Psychology, Public Policy, and Law

1076-8971/11/$12.00 DOI: 10.1037/a0025821

PUNISHMENT WITHOUT REASON: Isolating Retribution in Lay Punishment of Criminal Offenders

Eyal Aharoni
University of New Mexico and The Mind Research Network

Alan J. Fridlund
University of California, Santa Barbara

Research has suggested that criminal punishment decisions are driven primarily by retribution and that retributive judgments are achieved by a process of abstract moral reasoning. However, problems with construct validity limit confidence in these conclusions. Study 1 ($N = 254$) used experimentally manipulated vignettes to isolate retributive motives. Participants' sentencing recommendations were strongly provoked by indices of retribution (criminal intent) even when the most common consequentialist reasons for punishment (offender dangerousness and publicity of punishment) were minimized. In an exploratory fashion, Study 2 ($N = 49$) used a semistructured interview to examine whether participants would persist in punishing a hypothetical offender without explicit reasons. Participants persisted in their original punishment judgments even when unable to justify the reasons for these judgments. These results increase confidence that lay punishment is motivated by retribution, but also suggest that this motive may be better explained by fallible, heuristic processes than by abstract moral reasoning. Implications for legal policy are discussed.

Keywords: retribution, deterrence, punishment, sentencing, intention, moral reasoning, heuristic

It has long been assumed that prison deters criminal offenders. However, recent research in the forensic sciences has shown that, contrary to popular belief, prison may in fact increase the likelihood that offenders will engage in further crime after they are released (Chen & Shapiro, 2007; Cid, 2009; Grogger, 1991). This counterintuitive effect has even been observed experimentally, whereby offenders with similar sentences and criminal histories were arbitrarily assigned to different security levels. Those assigned to higher security levels were likelier to return to prison on a new violation after release (Chen & Shapiro, 2007). When incarceration is not just ineffective but even criminogenic, such findings force us to question why we desire certain forms of criminal punishment when doing so might actually occasion greater postrelease risk. Do we always have good reasons for persisting with such forms of punishment? Answers to this question could assist voters, legislators, and jurors who must evaluate whether the benefits gleaned from criminal punishment justify a potential increased risk to society.

Eyal Aharoni, Department of Psychology, University of New Mexico & The Mind Research Network, Albuquerque, New Mexico, Alan J. Fridlund, Department of Psychology, University of California, Santa Barbara. AQ: 2

We thank Michael Gazzaniga, Daphne Bugental, Aimee Libeu, Jonathan Haidt, Nick Stadlberger, Walter Sinnott-Armstrong, Andrew Delton, and members of the MacArthur Foundation's Law & Neuroscience Project, the Center for Evolutionary Psychology.

Correspondence concerning this article should be addressed to Eyal Aharoni, The Mind Research Network, 1101 Yale Boulevard Northeast, Albuquerque, NM 87106. E-mail: eaharoni@unm.edu

tapraid5/zdn-law/zdn-law/zdn00411/zdn2100d11z | xppws | S=1 | 10/5/11 | 2:37 | Art: 2011-0363

Deontological Versus Consequentialist Punishment

There are many possible reasons for seeking criminal punishment. These tend to fall into one of two categories: deontological and consequentialist. The deontological motive for punishment is often referred to as retribution, just deserts, or moralistic punishment. This form of punishment is considered deontological because its proponents consider it morally obligatory and self-justified. Emanuel Kant famously defended this stance as a "categorical imperative" to punish—not for any practical end-goal, but for its own sake (1785/1998). The severity of deontological punishment is presumed to be measured against offenders' moral blameworthiness for their past act, regardless of their future dangerousness. Later theorists articulated several varieties of this basic argument, most of which depend on a basic desire to restore a moral balance of some kind (Cottingham, 1979; Packer, 1968).

