7 ½ Deaths of Evelyn Hardcastle* (Turton, 2018), readers become acquainted with Aiden Bishop, a fictional character who describes his time spent at Blackheath Manor, a futuristic prison. While at Blackheath Manor, Aiden muses that putting people behind bars “can’t build better men” and “can only break what goodness remains” (p. 436). In other words, Aiden believes that this form of punishment is ineffective in catalyzing moral improvement.

Here, we join other scholars in conceptualizing punishment as a “behavior aimed at those who cause harm or violate social norms” (Deutchman, Bračić, Raihani, & McAuliffe, in press, p. 2). Scholars sometimes describe punishment as “costly,” meaning that punishers *pay* a cost (e.g., putting themselves at risk for retaliation) for a transgressor to *incur* a cost (e.g., loss of resources; Fehr & Gächter, 2002; McAuliffe, Jordan, & Warneken, 2015; Salali, Juda, & Henrich, 2015). Prior work has delineated how punishment differs from other responses to perceived moral transgression, such as logical consequences. Whereas punishment need not directly address the transgression-induced outcome (e.g., Mageau et al., 2018; Nelsen, 1985; Robichaud & Mageau, 2019), logical consequences require transgressors to respond to the harm they have caused (Ginott, 1965). For example, parents who employ a punitive approach to transgression may decide to take away their child’s television privileges because she purposefully broke a different object. This approach qualifies as punishment because the parent took away a privilege (watching television) that was unrelated to the child’s misdeed (breaking something else). In contrast, parents who uphold a logical consequences approach to transgression may require their child to repair an object that she purposefully broke.

While punishment is just one possible response to perceived moral transgression, it served as the focus of the current studies because it is a common reaction to wrongdoing. Specifically, Study 1 focused on children's and adults' views about the impact of one particularly severe form of punishment (incarceration) on moral character. Of course, incarceration is just one type of punishment, and laypeople's reasoning about incarceration may differ from their reasoning about other forms of punishment. For example, adults may be especially pessimistic about the impact of incarceration on moral character because they have negative stereotypes about what it is like to spend time in prison or jail. Nevertheless, the current work initially probed views about the impact of incarceration on moral character because such punishment is a common response to perceived transgression within the United States criminal legal system (e.g., Alexander, 2012, Forman, 2017; Travis, Western, & Redburn, 2014; Van Cleve, 2016). Study 2 compared incarceration with a less severe form of punishment (time-out) to probe the extent to which results generalized across different forms of punishment. Together, these studies afforded us the opportunity make theoretical contributions spanning developmental psychology, moral cognition, and experimental jurisprudence. We outline these contributions in the sections below.

How Might Age-Related Changes in Socio-Moral Cognition Shape Views About the Impact of Punishment on Moral Character?

Adults living in the United States may understand punishment as communicating negative information about punished individuals' moral character (Dunlea & Heiphetz, in press; Kleinfeld, 2016; Van Cleve, 2016; Yankah, 2004). Though people sometimes describe incarceration as paying a debt for a moral wrong, predominant cultural narratives in the United States sometimes portray those directly implicated in the criminal legal system as morally

bankrupt and forever unable to repay their debt. Such individuals are often depicted “as remaining criminal for life” and as having “inelastic” immoral character (Yankah, 2004, p. 1027; see also Dunlea & Heiphetz, in press; Kleinfeld, 2016; Van Cleve, 2016). As such, adults may view punishment as ineffective in catalyzing moral improvement. However, the developmental trajectory of this view remains unclear.

The current work recruited elementary school-aged children and adults in order to test between two competing possibilities regarding how views about the impact of punishment on moral character may change with age. On the one hand, elementary schoolers may be *less* likely than adults to view punishment as effectively improving moral character. This possibility is rooted in past work on *psychological essentialism*—the tendency to view others’ characteristics as stemming from internal, immutable, biologically-based “essences” (Gelman, 2003; Medin & Ortony, 1989). Although both children and adults endorse the idea that essences exist (for review, see Rhodes & Mandalaywala, 2017), elementary schoolers typically report more essentialist perspectives than do adults (e.g., Chalik, Leslie, & Rhodes, 2017; Heiphetz, Gelman, & Young, 2017; Hussak & Cimpian, 2019; Taylor, Rhodes, & Gelman, 2009). Crucially, recent work has documented age-related decreases in essentialist reasoning within the moral domain (Heiphetz, 2020). Such studies have shown that children are more likely than adults to attribute punishment to an internal “bad” essence (Dunlea & Heiphetz, 2020) and to perceive moral badness as arising from an internal, unchanging source (Heiphetz, 2019). Given that children are especially prone to viewing human characteristics—including moral character—as innate and unchangeable, children in elementary school may be less likely than adults to report that punishment is effective in catalyzing moral growth.

On the other hand, elementary schoolers may be *more* likely than adults to view punishment as an effective mechanism to improve moral character. This possibility is consistent with work showing that children in elementary school are typically more optimistic than adults (Boseovski, 2010; Boseovski & Lee, 2006; Heyman & Giles, 2004; Lockhart, Nakashima, Inagaki, & Keil, 2008). In one study probing views of trait change outside the domain of punishment, children were more likely than adults to report that people's negative qualities change for the better over time (i.e., to perceive that mean people will become nicer over time, Lockhart et al., 2008). Though past experiments have not focused on whether children view punishment as catalyzing positive growth, sociological data provide initial support favoring this possibility. In one diary study probing children's justifications for punishment (Twum-Danso Imoh, 2013), several children spoke of punishment as a vehicle for moral improvement. For example, one child reported that, without punishment, an unruly child will likely "grow up to become a bad person" (p. 479). One interpretation of these types of responses is that punishment can transform an individual with immoral character into someone who is virtuous. Speaking more directly to this interpretation, another respondent noted that "punishment is one of the ways through which a child can be corrected, so it is good" (p. 479).

In addition to testing between these competing developmental possibilities, the current work builds on theoretical models of punishment. Previous research suggests a unidirectional link between perceived immoral character and punitive outcomes: whereas perceived immorality typically augments the likelihood of punitive outcomes (e.g., receiving blame), perceived goodness attenuates such outcomes (Nadler, 2012; Nadler & McDonnell, 2011). For instance, in one line of work, participants learned about two people—one described as "good" and the other as "bad"—who both committed the same transgression (starting a deadly fire, Nadler &

McDonnell, 2011). Although each individual committed the same transgression, participants reported that the “bad” person was more responsible, more blameworthy, and less likable than the “good” person. In other words, perceived immoral character increased punitive outcomes even when behavior was held constant. This finding suggests a directional link between perceived immorality and punishment-related outcomes. If children and adults perceive punishment as changing moral character, this would suggest that the link between perceived immorality and punishment is *bidirectional*. In sum, the present work adds to past theoretical models of punishment by investigating the extent to which punishment impacts perceived moral character.

Why Study Punishment in the Context of the Criminal Legal System?

The current work used the criminal legal system as an example domain in which to investigate children’s and adults’ views about punishment. Some scholars working in the legal tradition conceptualize punishment as “expressive,” i.e., as both an action and a mechanism for social messaging (e.g., Duff, 2011; Feinberg, 1965; Kahan, 1996; Markel, 2011; Murphy & Hampton, 1988). For instance, scholars have argued that punishment expresses information about community norms (Duff, 2011; Markel, 2011) or social hierarchies (i.e., the social standing of victims relative to transgressors, Murphy & Hampton, 1988). Importantly, scholars writing on the expressive function of punishment typically conceptualize punishment as severe. For instance, punishment has been equated with “hard treatment” (Feinberg, 1965, p. 397) and depicted as requiring “pain or other consequences normally considered unpleasant” (Hart, 1959, p. 4). Given that severe punishment is typically confined to formal systems such as incarceration, scholarship discussing the expressive function of punishment is often couched within the criminal legal system context.

Although this past work has argued that punishment carries communicative weight, few programs of research within psychology have empirically tested how laypeople interpret punishment's messages. The studies that have employed experimental methods to answer related questions, however, have largely focused on testing adults (Bilz, 2016; Ho, Cushman, Littman, & Austerweil, 2019; Okimoto & Wenzel, 2011; for an exception with children, see Bregant, Shaw, & Kinzler, 2016). The current work leveraged experimental methods to investigate how both children and adults reason about the impact of one particularly severe form of punishment (incarceration) on moral character. Specifically, we tested 6- to 8-year old children and adults in the same experimental paradigm. Testing elementary schoolers was important for two main reasons. First, testing children in this age range allowed us to extend, and compare our results with, previous work examining elementary schoolers' essentialism (e.g., Heiphetz, 2019; Hussak & Cimpian, 2019; Taylor et al., 2009) and, separately, their optimism (e.g., Boseovski, 2010). As mentioned in the section above, these separate bodies of scholarship suggest that both essentialist reasoning and optimism are high during the elementary school years. Thus, testing children in this age range allowed us to examine the extent to which children's optimism overwhelms their tendency to believe human characteristics such as "badness" are immutable. Second, past work suggests that 6- to 8-year-olds can reason and respond to questions about individuals who have been implicated in the criminal legal system (Bregant et al., 2016; Dunlea & Heiphetz, 2020; Dunlea, Wolle, & Heiphetz, in press). As such, we were able to ask children questions about individuals who received a relatively severe form of punishment (incarceration) that is specific to criminal legal system context and compare their inferences about such individuals with their inferences about people who received a relatively less severe form of punishment (time-out).