The consequentialist motive for punishment, first promulgated by Beccaria (1764/1986) and Bentham (1789/1988), is a practical approach to protect social welfare by attempting to control future behavior, regardless of the issue of moral guilt. The primary consequentialist goals are incapacitation (i.e., to physically control behavior), specific deterrence (i.e., to disincentivize the offender from recidivating), and general deterrence (i.e., to disincentivize society at large from engaging in criminal activity (Miller & Vidmar, 1981; Packer, 1968; Vidmar, 2000).[1] Fn1

In Western law, both consequentialist and deontological motives for punishment play prominent roles in legislative policies on how judges and jurors ought to formulate criminal punishment decisions (e.g., Model Penal Code § 1.02(2), 1962). However, psychologically, each motive should be responsive to different conditions, and therefore, should yield different degrees of punishment depending on the particular factors under consideration.

Evidence on Punitive Motives

Although legislative doctrine typically permits both motives for criminal punishment, experimental research has made a consistent case that retribution plays a stronger role in actual punishment decisions (Bohm, 1992; Carlsmith, Darley, & Robinson, 2002; Darley, Carlsmith, & Robinson, 2000; Hamilton & Rytina, 1980; Harlow, Darley, & Robinson, 1995; McFatter, 1982; Rucker, Polifroni, Tetlock, & Scott, 2004; Tyler & Boeckmann, 1997; Warr et al., 1983).This pattern may be particularly true of more violent crimes (e.g., Viney, Parker–Martin, & Dotton, 1988).

In one highly representative study, Darley and colleagues made an earnest attempt to distinguish retributive motives for punishment from the consequentialist motive to incapacitate offenders (Darley et al., 2000). They employed the classic method for this type of research, asking participants to make sentencing judgments about hypothetical offenders, in a fashion not so different from the

[1] The rehabilitative motive is also often recognized as an independent consequentialist motive for punishment. However, no such distinction was made for the purposes of this study because of its functional similarity to specific deterrence.

penalty phase of actual criminal trials (albeit that in the real world, judges have to abide by numerous, sometimes conflicting, punishment guidelines). Case scenarios were designed to independently rouse the sentiments of just deserts (retribution) and the consequentialist motive of incapacitation. Retribution was theorized to be evoked by manipulating the moral seriousness or gravity of the offense (high vs. low). Incapacitation was theorized to be evoked by manipulating the offender's stipulated future dangerousness (high vs. low). These manipulations were matched orthogonally in separate vignettes. These authors found that punishments were most strongly evoked by the crime seriousness manipulation, not the dangerousness manipulation, and so they concluded that their findings substantiated the primacy of retribution in lay punishment decisions (Darley et al., 2000).

Other authors have reported similar conclusions. For example, Carlsmith and colleagues (2002) showed that when a hypothetical transgressor was coincidentally stricken by paralysis after committing a wrongful act, participants remained unsatisfied and persisted in punishment. This suggests that another motive for punishment (possibly, retribution) was at work when the motives to deter or incapacitate the offender were abrogated.

A recent attempt to isolate retribution was reported in a study by Keller and colleagues (Keller, Oswald, Stucki, & Gollwitzer, 2010), in an extension of Carlsmith, 2006. Participants had to select from a list the questions that would need to be answered about a particular criminal scenario in order to arrive at a satisfying sentencing decision. Different stimuli questions had been furnished to represent each of the primary motives for punishment (e.g., retribution, incapacitation, specific deterrence, and general deterrence). For instance, the crime's "magnitude of harm" (i.e., "How big is the financial, physical, and psychological harm the offender has caused committing the crime?") was preclassified as a good example of a retributive concern. These preclassifications had been determined using data from a pilot sample's self-report. The items most frequently selected by the experimental sample were from the list that had been previously classified as retributive.

Other studies have reported similar results in support of retribution (Bohm, 1992; Hamilton & Rytina, 1980; Harlow et al., 1995; McFatter, 1982; Rucker et al., 2004; Tyler & Boeckmann, 1997; Warr et al., 1983).

Limitations of Previous Research

Despite their many strengths, these studies have faced serious challenges deconfounding retribution from the consequentialist motives. Consider, for instance, Darley et al.'s conclusion that the punishment of increasingly serious crimes demonstrates retribution (2000). This inference fatally ignores the fact that people motivated by the practical desire to deter or incapacitate the offender should also focus their greatest effort on preventing the most serious of crimes (See Warr, Meier, & Erickson, 1983). This is because crime seriousness is quite informative about crime costliness (e.g., degree of physical, emotional, and financial damages), which, theoretically, is of great concern for consequentialists even when dangerousness is fairly low. So, to the extent that increased crime seriousness inspires increased punishment, this could just as well signify evidence

of consequentialist motives (Aharoni, 2009). It is impossible to exclude this interpretation from the pattern of punishment observed in the Darley et al. study.

Studies like Carlsmith et al. (2002) have circumvented this problem by showing that participants desire punishment even when the offender is paralyzed, ruling out the consequentialist motive of incapacitation. Nonetheless, such a design cannot rule out other consequentialist concerns such as that of general deterrence.

Keller et al. (2010) attempted to overcome some of these limitations by distinguishing retribution from all the primary consequentialist motives. However, the stimuli that were fashioned to indicate the separate motives for punishment were ultimately based on the verbal reports of pilot participants and were never independently manipulated. For example, without a careful experimental manipulation, it is impossible to distinguish study participants' "retributive" concerns about the magnitude of harm from the consequentialist motive to minimize harm. Therefore, it remains possible that apparent evidence for one motive in fact served to satisfy other motives.

Other studies have faced similar difficulties in isolating the motives for punishment, largely due to problems in construct validity and reliance on self-report (e.g., Bohm, 1992; Hamilton & Rytina, 1980; McFatter, 1982; Rucker, Polifroni, Tetlock, & Scott, 2004; Tyler & Boeckmann, 1997; Warr et al., 1983). Thus, despite early confidence in the primacy of retribution, consistent methodological problems reveal a continued demand to demonstrate that retributive punishment does in fact operate in the absence of practical reasons for punishment.

Rational Versus Nonrational Punishment

Research in motives for punishment faces another challenge besides simply determining whether lay punishment is better characterized by deontological or consequentialist punishment. This second challenge is to determine whether moralistic punishment judgments are typically conceived by way of explicit rational deliberation or instead by nonrational, automatic, heuristic processes. For rationalists like Kant, moral judgment was an inviolate logical deductive chain (see Kant, 1790/1952). Arguments that the practice of retributive punishment may be a rational form of jurisprudence also appear in modern scientific discourse. For instance, researchers have argued that because lay people punish retributively, criminal legislation should similarly endorse retribution so as not to alienate its citizens and risk reactions of vigilantism (e.g., Darley et al., 2000).

The rationalist view of moral reasoning stands in stark contrast with long-standing findings by social and cognitive scientists, however, who argue that human reasoning is often better characterized by nonrational heuristic processes than by rational deduction. Heuristics are simple cognitive rules that guide everyday decision making. They occur relatively unconsciously, quickly, and automatically. As such, they tend to afford us a quick and dirty alternative to complex reasoning and deliberation, and tend to serve our immediate needs in many ordinary situations. However, they often lack the flexibility and innovation characteristic of higher-order reasoning processes (see Gigerenzer, Todd, & the ABC Research Group, 1999).

For instance, the availability heuristic describes the tendency to judge the frequency of an event based on the ease with which an example can be brought to mind (Tversky & Kahneman, 1974). Sometimes what comes to mind first is also the most frequent. For example, most Americans can rely on their recent past experience to correctly judge that the name Mary is more common in their home country than the name Mohammed. However, this rule would be incorrect if used to judge the global prevalence of these two names.

Another feature of automatic judgments is that they can be resistant to change. For example, studies of belief preservation have shown that we tend to adhere to preexisting folk theories (e.g., that risk takers make better fire fighters) more strongly in the face of evidence to the contrary, even when that evidence is factually overwhelming (e.g., Lepper, Anderson, & Ross, 1980). In this case, continued belief in the proposition is not strictly rational because the proposition is knowingly tenuous. Persistence, therefore, is held to signify heuristic processing. In this view, explanations for one's decisions are simply post hoc rationalizations (see Gazzaniga, 1998; Gazzaniga, 2000).