In addition to the theoretical contributions laid out above, studying punishment in the context of the criminal legal system addresses an important topic that is understudied in psychology. Scholarship at the intersection of psychology and law has documented discrimination and prejudice toward those who have been implicated in the criminal legal system (e.g., Banks, Eberhardt, & Ross, 2006; Haney, 2012; Moore, Stuewig, & Tangney, 2016; Richardson & Goff, 2013). While this literature has underscored the far-reaching implications of legal punishment in the United States, relatively less work has examined the psychological underpinnings of *why* formerly incarcerated individuals may experience such negativity long after experiencing incarceration. This negativity may partially arise from the view that incarceration within the United States context fails to improve people (Dunlea & Heiphetz, in press; Kleinfeld, 2016; Van Cleve, 2016; Yankah, 2004). The current work addressed this possibility by examining how laypeople view those who have experienced incarceration.

Overview of Current Work

The current work examined laypeople's views about the impact of certain types of punishment on perceived moral character. Study 1 examined this question by asking children and adults to indicate the extent to which "nice" and, separately, "mean" individuals' moral qualities changed following one particularly severe type of punishment—incarceration. By testing children and adults in the same paradigm, the present work provided insight into how judgments about the impact of punishment change throughout development. Study 2 built on the results of Study 1 by investigating the extent to which children's views generalize across different forms of punishment.

Study 1

Study 1 investigated the extent to which children and adults view punishment as driving moral change. To do so, we told participants about both a morally good (“nice”) and, separately, a morally bad (“mean”) individual. Participants in the experimental condition learned that these individuals were punished for breaking the law, whereas the participants in the control condition learned that these individuals went on a business trip. Participants indicated how morally good each individual would be both during and after the incarceration or trip. Data for this study were collected between Fall 2017 and Spring 2018.

Method

Participants. Participants included 94 children between six and eight years old ($M_{age}=6.92$ years, $SD_{age}=.79$ years; 57% female, 43% male). Children’s parents completed a demographic questionnaire in which they identified their children as White or European-American (45%), Black or African-American (17%), Asian or Asian-American (14%), multiracial (14%), and “other” (6%); the remaining parents did not answer this question. Parents reported their child’s ethnicity using a separate question, and 18% of parents identified their children as Hispanic or Latinx. Data from five additional children were excluded because they did not comprehend the experimental items. Children were recruited from a departmental database and from a children’s museum in a large city in the northeastern United States. Here and in Study 2, families signed up for inclusion in the departmental database either in person (at public street fairs, public parks, and the aforementioned children’s museum) or by visiting our laboratory’s website. Given our recruitment strategy, any families with eligible children could participate. All children received a small prize for participating.

Participants also included 94 adults between 18 and 52 years old ($M_{age}=22.68$ years, $SD_{age}=5.74$ years; 66% female, 34% male). Adult participants self-identified as White or

European-American (44%), Black or African-American (19%), Asian or Asian-American (26%), Native American or Pacific Islander (1%), multiracial (7%), or “other” (3%). Additionally, 10% of adults self-identified as Hispanic or Latinx. Data from three additional adults were excluded because they did not correctly answer an attention check item asking them to describe any of the characters presented throughout the study. Adults were recruited through the psychology department’s participant pool and from the greater metropolitan community. Adults who participated via the university participant pool received .5 credits, and adults from the greater community received a small prize, such as a piece of candy.

Thirty-four adults reported that they knew at least one person who has previously served time in a jail or prison. Additionally, seven parents reported that their child knew at least one person who has experienced incarceration. However, no significant differences in responses emerged between individuals who knew at least one incarcerated individual and those who did not (see Supplementary Materials for relevant exploratory analyses concerning this variable). We also conducted a series of exploratory analysis examining the extent to which participant race and ethnicity predicted participant responses. Although members of racial and ethnic minority groups often have very different experiences in the criminal legal system than majority group members (e.g., Alexander, 2012; Banks et al., 2006; Richardson & Goff, 2013; Van Cleve, 2016), neither of these variables reliably predicted participants’ responses (see Supplementary Materials for relevant analyses).

Procedure. Here and in Study 2, an experimenter interviewed children individually in a quiet room located in a children’s museum or in a developmental psychology laboratory. First, the experimenter said that he or she would ask children questions and that there were no right or wrong answers. Further, the experimenter specified that he or she would read sentences about

other people and that children would indicate how much they agreed with each sentence. The experimenter then introduced children to a five-point scale consisting of stick figures arrayed from smallest to largest on a sheet of paper and instructed children on how to use the scale (e.g., asking them to point to the smallest picture if they didn't agree at all with a sentence the experimenter said). The remaining labels were "agree a little bit," "agree a medium amount," "agree a lot," and "agree completely." The experimenter then asked children two test items to gauge their understanding of the scale ("Can you show me where you would point if you didn't agree with the answer at all?"; "Can you show me where you would point if you agree a medium amount?"). On average, children used the scale correctly: 99% correctly answered the item asking where they would point if they "don't agree with the answer at all" and 90% correctly answered the item asking where they would point if they "agree a medium amount." Participants who answered incorrectly received corrective feedback and, subsequently, received a second chance to respond to the item. All participants who did not answer correctly initially provided the correct answer on their second try. See Supplementary Materials for relevant study materials associated with this article, including the scale used to elicit children's responses.

Following these instructions, the experimenter showed children pictures of stick figures on a PowerPoint display. Here and in Study 2, we referred to each individual using male pronouns because most people incarcerated in the United States are male (Bronson & Carson, 2019). The experimenter then pointed to each individual, one at a time, and described him as having either good or bad moral character. For example, the experimenter described the "nice" (morally good) individual as liking to "help others" and the "mean" (morally bad) individual as liking to "start fights with other people." In addition to learning about one "nice" and one "mean" individual, participants learned about a "religious" individual and an "atheistic"

individual. Including the latter two individuals in the present study allowed us to investigate how participants' views of religious qualities compared to their views of moral qualities. This question did not directly concern the main research question; therefore, participants' views of religious qualities will not be discussed further in the Main Text (see Supplementary Materials for relevant experimental items and analyses).

After providing information about an individual's moral character, the experimenter told participants that the individual had broken the law and gone to jail (punishment condition) or that the individual had gone on a business trip (control condition); that is, condition type was a between-participants variable. We chose a business trip as the control condition because, like incarceration (Travis, 2005), work-related absences can be lengthy but are typically temporary. To aid comprehension, children in the punishment condition saw a series of PowerPoint animations depicting the punished individual entering and subsequently being carried away from his home by a police car. After watching the series of animations, children viewed an image of the punished individual standing in a jail cell. Children in the control condition saw a series of PowerPoint animations depicting the control individual entering and subsequently being carried away from his home by an airplane. After watching these animations, children viewed an image of the control individual standing in a business office.

The remainder of Study 1 progressed in two parts. During Part I, the experimenter asked participants in the punishment condition to rate their agreement with four statements regarding each punished individual's moral character (e.g., "How much do you agree that now, Frank is a good person deep, deep down inside?"). Each of the four items highlighted positively valenced characteristics: one item focused on individuals' good "essences," one item focused on individuals' kind behaviors, and two items focused on individuals' prosocial intentions (see

Supplementary Materials for all experimental items). We framed items in terms of individuals' positively, instead of negatively, valenced characteristics because we did not want to introduce or reinforce any negative preconceptions about individuals who have come into contact with the criminal legal system.

The experimenter pointed at the individual standing in the jail cell while asking each item. The experimenter began Part II by saying, "Frank stays in jail for a really, really long time. Eventually, Frank finishes all the time that he needed to spend in jail, and he is allowed to go back home." To ensure that children understood that the punished individuals left jail, the experimenter then showed children corresponding images of each individual standing near his home. After providing this information, the experimenter once again asked participants to rate their agreement with the same four statements regarding each individual's moral character (e.g., "How much do you agree that now, Frank is a good person deep, deep down inside?"). The purpose of asking the same set of experimental items in Parts I and II was to examine the extent to which participants believe that moral character changes as a function of punishment. Critically, measuring perceived moral character at baseline (i.e., at the beginning of individuals' punishment) and following the punishment allowed us to tabulate the extent to which participants perceived punishment as eliciting moral improvement, moral decline, or no change in moral character.

The procedure of the control condition closely mirrored that of the punishment condition. Here, participants rated their agreement with four statements regarding each individual's moral character at the beginning (Part I) and conclusion (Part II) of the business trip. As in the punishment condition, we employed visual aids to ensure children understood that the individuals in the control vignettes had returned from their business trips. Specifically, the

experimenter showed children corresponding images of each control individual standing near his home.

Adults completed this procedure online and selected the scale label that best matched their response (i.e., they viewed only the verbal labels, not the stick figures used to represent the scale to children). Moreover, adults did not view the visual stimuli used to aid children's comprehension of the story; instead, they only read descriptions of each scenario. We made these changes because adults are better able than children to attend to verbal information and do not require pictures to draw their attention to stimuli.