With regard to moral decisions in particular, does explicit reasoning and deduction guide our punitive judgments, or could it be that our moral judgments come first and practical reasons are fabricated retroactively in attempt to interpret those judgments (Haidt, 2001)? If the latter, we would be burdened to question whether those intuitions are always in society's best interest.

Current research suggests that noninstrumental heuristic processing frequently guides moral decision making, as demonstrated by studies of the well-known trolley problem. For instance, Greene and colleagues (Greene, Nystrom, Engell, Darley, & Cohen, 2004; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001) showed that when deciding whether to kill one innocent bystander in order to save five others, study participants are far less likely to do so when they are required to use personal contact to kill the bystander rather than a mechanical switch, even though the outcomes are equivalent. Yet, when confronted about this inconsistency, participants are often unable to articulate a rationale.

Another provocative demonstration of heuristic moral reasoning derives from research on "moral dumbfounding" (see Haidt, Bjorklund, & Murphy, 2000; Haidt, 2001). In this research, adult participants were given scenarios that most people consider to be wrong, such as consensual incest among adult siblings. The scenarios were designed to circumvent ready objections (e.g., the act was consensual, effective protection was used, the siblings had no negative feelings afterward, and the act was never repeated or disclosed).

The interviewer then delivered the key question: Is this act still wrong? A large majority of the participants answered affirmatively. The interviewer then challenged them further to explain why, given that their objections were all inapplicable. Participants typically found it difficult to defend their condemnation via logical argument, resorting instead stating some version of "Gosh, I can't explain it—it just seems wrong!" Only about 20% of participants concluded that the act was permissible, mostly by virtue of the act's benign consequences. Continued adherence to the original moral rule is noninstrumental and not strictly rational in this context because the reasons for the rule no longer apply. The

authors concluded that this persistence of judgment without articulable reasons was evidence of a moral intuition—in this case, an intuition to avoid incest.[2] Fn2

These considerations—evaluating the degree to which punishment judgments are both retributive and heuristic—gave rise to the present studies. In Study 1, we sought to improve the construct validity of the measurement of retribution by addressing two methodological challenges.

Study 1 Approach and Hypotheses

The first challenge was to design a manipulation that could evoke retribution relatively free of consequentialist concerns. In his *Groundwork of the Metaphysics of Morals*, Kant argued, in determining what is morally good or bad, an actor's intention was more important than the consequences of his action (1785/1998). In modern criminal law, the offender's intent often plays a major role in determinations of culpability. Legal theory commonly defines intent as "the mental resolution or determination to do [an act]" (Garner, 2004). Similarly, cognitive scientists have defined an intention as a mental state responsible for translating goals into actions (see Haggard, 2005). Research has shown that beliefs about another's intention to harm, beyond any harmful consequences of his actions, best predict aggressive retaliation and punishment, at least for more violent offenses (Baron & Richardson, 1994, pp. 145, 210; Blum West, 1985; Geen, 1968; Greenwell & Dengerink, 1973; Surber, 1977; White, 1975). A similar pattern is seen in the tendency to assign greater liability, under certain conditions, to offenders of attempted crimes (Robinson & Darley, 1995, pp. 14–28). This tendency to base moral judgments and punishments on the nature of the offender's intentions appears to develop at a very early age (Costanzo, Coie, Grumet, & Farnill, 1973; Darley, Klosson, & Zanna, 1978; Karniol, 1978; Piaget, 1948).

In naturalistic settings, an offender's intentions could cue not just his culpability, but also his future dangerousness (relevant to specific deterrence). However, this possibility can be experimentally controlled by designing scenarios that involve malicious intent but not future danger, and inversely, future danger without malicious intent—all independent of the crime's costliness. These properties, however, are not true of previous manipulations of crime seriousness; these, as we argued above, cannot be easily dissociated from a crime's costliness.

Designing scenarios that involve future danger without malicious intent is no easy task. This is because purely unintentional (i.e., accidental) offenses are not likely to recur, and may not even qualify as offenses in courts of law. This obstacle may explain why some researchers have operationally defined low intent in terms of mental illness (e.g., Darley et al., 2000). Although mental illness is admittedly a poor proxy for low intent, the primary purpose of this manipulation was simply to characterize a condition in which the mental state leading to the offense would mitigate punishment in a psychological sense, albeit at the cost of

[2] We use the term "intuition" in the sense used by Haidt and colleagues—namely, a nonrational, nondeliberate, unconscious, inclination motivating judgment. This is not to suggest that such processes lack usefulness or an underlying rationale. To the contrary, common intuitions likely served to solve genuine problems, just in the way that intuitions to avoid incest likely lessened genetic complications resulting from pathogen load.

legal realism. Indeed, mental illness has generally been found to be mitigating in courts (Monahan & Hood, 1976). Thus, for purposes of replication and experimental control, we chose to uphold this practice despite its obvious limit in ecological validity.

The second challenge for researchers of retribution is to design a manipulation that can independently modulate the desire to deter a general audience. One way to remove this "general deterrence" confound is to vary independently how public or private the punishment will be. In one study, conducted by Baron and Ritov (1994), participants assigned civil penalties to a company found liable for a negligent act. Participants fined the company above and beyond what would be needed for full compensation even when the company was going out of business, eliminating the need for specific deterrence, and even when all information about the penalty was kept private, ruling out general deterrence concerns.

For our purposes, we wished to deal similarly with the general deterrence confound but in the context of criminal behavior. We achieved this by stipulating in the "private" condition that the victim of the crime was 17 years old at the time the case was filed, and state law does not hold public trials or publish hearings, transcripts, or sentences that involve minors under the age of 18. Although such legal actions admittedly are historically rare, this manipulation was developed at the expense of maximal realism because of the experimental control it afforded. As with the low intent manipulation, its credibility can be assessed using standard manipulation checks.

An additional way to clarify the unique contribution of retribution could be to examine whether the desire to make the offender suffer mediates the effect of intent on punishment. Of importance, the induction of suffering is not by itself sufficient to indicate retribution because suffering can also be used instrumentally to control an offender's future behavior. However, this corporeal component of punishment has also been integral to retributive punishment (Bohm, 1992; Cottingham, 1979). Thus, if independent evidence for retribution is observed, showing that this effect is mediated by the desire for suffering would serve to confirm the importance of this aspect of retributive punishment.

Study 1 evaluated the hypothesis that (a) punishment recommendations will track criminal intent even when the two consequentialist concerns are minimized, and that (b) more intentional offenses will warrant greater suffering. If so, this would lend needed support to the argument that the punishment of criminal offenders under such conditions is frankly retributive. In an effort to assess the external validity of our student sample, we also sought to examine whether politically liberal versus conservative participants (determined by self-report) differ in the leniency of their punishment decisions, with liberal students favoring leniency.

Study 2 Approach and Hypothesis

In Study 2, we sought to explore whether the construction of retributive punishment judgments is better characterized by rational deduction or nonrational, heuristic processes. Drawing on the work of Haidt, Carlsmith, and others, this study examined whether typical adults tend to retract their support for punishment

after the common practical reasons to punish (e.g., specific deterrence, general deterrence) have been discredited.

To achieve this, a vignette describing an egregious crime was designed such that the common reasons for punishing someone were inapplicable to the scenario. We measured whether participants would nonetheless recommend punishment for the offender. Because the practical motives for punishment can take so many forms, responses were probed using a semistructured interview format. At this early, exploratory phase, this design enabled us to minimize any assumptions about construct validity inherent to experimental manipulations.

Study 2 evaluated the hypothesis that a proportion of participants greater than chance would favor punishment—as opposed to release—even when disarmed of their justifications.[3] If so, this persistence could provide an initial suggestion that these motives for punishment are driven more by nonrational heuristic processes than by rational deliberation. Fn3

Study 1: Isolating Retribution: The Roles of Intent, Dangerousness, and Publicity in Punishment