Both children and adults answered all items about each individual before moving on to the next individual. For example, after answering all items about the “mean” individual, participants completed the same procedure for the “nice” individual. The order in which participants learned about each individual was counterbalanced across participants, as was the order of experimental items regarding each individual. Based on recommendations to include approximately 50 participants per cell in psychological research (Lakens & Evers, 2014; Simmons, Nelson, & Simonsohn, 2013), we aimed to recruit approximately 50 participants in each condition (punishment condition: $n_{\text{children}}=48$; $n_{\text{adults}}=49$; control condition: $n_{\text{children}}=46$; $n_{\text{adults}}=45$). The number of participants per cell is comparable to prior work in developmental psychology (Heiphetz, 2019; Heiphetz, Lane, Waytz, & Young, 2018; Rhodes, Leslie, Saunders, Dunham, & Cimpian, 2018; Shaw & Olson, 2015).

Results

Here and in Study 2, we adjusted analyses that included multiple comparisons using a Bonferroni correction. Additionally, we conducted sensitivity power analyses in G*Power (power=80%, alpha=.05; Faul, Erdfelder, Lang, & Buchner, 2007) to ensure that the current

sample sizes provided enough power to detect reliable effects. Unless otherwise noted, all significant pairwise comparisons yielded p values below the Bonferroni-corrected significance threshold and effect sizes that were above the sensitivity analysis-generated threshold. See Supplementary Materials for additional information regarding the adjusted alpha level and sensitivity analysis generated threshold for each set of pairwise comparisons. In addition to the main analyses reported below, we investigated whether participant age predicted responses in our data. This variable did not reliably predict participants' responses. See Supplementary Materials for these analyses and descriptive statistics for each item in each of the studies.

We averaged participants' responses to the four items regarding each individual's moral character at the beginning of his incarceration/trip to create a composite score ("nice" individuals: $\alpha=.81$; "mean" individuals: $\alpha=.89$). We did the same for the four items regarding each individual's moral character after the incarceration/trip ("nice" individuals: $\alpha=.84$; "mean" individuals: $\alpha=.95$). For ease of interpretation, the main dependent variable was the difference between these two composite values, indicating perceived moral change. In addition to conducting analyses using a difference score, we analyzed our data using raw means, and our interpretation of the results is consistent across analytic approaches. See Supplementary Materials for these analyses and the associated descriptive statistics.

Evaluations of "nice" individuals. We investigated participants' responses using two types of analyses (Figure 1). First, we examined the extent to which participants reported that "nice" individuals changed after punishment and after going on a business trip. We used a series of one-sample t -tests to compare perceived change scores to 0 (indicating, on average, no perceived moral change) in each condition among children and, separately, adults. Children did not report that "nice" individuals changed as a function of punishment ($t(47)=1.72, p=.092$,

Cohen's $d=.25$, 95% CI_{diff} : [-.03, .37]) or of going on a business trip ($t(45)=-1.84$, $p=.072$, Cohen's $d=-.27$, 95% CI_{diff} : [-.30, .01]). Like children, adults did not report that "nice" individuals changed as a result of going on a business trip ($t(44)=.71$, $p=.479$, Cohen's $d=.11$, 95% CI_{diff} : [-.06, .13]). However, unlike children, adults reported that "nice" individuals became less "nice" after punishment ($t(47)=-3.21$, $p=.002$, Cohen's $d=-.46$, 95% CI_{diff} : [-.43, -.10]).

Next, we conducted a 2 (Participant Age: child vs. adult) x 2 (Condition: punishment vs. control) between-participants ANOVA in order to investigate whether children and adults reported different magnitudes of change in each condition. This analysis revealed a Participant Age x Condition interaction ($F(1, 183)=14.70$, $p<.001$, $\eta_p^2=.07$). Neither of the main effects reached significance ($ps \geq .101$). To examine the Participant Age x Condition interaction, we conducted two sets of tests. First, we compared the magnitude of change children expected in the punishment condition with the magnitude of change children expected in the business trip condition, and we conducted an analogous comparison among adults. Both children and adults distinguished between the punishment and the business trip conditions. However, they did so in different ways. Children reported that "nice" individuals became "nicer" after receiving punishment than after going on a business trip ($p=.006$, Cohen's $d=.51$, 95% CI_{diff} : [.09, .54]). However, this effect size was smaller than the smallest effect size that could be detected given the present samples; thus, caution is warranted in interpreting this result. In contrast, adults reported that "nice" individuals became less "nice" after receiving punishment than after going on a business trip ($p=.009$, Cohen's $d=.64$, 95% CI_{diff} : [-.52, -.08]).

Second, we investigated whether children and adults reported different magnitudes of change in the punishment versus business trip condition. Children reported stronger increases in moral goodness after punishment than did adults ($p<.001$, Cohen's $d=.69$, 95% CI_{diff} : [.22, .66]).

A significant difference in perceived change did not emerge between children and adults in the business trip condition ($p=.129$, Cohen's $d=.41$, 95% CI_{diff} : [-.40, .05]).

Evaluations of “mean” individuals. Once again, we investigated participants' responses using two types of analyses (Figure 1). First, we examined the extent to which participants perceived “mean” individuals to change after receiving punishment and, separately, after going on a business trip. As above, we used a series of one-sample t -tests to compare perceived change scores to 0 in both conditions among children and, separately, adults. Children reported that “mean” individuals became “nicer” after receiving punishment ($t(46)=7.21$, $p<.001$, Cohen's $d=1.05$, 95% CI_{diff} : [.78, 1.39]) and after going on a business trip ($t(45)=2.91$, $p=.006$, Cohen's $d=.43$, 95% CI_{diff} : [.11, .63]). However, these effects did not emerge among adults (punishment: $t(47)=.51$, $p=.612$, Cohen's $d=.07$, 95% CI_{diff} : [-.11, .18]); business trip: $t(42)=1.02$, $p=.315$, Cohen's $d=.16$, 95% CI_{diff} : [-.06, .19]).

Next, we conducted a 2 (Participant Age: child vs. adult) x 2 (Condition: punishment vs. control) between-participants ANOVA in order to investigate whether children and adults reported different magnitudes of change in each condition. This analysis revealed main effects of Participant Age ($F(1, 180)=37.43$, $p<.001$, $\eta_p^2=.17$) and Condition ($F(1, 180)=9.66$, $p=.002$, $\eta_p^2=.05$). These main effects were qualified by a Participant Age x Condition interaction ($F(1, 180)=11.27$, $p=.001$, $\eta_p^2=.06$). To further examine this interaction, we conducted two sets of tests. First, we compared the magnitude of change children expected in the punishment condition with the magnitude of change children expected in the business trip condition, and we conducted an analogous analysis among adults. Children expected stronger increases in moral goodness after punishment than after a business trip ($p<.001$, Cohen's $d=.75$, 95% CI_{diff} : [.41, 1.02]). This pattern did not emerge among adults ($p=.862$, Cohen's $d=.06$, 95% CI_{diff} : [-.34, .28]). Second, we

investigated whether children and adults reported different magnitudes of change in each condition. Children reported stronger increases in moral goodness after punishment than did adults ($p < .001$, Cohen's $d = 1.30$, 95% CI_{diff} : [.75, 1.35]). No significant difference in perceived change emerged between children and adults in the business trip condition ($p = .056$, Cohen's $d = .45$, 95% CI_{diff} : [-.01, .62]).

Discussion

The purpose of Study 1 was to examine children's and adults' judgments about how "nice" and, separately, "mean" individuals' moral attributes might change as a result of being punished (i.e., going to jail). Three main results emerged. First, children, unlike adults, reported that "mean" individuals became "nicer" after punishment. Second, children reported that punishment catalyzed a greater increase in "mean" individuals' moral goodness than did going on a business trip. Third, whereas adults expected "nice" individuals to become less "nice" following punishment, we did not observe this effect among children. These findings extend past work suggesting that children are more optimistic about others than are adults (e.g., Boseovski, 2010) by demonstrating that children are even optimistic about those whom many people perceive to have committed severe moral transgressions (e.g., people who broke the law). Crucially, this work moves beyond past research demonstrating that children believe negative characteristics change for the better over time (e.g., Lockhart et al., 2008) by showing that children reported more moral growth in "mean" individuals after punishment than after going on a business trip. In other words, the present work suggests that children may view severe punishment (incarceration) as a potent driver of moral improvement.

Study 2

Study 2 built on Study 1 in several ways. The primary contribution of Study 2 was to investigate whether the main pattern of results from Study 1 would generalize to different types of punishment. Based on the results from Study 1, it is unclear whether children view only extremely severe punishments (e.g., incarceration) as a vehicle for positive moral change, or whether this effect generalizes to less severe punishments. On the one hand, children sometimes engage in proportional reasoning. For example, they report that people who are more causally responsible for a crime deserve more punishment than those who are less causally responsible (Finkel, Liss, & Moran, 1997). Consequently, children may assume proportionality when judging the link between punishment severity and perceived moral change; namely, they may judge that relatively severe punishments elicit moral improvement more effectively than do relatively mild punishments. On the other hand, children's judgments sometimes strongly depend on the presence of particular factors, even if those factors are only present in small amounts. For instance, children in elementary school sometimes categorize Black-White multiracial individuals as Black despite the presence of both Black and White ancestors (e.g., Roberts & Gelman, 2017). Similarly, elementary school aged children may judge redemption on the basis of only small amounts of punishment, reasoning that relatively mild punishment can still result in moral improvement. To test between these possibilities, Study 2 compared perceived moral improvement after a relatively minor punishment (time-out) versus a relatively severe punishment (incarceration).

Probing views of moral character following severe versus relatively minor punishment also has implications for theories of jurisprudence. Some punishment theorists assume that punishment must be severe in order to be communicative (Feinberg, 1965; Hart, 1959; Murphy & Hampton, 1988). However, whether or not this is actually the case is an empirical question. By

examining what severe and relatively minor punishments signal about a person's moral character, the present work empirically tested an assumption made by some punishment theorists.

In addition to testing generalizability, Study 2 also made several secondary contributions. First, Study 2 allowed us to determine the extent to which the pattern of results from Study 1 would replicate in a new sample of children. Second, Study 2 changed the point at which we collected the "baseline" measure of individuals' moral character. In Study 1, we asked participants about individuals' moral character at the beginning of their incarceration and, separately, when they returned home. In Study 2, we asked participants about individuals' moral character *before* receiving punishment and, separately, when they returned home. Changing the "baseline" about which we asked allowed us to more precisely capture the degree of perceived change following each type of punishment. Third, Study 2 included a more nuanced measure to capture participants' responses. Data for Study 2 were collected between Summer and Fall 2018.

Method

Participants. Recruitment procedures were identical to Study 1, with two main exceptions. First, we only recruited children in Study 2. We made this decision because only children reported that "mean" people changed in accordance with punishment in Study 1, and we wanted to further probe the potential limits of this effect in Study 2. Second, we recruited children for Study 2 from a departmental database and at a partnering developmental psychology laboratory in a large city in the northeastern United States.

Our final sample included 77 children between 6 and 8 years old ($M_{\text{age}}=6.92$ years, $SD_{\text{age}}=.85$ years; 60% female, 40% male). Parents identified their children as White or European-American (25%), Black or African-American (17%), Asian or Asian-American (16%), Native

American or Pacific Islander (4%), Multiracial (23%), and Other (8%); the remaining parents did not indicate their children's racial group membership. Additionally, 23% of parents identified their children as Hispanic or Latinx. Data from three additional participants were excluded because the child did not speak English ($n=1$), the child had a developmental disability ($n=1$), or the parent interfered during testing ($n=1$). Nine parents reported that their child knew at least one person who has experienced incarceration. However, no significant differences in responses to the main dependent variables emerged between participants who knew at least one incarcerated individual and those who did not (see Supplementary Materials for relevant exploratory analyses). Moreover, we also conducted a series of exploratory analysis examining the extent to which participant race and ethnicity predicted participant responses. As in Study 1, neither of these variables reliably predicted participants' responses (see Supplementary Materials for relevant analyses).

Procedure. First, the interviewer said that he or she would ask children questions and that there were no right or wrong answers. The interviewer further specified that he or she would show the child pictures of people on a computer screen, tell the child information about the people, and subsequently ask questions about them. The remainder of the study proceeded in four main parts, which we describe below. The entire procedure is also illustrated in Figure 2.

During Part I, the experimenter displayed an image of a pair of individuals on a laptop computer screen and described the pair as being either “nice” or “mean” (e.g., “Here are two people—Frank and Bobby. Frank and Bobby like to start fights with other people. They are *both* very mean”). These descriptions closely matched those used in Study 1. After describing the pair of individuals, the experimenter asked participants four items about each individual's moral character. The content of the moral character items in Study 2 was identical to those used in

Study 1 (as in Study 1, each of the four items highlighted positively valenced characteristics); however, the method of eliciting responses differed across the two studies. Instead of eliciting participants' responses using a five-point scale as in Study 1, participants in Study 2 responded to yes-or-no items (e.g., "Do you agree that Frank is a good person deep, deep down inside? Yes, or no?"). The experimenter followed up each item with a more fine-grained item ("How sure are you? Are you very sure, kind of sure, or not very sure?"). We assigned responses numerical values from -3 (indicating least optimism about individual's moral character, e.g., very sure that the individual was *not* morally good) to +3 (indicating most optimism about individual's moral character, e.g., very sure that the individual *was* morally good). We adapted this two-step approach from prior research in developmental psychology (e.g., Bregant et al., 2016; Hussak & Cimpian, 2018). Implementing this approach allowed us to capture more nuanced responses from participants in a way young children could understand. Participants answered items about one individual in the pair before moving on to items about the other individual in the pair.

After answering items about both individuals, participants proceeded to Part II. Here, they learned that each individual in the pair—regardless of whether the pair was described as "nice" or "mean"—broke the law. The experimenter specified that each individual committed the same transgression (e.g., "One day, Frank and Bobby *both* broke the law. They did the exact same thing"). The purpose of providing this instruction was to ensure that participants did not infer that individuals experienced different punishments because they committed different transgressions. Next, participants learned that each individual received punishment. Unlike in Study 1, where individuals were punished by going to jail, individuals in Study 2 were punished in one of two ways. One individual in each pair was described as receiving a relatively severe punishment (going to jail) whereas the other individual was described as receiving a relatively

less severe punishment (going to time-out). To ensure that children understood that going to jail was relatively more severe than going to time-out, the experimenter provided detailed descriptions of both the jail and the “time-out house.” For example, participants learned that individuals who went to jail could never choose what they did, but that individuals who went to time-out could sometimes choose what they did. Additionally, children viewed pictures of each punishment environment as it was described to help them differentiate between the two types of punishment. When learning about individuals who were incarcerated, participants viewed pictures of stick figures standing in a jail cell. When learning about individuals who went to time-out, participants viewed pictures of stick figures standing in a room without jail bars (see Supplementary Materials for an example).

After describing the punishment each individual received, the experimenter asked children the first of two test items to gauge their understanding of the story (e.g., “Can you remind me, where was Frank taken? Was he taken to jail, or was he taken to the ‘time-out house?’”). Each of the test items focused on one individual in the pair. Following the first test item, participants answered four yes-or-no items probing their views of how severely the individual in that item was punished (e.g., “Do you agree that going to jail is a big punishment for Frank? Yes, or no?”). The experimenter followed up each item with a more fine-grained item (“How sure are you? Are you very sure, kind of sure, or not very sure?”). We assigned responses numerical values from -3 (indicating the least severity, e.g., very sure that the punishment is not severe) to +3 (indicating most severity, e.g., very sure that the punishment is severe).

During Part III, participants rated the extent to which the individual highlighted in the first test item wanted to avoid a similar punishment in the future. In doing so, participants answered four yes-or-no items (e.g., “Do you agree that Frank does not want to live in the jail

again? Yes, or no?”). The experimenter followed up each item with additional, more fine-grained items (“How sure are you? Are you very sure, kind of sure, or not very sure?”). We assigned responses assigned numerical values from -3 (indicating least desire to avoid punishment, e.g., very sure that the individual does *not* want to avoid punishment in the future) to +3 (indicating most desire to avoid punishment, e.g., very sure that the individual *wants* to avoid punishment in the future). The purpose of asking these items was to test whether the desire to avoid future punishment could explain the relation between punishment type and the degree to which individuals are perceived to change as a result of punishment. However, these items did not become a central component of our analyses because we did not find a significant difference in children’s responses to items probing moral change following jail compared to those probing moral change following time-out (see Results section for additional details). Given this non-significant difference, we do not discuss the mediation analyses further.

During Part IV, the experimenter told participants that the individual discussed in the first test item finished serving time in either the jail or the “time-out house” and returned home. Participants then answered four yes-or-no items about the punished individual’s positive moral characteristics; these items were identical to the items asked about the same individual in Part I. The purpose of having a “baseline” measure of moral character (i.e., before individuals were punished) and a post-punishment measure of moral character was to examine the extent to which participants viewed punishment as a vehicle for moral change. As in Study 1, measuring perceived moral character at baseline (i.e., before individuals’ punishment) and following the punishment allowed us to tabulate the extent to which participants perceived punishment as eliciting moral improvement, moral decline, or no change in moral character.

After recording children's responses to items in Parts I-IV, the experimenter said, "Okay, so now I am going to tell you the rest of the story about [name of the second individual in the first pair]." Following these instructions, the experimenter asked children an additional test item about how the second individual in the pair was punished (e.g., "Can you remind me, where was Bobby taken? Was he taken to jail, or was he taken to the 'time-out house'?"). If children provided an incorrect response, the experimenter reminded participants of the correct answer. Next, the experimenter reminded participants about the details of the second individual's punishment (e.g., "So remember, when he was in the 'time-out house,' people sometimes told Bobby what to do"). The experimenter then asked participants items analogous to those asked about the first individual in Part II (four items about how severely the *second individual* was punished), Part III (four items about how much the *second individual* wanted to avoid a similar punishment in the future), and Part IV (four items about the *second individual's* moral character after punishment). After answering all items about the first pair of individuals (in the example above, the "mean" pair), participants completed the same procedure for the second pair of individuals (in this case, the "nice" pair).

The following items were counterbalanced across participants in Study 2: (1) the order in which individuals within a pair were introduced, (2) the order in which participants learned about the "nice" and "mean" pairs of individuals, (3) the order of experimental items (e.g., items about whether an individual is morally good), (4) the placement of individuals within each trial (e.g., the individual who went to jail was sometimes on the left side of the screen and sometimes on the right side).