Method

Participants

Participants consisted of 254 (71 male and 183 female) psychology students from University of California, Santa Barbara (UCSB) who received introductory psychology course credit for their participation. They were age-qualified for U.S. jury selection, with ages ranging from 18–24. Other demographic information collected included ethnicity (Caucasian: 54.3%; Hispanic: 16.5%; Asian: 15.7%; African American: 3.1%; Other: 7.1%; No answer: 3.1%), religious affiliation (Christianity: 26.8%; Catholicism: 21.3%; Judaism: 8.7%; Buddhism: 4.3%; Other: 2.8%; None: 31.9%; No answer: 4.3%), church attendance (mean [*M*] = 12.1 times per year; standard deviation [*SD*] = 21.4), political party affiliation (Democratic: 66.1%; Republican: 15.4%; Independent: 7.5%; Libertarian: 4.7%; Other: 6.3%), and political conservatism ($M = 4.41$; $SD = 2.27$) where 1 = *very liberal* and 11 = *very conservative*. It is noteworthy that this trend toward liberal attitudes is significantly different from the 5.5 neutral point on the liberal–conservative scale, $t(253) = -7.68$, $p < .001$. Seven participants reported having served jury duty.[4] Fn4

Design

We employed a 2 × 2 × 2 between-subjects factorial design that crossed intent, dangerousness, and publicity of punishment in a mock case about a violent crime. Intent was characterized as a homicide that was either performed willfully and deliberately or as a result of mental illness caused by a brain tumor (adapted

[3] The chance performance criterion is defined in the Design section.

[4] Thirty-eight participants reported having a family member who had served as a judge or lawyer. Ninety-eight said they or family member have been the victim of violent crime. Thirty-seven said they or family have been prosecuted for a crime. None of these factors significantly influenced the magnitude of punishment recommended.

from Study 2 of Darley et al., 2000). Dangerousness was characterized by expert testimony that the offender either was extremely dangerous or not at all.

The publicity of punishment was manipulated by stipulating one of two conditions. Condition 1 stipulated that the victim of the crime was 17 years old at the time the case was filed, and state law does not hold public trials or publish hearings, transcripts, or sentences that involve minors under the age of 18. In case participants' perceptions of privacy would be obscured by the mere presence of the experimenter, participants in this condition were also assured that any identifying information about their responses would be kept private from the study's experimenters.

For Condition 2 of the publicity variable, the age of the victim was held constant, but the courtroom was described as containing a full audience including local media, and participants were told that the hearing, transcript, and sentence would be made publicly available. Furthermore, participants in this condition were told that their responses would be (confidentially) identified, read, and evaluated by the study's experimenters.

In both publicity conditions, the instructions stated that the case called for a bench trial whereby the participant would act as sole judge. Participants were run individually in private rooms. Extraneous social audience primes such as mirrors and photographs of people were removed from the participant's laboratory environment.

Measures

Eight different vignettes—one for each condition—described a unique case about an offender stipulated to have committed a homicide. The vignettes and the punishment scales were modeled after those reported in previous studies (Darley et al., 2000 and Carlsmith et al., 2002). The primary scales of interest were sentence length ("What do you think is an appropriate sentence for this offense?" 1 = "0 days," 2 = "1 day," 3 = "2 weeks," 4 = "2 months," 5 = "6 months," 6 = "1 year," 7 = "3 years," 8 = "7 years," 9 = "15 years," 10 = "30 years," 11 = "life"), and two Likert-type scales measuring punishment severity ("How severe a punishment should be given for this offense?") and offender suffering ("How much should the offender should suffer for this offense?"), from 1 (*not at all*) to 11 (*extremely*). We excluded a death sentence option at the end of the sentencing scale out of concern that this personally and politically loaded issue might distort participants' responding to the remainder of the scale. Although the sentencing scale does not conform to an interval level, similar scales have reliably demonstrated a strong positive correlation with the severity measure, and have yielded no important differences from it (Darley et al., 2000 and Carlsmith et al., 2002). Therefore, the sentencing scale is reported below for purposes of replication and to assist in the interpretation of the severity scale results.