Results

Perceived severity of time-out versus jail. To determine whether our manipulation of punishment severity was effective, we first compared children's views of the severity of going to jail versus time-out. After determining that each set of items probing the perceived severity of going to jail ($\alpha_{\text{nice}} = .68$, $\alpha_{\text{mean}} = .74$) and time-out ($\alpha_{\text{nice}} = .79$, $\alpha_{\text{mean}} = .71$) had acceptable reliability, we collapsed items in each condition into separate variables; see Supplementary Materials for descriptive statistics for each set of items in Study 2. Subsequently, we used a series of paired-samples *t*-tests to compare punishment severity scores in each condition. The manipulation of perceived severity across punishment types worked as intended: children viewed jail as a more severe punishment than time-out in both conditions ("nice" condition: $t(73) = 16.43$, $p < .001$, Cohen's $d = 1.91$, 95% CI_{diff} : [3.26, 4.16]; "mean" condition: $t(74) = 19.19$, $p < .001$, Cohen's $d = 2.22$, 95% CI_{diff} : [3.29, 4.05]).

Evaluations of "nice" individuals. We investigated participants' responses using two types of analyses (Figure 3). First, we examined the extent to which participants reported that "nice" individuals changed after each type of punishment. Items probing views of moral goodness at each time point had acceptable reliability in both conditions (jail: $\alpha_{\text{before}} = .69$, $\alpha_{\text{after}} = .86$; time-out: $\alpha_{\text{before}} = .86$, $\alpha_{\text{after}} = .76$). Thus, we collapsed across items for each time point. As in Study 1, the main dependent variable was the difference in perceived moral character before and after punishment. Subsequently, we used a series of one-sample *t*-tests to compare perceived change scores to 0 in both conditions. Children did not report that "nice" individuals' characteristics changed after experiencing either form of punishment (going to jail: $t(73) = -1.70$, $p = .093$, Cohen's $d = -.20$, 95% CI_{diff} : [-.41, .03]; time-out: $t(73) = -2.23$, $p = .029$, Cohen's $d = -.26$, 95% CI_{diff} : [-.54, -.03]; note that the effect for the time-out condition drops to non-significance after applying a Bonferroni correction). Next, we conducted a paired-samples *t*-tests to compare

perceived change scores in the jail and time-out conditions. No significant difference in perceived change emerged ($t(73)=.87, p=.388$, Cohen's $d=0.10$, 95% CI_{diff} : [-.12, .32]).

Evaluations of “mean” individuals. Once again, we investigated participants' responses using two types of analyses (Figure 3). First, we examined the extent to which participants reported that “mean” individuals became “nicer” after each type of punishment. After determining that items probing views of moral character at each time point had acceptable reliability in both conditions (jail: $\alpha_{before}=.66; \alpha_{after}=.87$; time-out: $\alpha_{before}=.79; \alpha_{after}=.86$), we collapsed across items for each time point. Again, the main dependent variable was the difference in perceived moral character before and after punishment. First, we used a series of one-sample t -tests to compare perceived change scores to 0 in both conditions. Children reported that “mean” individuals became significantly “nicer” after going to jail ($t(74)=14.15, p<.001$, Cohen's $d=1.63$, 95% CI_{diff} : [2.80, 3.72]) and, separately, after going to time-out ($t(74)=12.71, p<.001$, Cohen's $d=1.47$, 95% CI_{diff} : [2.58, 3.54]). Next, we conducted a paired-samples t -tests to compare perceived change scores in the jail and time-out conditions. A significant difference did not emerge ($t(74)=1.03, p=.305$, Cohen's $d=.12$, 95% CI_{diff} : [-.18, .60]).

Discussion

Study 2 conceptually replicated and extended the results from Study 1 among a new sample of children. As in Study 1, children in Study 2 reported that “mean” individuals become “nicer” after punishment. Also as in Study 1, we did not observe that children in Study 2 reported that “nice” individuals would change as a result of punishment. In other words, children in Study 2 appeared to view punishment as a vehicle for positive moral change. Additionally, Study 2 examined whether the effects from Study 1 depend on the severity of punishment. Although children in Study 2 reported that going to jail was a more severe punishment than

going to time-out, they expected “mean” individuals to become “nicer” regardless of whether they went to time-out or jail. Importantly, the degree of reported moral improvement did not significantly differ across punishment contexts. However, null effects are difficult to interpret; it is possible that the degree of reported moral improvement differs across punishment contexts, and the current work failed to capture this difference. Thus, caution is warranted in interpreting this result. Nevertheless, these findings suggest that children may view multiple forms of punishment as a signal of redemption.

General Discussion

The current work examined children’s and adults’ views about the impact of punishment on moral character. In Study 1, elementary schoolers and adults reported on the extent to which one especially severe form of punishment (incarceration) impacted “nice” and, separately, “mean” individuals’ moral character. Children, but not adults, reported that “mean” individuals became “nicer” following severe punishment. Moreover, we did not find evidence that children viewed “nice” individuals as changing following severe punishment. Adults, unlike children, reported that “nice” individuals became less “nice” following severe punishment. These findings suggest that children, but not adults, may view at least one type of severe punishment as a vehicle for positive moral change.

Study 2 built on these results by investigating the extent to which children’s beliefs about the impact of punishment on moral character depend on punishment severity. Here, children reported on the extent to which “nice” and, separately, “mean” individuals’ moral character changed following a relatively minor punishment (time-out) and, separately, a relatively severe punishment (incarceration). Despite acknowledging that incarceration is a more severe type of punishment than time-out, children reported that “mean” individuals became “nicer” regardless

of punishment type. Importantly, the degree of reported moral improvement among “mean” individuals did not significantly differ across punishment types, suggesting that children may conceptualize punishment—regardless of how severe—as a signal of moral redemption.

The present findings make theoretical contributions to several bodies of scholarship. First, the current work expands scientific understanding of social cognitive development. Past work led to two competing possibilities regarding how views about the impact of punishment on moral character might differ among children and adults. On the one hand, elementary schoolers are more likely than adults to view the social world through an essentialist lens (e.g., Chalikh et al., 2017; Hussak & Cimpian, 2019; Heiphetz et al., 2017; Taylor et al., 2009), including when reasoning about moral character (Heiphetz, 2019, 2020). In other words, elementary schoolers are typically more likely than adults to view human characteristics—including moral character—as unchangeable, innate, and rooted in biology. Thus, elementary schoolers could be less likely than adults to report that punishment changes moral character. On the other hand, elementary schoolers typically express more optimism than adults (e.g., Boseovski, 2010). Thus, compared to adults, elementary schoolers could be more optimistic that punishment may improve moral character. The results of Study 1 were consistent with the second possibility, showing that elementary schoolers—but not adults—viewed one especially severe type of punishment (incarceration) as helping “mean” individuals become “nicer.” This finding jointly contributes to theories regarding the development of essentialism as well as optimism by showing that children’s positivity may overpower their tendency to apply an essentialist framework to moral character.

Second, the present work adds to theoretical models of punishment by demonstrating that people expect punishment to change others’ moral characteristics. In both Studies 1 and 2,

children reported that “mean” individuals became “nicer” following punishment. Moreover, adults in Study 1 reported that “nice” individuals’ positive moral characteristics worsened following punishment. These findings are noteworthy in light of past work investigating the link between perceived moral character and punishment. Namely, past work has shown that perceived immoral character increases punitive outcomes even when behavior is held constant (e.g., Nadler & McDonnell, 2011). The current work extends this research by showing that punishment also impacts perceptions of punished individuals’ moral character. In other words, the current work demonstrates the reverse directional link by showing that punishment impacts perceived moral character.

Third, the current work contributes to work in experimental jurisprudence. Many legal theorists discuss the expressive nature of punishment (e.g., Feinberg, 1965; Kahan, 1996; Murphy & Hampton, 1988). However, relatively less research has empirically examined how laypeople interpret the message communicated via punishment. The current work addressed this topic and found that children, but not adults, may interpret punishment as expressing a social message that “mean” individuals have changed for the better. Put slightly differently, the present work suggests that children may conceptualize punishment as a vehicle for moral redemption; however, over development, the conceptualization of punishment-as-redemption declines. In addition to elucidating how people interpret punishment’s message, the present work examined the assumption made by some expressive punishment theorists that punishment needs to be severe in order to be communicative (Feinberg, 1965; Hart, 1959; Murphy & Hampton, 1988). The results of Study 2 suggest that there may be nuance to this view. While it is possible that only messages emitted by severe punishments are audible to adults, the current work suggests that, among children, punishment need not be severe in order be communicative. Here, children

reported that both severe (incarceration) and relatively minor (time-out) punishments help “mean” individuals become “nicer.” Thus, at least among children, punishment does not need to be severe in order to be communicative.

Limitations and Directions for Future Research

The present work provides critical insight into how laypeople conceptualize the impact of punishment on moral character. In doing so, the current research clarifies an understudied topic within psychology (incarceration) and leverages experimental methods to answer questions that have traditionally been the purview of legal scholars and philosophers (e.g., what punishment communicates). Nevertheless, the present work is limited in some ways, and several additional questions remain ripe for future investigation.

One limitation of the current work is that it did not focus on *why* children and adults reported different views of the impact of severe punishment on moral character. Thus, a fruitful direction for future research concerns identifying potential mechanisms underlying this pattern of results. One possibility—the *optimism account*—posits that domain-general age-related changes in social cognition may underlie differences in children’s and adults’ responses. As discussed in the Introduction, elementary schoolers—the age range tested in the current research—are typically more optimistic than adults (e.g., Boseovski, 2010). Importantly, past research has documented age-related declines in optimism regarding different types of characteristics (moral and non-moral) in several domains. For instance, children are more likely than adults to endorse the idea that people’s physical appearance, intellectual ability, and moral characteristics will improve substantially over time (e.g., predicting that people will become better looking, smarter, and nicer with time, Heyman & Giles, 2004; Lockhart et al., 2008). Thus, domain-general age-related decreases in optimism may help explain why children—but not adults—reported that

“mean” individuals became “nicer” after incarceration and, separately, why only adults expected “nice” individuals’ moral character to worsen following punishment.

Alternatively, the *social input account* posits that cultural messages shape conceptual representations of carceral facilities over development, which, in turn, may alter views about the impact of incarceration on moral character. Prior work argues that people living in the United States primarily “form their impressions of crime and the criminal legal system based on what they hear, read, and see in the media” (Yousman, 2009, p. 1). Importantly, people living in the United States typically consume media that portray prisons and jails as dangerous, torturous, and rife with violence (Bennett, 2006; Yousman, 2009). Given that social input shapes conceptual development (e.g., Chalik et al., 2017; Gelman, 2009), such messages could influence people’s views about the impact of incarceration on moral character across age. Specifically, as children grow into adults, they may come to view carceral environments as less rehabilitative because of the negative messages they hear about incarceration (e.g., through the media). Thus, adults in the current work may have reported pessimism about the impact of incarceration on moral character because over time they have consumed—and subsequently accepted—media messages portraying carceral facilities as inhospitable to moral improvement.

Future work can test between these two candidate mechanisms by modifying the procedure to Study 2 to include both children and adults in the same experimental paradigm. As previously mentioned, some evidence suggests that age-related declines in optimism are domain-general (e.g., Boseovski, 2010); thus, adults’ relative pessimism about incarceration may also generalize to other types of punishment. Consistent with the *optimism account*, adults may be less likely than children to report that either incarceration or time-out improves moral character. However, consistent with the *social input account*, adults’ negativity may be specific to

incarceration (given that media transmit far more negative messages about incarceration than time-out). If this is the case, (1) adults may report more pessimism about the impact of incarceration on moral character than children, and (2) children and adults may report similarly optimistic views about the impact of time-out on moral character.

Another way to test between the two aforementioned potential mechanisms includes employing cross-cultural methods. Legal scholars have described criminal punishment in Europe as having a different “flavor” than criminal punishment in the United States. Unlike in the United States, criminal punishment in Europe “embraces ideals of rehabilitation and forgiveness” (Kleinfeld, 2016, p. 1035). Individuals growing up in Europe may attend to such cultural messages linking punishment with rehabilitation and positive change. In turn, European adults may view punishment as indicative of moral redemption. Thus, there may be greater developmental stability in reasoning about the impact of punishment on moral character in Europe than in the United States. Such a result would suggest that social input (e.g., media, cultural messages) shapes people’s concepts about what it is like to spend time in a carceral facility and, thus, provide evidence in favor of the *social input account*.

In addition to testing between the *optimism* and *social input accounts*, future work can examine which aspects of United States carceral facilities underlie adults’ pessimism. For example, the severity associated with incarceration may drive adults’ pessimism regarding this specific type of punishment. When addressing this topic, it may be particularly useful to manipulate severity in punitive contexts other than incarceration. Doing so may be particularly helpful considering that several factors could potentially influence participants’ responses to items about incarceration. For example, a future study can ask participants to reason about a long time-out (a relatively severe punishment) and, separately, a short time-out (a relatively minor

punishment). If adults are more pessimistic about the long (versus short) time-out, this may suggest that adults view punishment severity—a key component of incarceration—as inimical to moral improvement.

Finally, the current work is limited in that it probed adults' views of only one process—namely, severe punishment—as a means for moral improvement. We focused on severe punishment because this is a common response to perceived moral transgression, especially in the United States (e.g., Alexander, 2012). However, future research can examine whether less punitive responses to transgression improve perceived moral character. The current work suggests that adults are pessimistic about one particularly severe form of punishment (incarceration). However, it is unclear whether adults are pessimistic specifically about incarceration or about punishment in general. Recent scholarship provides initial evidence in favor of the former possibility, showing that 14- to 18-year-olds view relatively mild forms of punishment as effective in preventing recidivist behavior (Robichaud & Mageau, 2020). These findings thus provide initial evidence that, even beyond the elementary school years, people may view some forms of punishment as effective in catalyzing moral improvement.

Such a finding may have important practical implications. Converging lines of evidence suggests that severe punishment can negatively affect psychological and physical well-being. For instance, incarcerated individuals are up to ten times more likely than non-incarcerated individuals to experience depression, anxiety, and other trauma-based symptoms (Haney, 2012). If people believe that less punitive responses are also effective in improving moral character, they may be amenable to supporting criminal legal policies that champion alternative approaches to severe punishment. Thus, understanding perceptions of punishment could be an important component of criminal legal reform.

Conclusion

Uniting work on developmental psychology, moral cognition, and experimental jurisprudence, the current research examined children's and adults' views about the impact of punishment on moral character. Children reported that punishment—regardless of how severe—catalyzed positive moral change among “mean” individuals. This finding suggests that, at times, children's optimism can overwhelm their tendency to view moral character as immutable. Unlike children, adults expected that “nice” individuals' positive qualities worsen following punishment. Further, we did not find evidence that adults expected punishment to help “mean” individuals become “nicer.” These findings marshal evidence suggesting that people in the United States become increasingly pessimistic about the impact of punishment on moral character with age. In doing so, the current work suggests that adults living in the United States may believe that redemption is not for everyone, or, at the very least, that a specific form of severe punishment (incarceration) is not the way to achieve it.

References

- Alexander, M. (2012). *The new Jim Crow: Mass incarceration in the age of colorblindness*. New York, NY: The New Press.
- Banks, R. R., Eberhardt, J. L., & Ross, L. (2006). Discrimination and implicit bias in a racially unequal society. *California Law Review*, *94*, 1169-1190.
- Bennett, J. (2006). The good, the bad and the ugly: The media in prison films. *The Howard Journal of Criminal Justice*, *45*, 97-115. doi: 10.1111/j.1468-2311.2006.00408.x
- Bilz, K. (2016). Testing the expressive theory of punishment. *Journal of Empirical Legal Studies*, *13*, 358-392. doi: 10.1111/jels.12118
- Boseovski, J. J. (2010). Evidence for “rose-colored glasses”: An examination of the positivity bias in young children’s personality judgments. *Child Development Perspectives*, *4*, 212-218. doi: 10.1111/j.1750-8606.2010.00149.x
- Boseovski, J. J., & Lee, K. (2006). Children’s use of frequency information for trait categorization and behavioral prediction. *Developmental Psychology*, *42*, 500-513. doi: 10.1037/0012-1649.42.3.500
- Bregant, J., Shaw, A., & Kinzler, K. D. (2016). Intuitive jurisprudence: Early reasoning about the functions of punishment. *Journal of Empirical Legal Studies*, *13*, 693-717. doi: 10.1111/jels.12130
- Bronson, J., & Carson, E. A., (2019). *Prisoners in 2017*. Retrieved January 9th, 2020, from <https://www.bjs.gov/content/pub/pdf/p17.pdf>
- Chalik, L., Leslie, S. J., & Rhodes, M. (2017). Cultural context shapes essentialist beliefs about religion. *Developmental Psychology*, *53*, 1178-1187. doi: 10.1037/dev0000301
10.1007/s12152-018- 9393-4

- Deutchman, P., Bračić, M., Raihani, N., & McAuliffe, K. (In press). Punishment is strongly motivated by revenge and weakly motivated by inequity aversion. *Evolution and Human Behavior*, 1-25. doi: 10.1016/j.evolhumbehav.2020.06.001
- Dunlea, J. P., & Heiphetz, L. (2020). Children's and adults' understanding of punishment and the criminal justice system. *Journal of Experimental Social Psychology*, 87, 103913. doi: 10.1016/j.jesp.2019.103913
- Dunlea, J. P., & Heiphetz, L. (In press). Moral psychology as a necessary bridge between social cognition and law. *Social Cognition*.
- Dunlea, J. P., Wolle, R. G., & Heiphetz, L. (In press). Enduring positivity: Children of incarcerated parents report more positive than negative emotions when thinking about close others. *Journal of Cognition and Development*.
- Duff, R. A. (2011). Retrieving retributivism. In M. D. White (Ed.), *Retributivism: Essays on theory and policy* (pp. 3-24). Oxford, UK: Oxford University Press.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191. doi: 10.3758/BF03193146
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415, 137-140. doi: 10.1038/415137a
- Feinberg, J. (1965). The expressive function of punishment. *The Monist*, 49, 397-423.
- Finkel, N. J., Liss, M. B., & Moran, V. R. (1997). Equal or proportional justice for accessories? Children's pearls of proportionate wisdom. *Journal of Applied Developmental Psychology*, 18, 229-244. doi: 10.1016/S0193-3973(97)90037-X

- Forman, J., Jr. (2017). *Locking up our own: Crime and punishment in Black America*. New York, NY: Farrar, Straus, & Giroux.
- Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York, NY: Oxford University Press.
- Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology*, *60*, 115-140. doi: 10.1146/annurev.psych.59.103006.093659
- Ginott, H. G. (1965). *Between parent and child: New solutions to old problems*. Oxford, UK: Macmillan.
- Haney, C. (2012). Prison effects in the era of mass incarceration. *The Prison Journal*, *92*, 1-24. doi: 0032885512448604.
- Hart, H. L. A. (1959). Prolegomenon to the principle of punishment. *Proceedings of the Aristotelian Society*, *60*, 1-26. doi: 10.1093/acprof:oso/9780199534777.003.0001
- Heiphetz, L. (2019). Moral essentialism and generosity among children and adults. *Journal of Experimental Psychology: General*, *148*, 2077-2090. doi: 10.1037/xge0000587
- Heiphetz, L. (2020). The development and consequences of moral essentialism. *Advances in Child Development and Behavior*, *59*, 165-194. doi: 10.1016/bs.acdb.2020.05.006
- Heiphetz, L., Gelman, S. A., & Young, L. L. (2017). The perceived stability and biological basis of religious beliefs, factual beliefs, and opinions. *Journal of Experimental Child Psychology*, *156*, 82-98. doi: 10.1016/j.jecp.2016.11.015
- Heiphetz, L., Lane, J. D., Waytz, A., & Young, L. L. (2018). My mind, your mind, and God's mind: How children and adults conceive of different agents' moral beliefs. *British Journal of Developmental Psychology*, *36*, 467-481. doi: 10.1111/bjdp.12231

- Heyman, G. D., & Giles, J. W. (2004). Valence effects in reasoning about evaluative traits. *Merrill-Palmer Quarterly*, *50*, 86-109. doi: 10.1353/mpq.2004.0004
- Ho, M. K., Cushman, F., Littman, M. L., & Austerweil, J. L. (2019). People teach with rewards and punishments as communication, not reinforcements. *Journal of Experimental Psychology: General*, *148*, 520–549. doi: 10.1037/xge0000569
- Hussak, L. J., & Cimpian, A. (2018). Investigating the origins of political views: Biases in explanation predict conservative attitudes in children and adults. *Developmental Science*, *21*, e12567. doi: 10.1111/desc.12567
- Hussak, L. J., & Cimpian, A. (2019). “It feels like it’s in your body”: How children in the United States think about nationality. *Journal of Experimental Psychology: General*, *148*, 1153–1168. doi: 10.1037/xge0000567
- Kahan, D. M. (1996). What do alternative sanctions mean? *The University of Chicago Law Review*, *63*, 591-653. doi: 10.2307/1600237
- Kleinfeld, J. (2016). Two cultures of punishment. *Stanford Law Review*, *68*, 933-1036.
- Lakens, D., & Evers, E. R. K. (2014). Sailing from the seas of chaos into the corridor of stability: Practical recommendations to increase the informational value of studies. *Perspectives on Psychological Science*, *9*, 278-292. doi: 10.1177/1745691614528520
- Lockhart, K. L., Nakashima, N., Inagaki, K., & Keil, F. C. (2008). From ugly duckling to swan?: Japanese and American beliefs about the stability and origins of traits. *Cognitive Development*, *23*, 155-179. doi: 10.1016/j.cogdev.2007.08.001
- Mageau, G. A., Lessard, J., Carpentier, J., Robichaud, J. M., Joussemet, M., & Koestner, R. (2018). Effectiveness and acceptability beliefs regarding logical consequences and mild

- punishments. *Journal of Applied Developmental Psychology*, 54, 12-22. doi: 10.1016/j.appdev.2017.11.001
- Markel, D. (2011). What might retributive justice be? In M. D. White (Ed.), *Retributivism: Essays on theory and policy* (pp. 3-24). Oxford, UK: Oxford University Press.
- McAuliffe, K., Jordan, J. J., & Warneken, F. (2015). Costly third-party punishment in young children. *Cognition*, 134, 1-10. doi: 10.1016/j.cognition.2014.08.013
- Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179-195). New York, NY: Cambridge University Press.
- Moore, K. E., Stuewig, J. B., & Tangney, J. P. (2016). The effect of stigma on criminal offenders' functioning: A longitudinal mediational model. *Deviant Behavior*, 37, 196-218. doi: 10.1080/01639625.2014.1004035
- Murphy, J. G., & Hampton, J. (1988). *Forgiveness and mercy*. New York, NY: Cambridge University Press.
- Nadler, J. (2012). Blaming as a social process: The influence of character and moral emotion on blame. *Law and Contemporary Problems*, 75, 1-31.
- Nadler, J., & McDonnell, M. H. (2011). Moral character, motive, and the psychology of blame. *Cornell Law Review*, 97, 255-304.
- Nelsen, J. (1985). The three R's of logical consequences, the three R's of punishment, and the six steps for winning children over. *Individual Psychology*, 41, 161-165.
- Okimoto, T. G., & Wenzel, M. (2011). Third-party punishment and symbolic intragroup status. *Journal of Experimental Social Psychology*, 47, 709-718. doi: 10.1016/j.jesp.2011.02.001

- 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*, e12509. doi: 10.1111/desc.12509
- Rhodes, M., & Mandalaywala, T. M. (2017). The development and developmental consequences of social essentialism. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*, e1437. doi: 10.1002/wcs.1437
- Richardson, L. S., & Goff, P. A. (2013). Implicit racial bias in public defender triage. *Yale Law Journal, 122*, 2626-2649.
- Roberts, S. O., & Gelman, S. A. (2017). Multiracial children's and adults' categorizations of multiracial individuals. *Journal of Cognition and Development, 18*, 1-15. doi: 10.1080/15248372.2015.1086772
- Robichaud, J. M., & Mageau, G. A. (2020). The socializing role of logical consequences, mild punishments, and reasoning in rule-breaking contexts involving multifaceted issues. *Social Development, 29*, 356-372. doi: 10.1111/sode.12395
- Salali, G. D., Juda, M., & Henrich, J. (2015). Transmission and development of costly punishment in children. *Evolution and Human Behavior, 36*, 86-94. doi: 10.1016/j.evolhumbehav.2014.09.004
- Shaw, A., & Olson, K. (2015). Whose idea is it anyway? The importance of reputation in acknowledgement. *Developmental Science, 18*, 502-509. doi: 10.1111/desc.12234
- Simmons, J., Nelson, L., & Simonsohn, U. (2013). *Life after p-hacking*. Paper presented at the annual meeting of the Society for Personality and Social Psychology. New Orleans, LA.

- Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys will be boys; cows will be cows: Children's essentialist reasoning about gender categories and animal species. *Child Development, 80*, 461-481. doi: 10.1111/j.1467-8624.2009.01272.x
- Travis, J. (2005). *But they all come back: Facing the challenges of prisoner reentry*. Washington, DC: The Urban Institute.
- Travis, J., Western, B., & Redburn, S. (Eds.). (2014). *The growth of incarceration in the United States: Exploring causes and consequences*. Washington, DC: National Academies Press.
- Turton, S. (2018). *The 7 ½ deaths of Evelyn Hardcastle*. London, UK: Bloomsbury Raven.
- Twum-Danso Imoh, A. (2013). Children's perceptions of physical punishment in Ghana and the implications for children's rights. *Childhood, 20*, 472-486. doi: 10.1177/0907568212471404
- Van Cleve, N. G. (2016). *Crook County: Racism and injustice in America's largest criminal court*. Stanford, CA: Stanford University Press.
- Yankah, E. N. (2004). Good guys and bad guys: Punishing character, equality and the irrelevance of moral character to criminal punishment. *Cardozo Law Review, 25*, 1019-1067
- Yousman, B. (2009). *Prime time prisons on U.S. TV: Representation of incarceration*. New York, NY: Peter Lang Publishing, Inc.

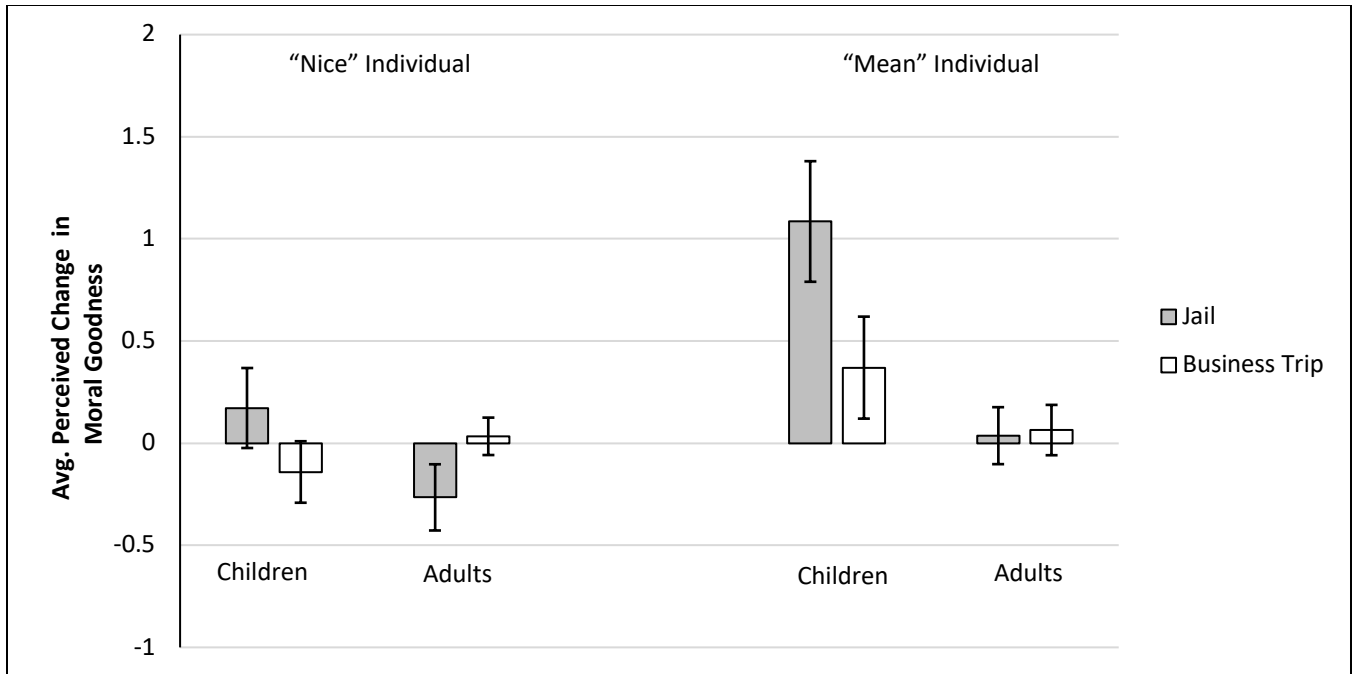


Figure 1. Average perceived change in moral goodness, Study 1. Positive values reflect perceived increases in moral goodness, whereas negative values reflect perceived decreases in moral goodness. Zero indicates no perceived change in moral goodness. Error bars represent 95% confidence intervals.

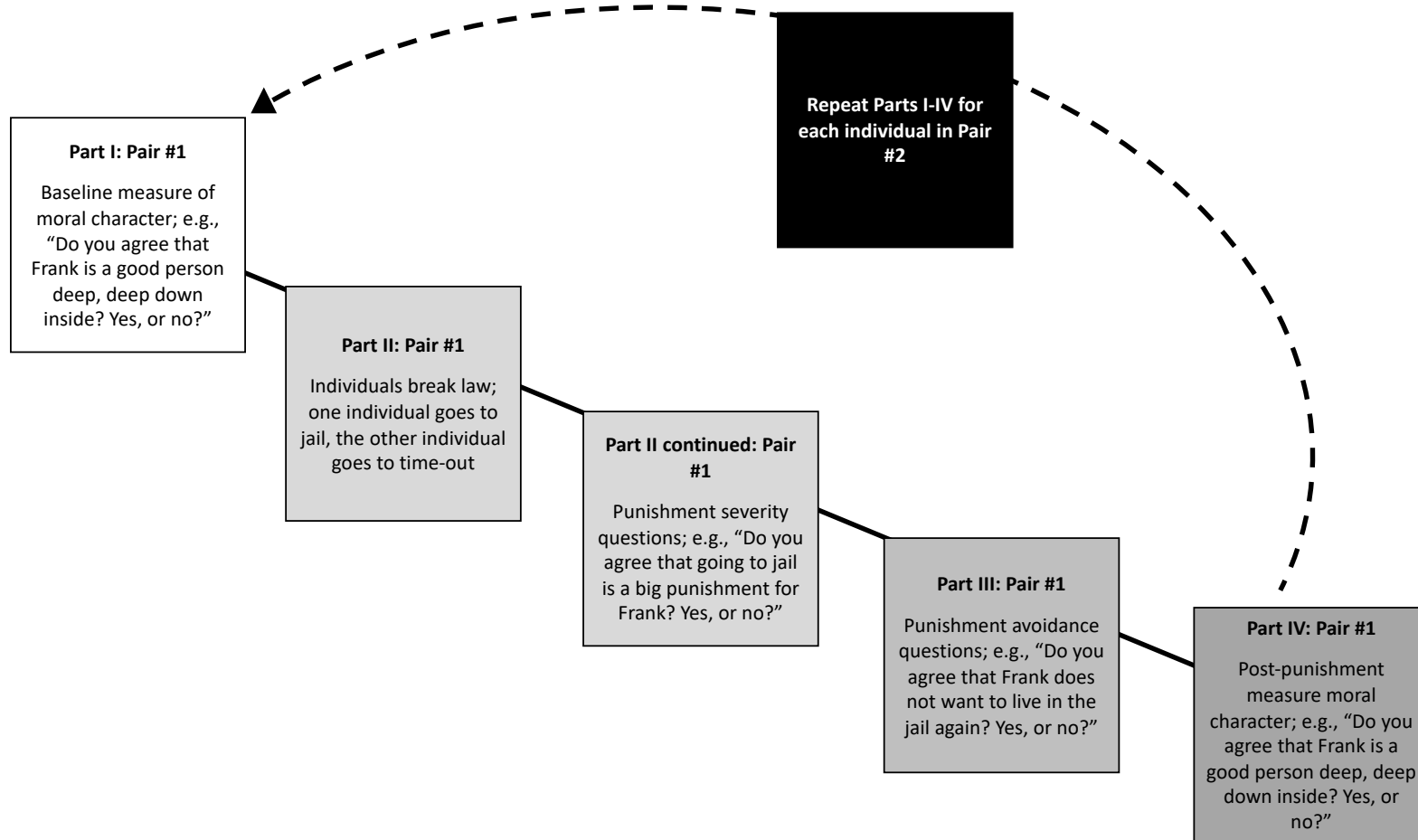


Figure 2. A schematic representation of Study 2’s procedure. As noted in Study 2’s procedure, participants finished responded to items in each block before moving on to the next block; all participants responded to items in each block in the order depicted above.

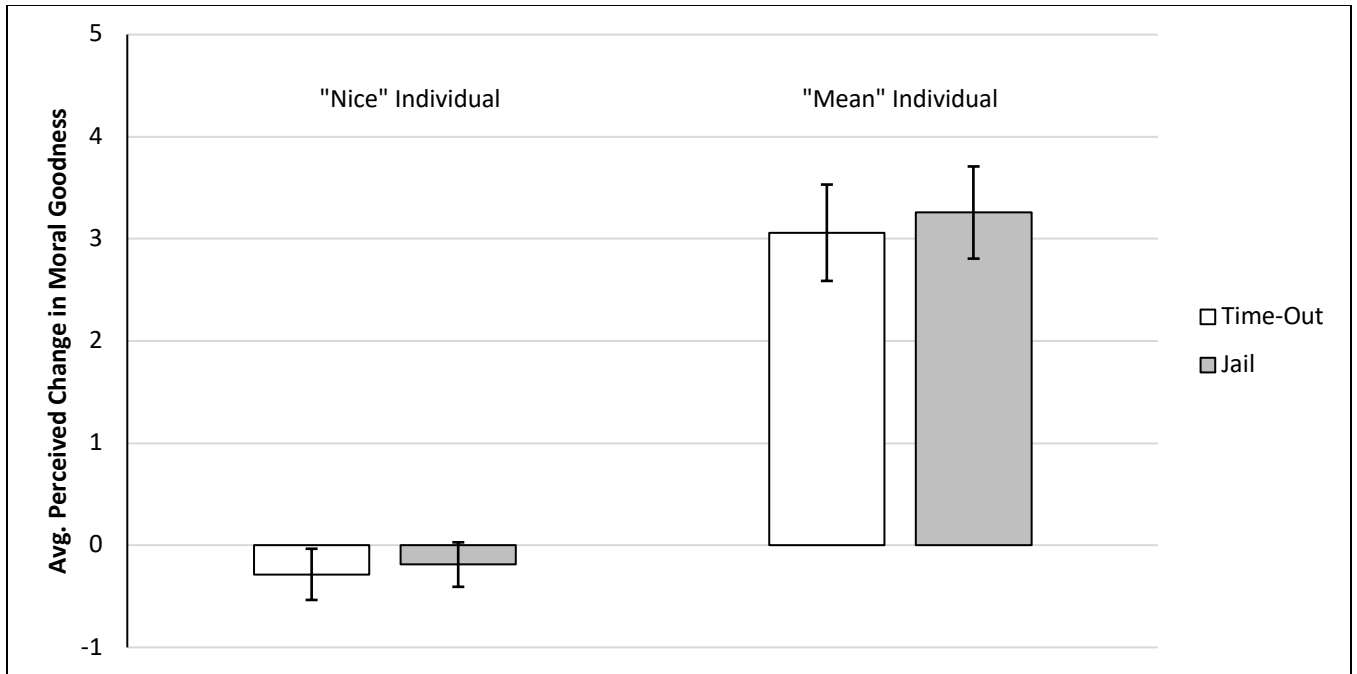


Figure 3. Average perceived change in moral goodness, Study 2. Positive values reflect perceived increases in moral goodness, whereas negative values reflect perceived decreases in moral goodness. Zero indicates no perceived change in moral goodness. Error bars represent 95% confidence intervals.