Materials and Apparatus

Laboratory cubicles and desktop computers were provided for participants during the experiment. All instructions, experimental stimuli, and measures were delivered by a text-based computer program. (See Appendix A for stimuli). Manipulation checks were also conducted to assess participants' belief in the

manipulations using similar 11-point scales (e.g., "To what extent do you think this offender intended to commit the crime?", "How likely do you think this offender is to commit a similar crime in the future?", "How much did you think your sentencing decision was going to be public record?", "How morally wrong do you think this crime was," and a composite measure of the crime's costliness, comprised of how "physically debilitating," "emotionally traumatic," and "financial costly" the crime was).

Procedure

Participants each read only one vignette according to their experimental condition. They then answered questions on a computer regarding sentence recommendations, manipulation checks, and participants' demographic characteristics.

Reliability

We examined the correlation between the sentence length and punishment severity measures. These were highly positively correlated, one-tailed $r = .77$, $p < .0001$. Reliability analysis further confirmed that these two measures were closely related, Cronbach's alpha = .87.

Results

Manipulation Checks

First, we sought to determine whether the three manipulations were effective. Using a three-way analysis of variance (ANOVA), participants indeed considered the "high-intent" crime as more intentional than the "low-intent" crime, $M_{high} = 8.63$ (2.47), $M_{low} = 5.06$ (2.46), $F(1, 247) = 132.96$, $p < .001$, $\eta_p^2 = .35$. Similarly, they rated the "high-dangerousness" offender more dangerous than the "low-dangerousness" offender, $M_{high} = 8.35$ (2.47), $M_{low} = 5.06$ (2.46), $F(1, 247) = 111.35$, $p < .001$, $\eta_p^2 = .31$. Finally, they were more confident that their punishment would be publicly disclosed when so stipulated by the publicity manipulation, $M_{high} = 8.46$ (2.78), $M_{low} = 4.60$ (2.78), $F(1, 247) = 123.15$, $p < .001$, $\eta_p^2 = .33$. Those in the "low publicity" condition also reported that they believed courts retain the authority to withhold information about the trial in order to protect a minor's identity, as shown by a one-sample t test, $M = 6.98$ (2.85); $t(126) = 5.86$, $p < .001$. As might be expected, an interaction was also observed in which participants rated the "high-intent" offender as somewhat more dangerous than the "low-intent" offender, $M_{high} = 7.49$ (3.25), $M_{low} = 6.18$ (3.21), $F(1, 247) = 20.87$, $p < .001$, $\eta_p^2 = .08$. This might potentially have complicated the interpretation of any punishments that track criminal intent. However, this interaction is substantially smaller than the effect of "intent" on perceived intent. Furthermore, as we show below, participants' punishment recommendations are not directly influenced by the dangerousness manipulation. Taken together, any ambiguities in the interpretation of the effects of intent on punishment should be better explained by evoked perceptions of intent rather than perceptions of dangerousness. No other interactions were found.

To examine whether criminal intent overlaps with the perceived costliness of the crime to the degree that the crime's perceived moral wrongness does, measures of these constructs were subjected to a simultaneous linear regression, with perceived intent and wrongness as independent variables. As predicted, moral wrongness explained a significant proportion (25.2%) of the variance perceived costliness of the crime, $\beta = .25$, $t(253), = 3.69$, $p < .001$, but criminal intent did not, $\beta = .02$, $t(253), = .27$, $p = .78$. (overall $R^2 = .06$, $p < .001$.) This asymmetry validates the previous reasoning, that criminal intent is less confounded by variation in perceived crime costliness than are perceptions of the moral seriousness of the crime.

Hypothesis Tests

Using ANOVA, we then examined whether intent predictably influenced punishment independent of the two consequentialist manipulations.[5] For the punishment severity ratings, the predicted effect of intent was observed. No independent effects were observed either for dangerousness or publicity. One small two-way interaction was observed between intent and dangerousness; pair-wise comparisons using Fisher's LSD tests revealed that offender dangerousness positively predicted punishment only when the offense was unintentional. In other words, when, and only when, the offender's crime could be attributed to a brain tumor, his punishment was determined by his stipulated dangerousness. The three-way interaction did not reach significance. However, pair-wise comparisons suggested a nonsignificant trend in which intent retained its predicted effect when dangerousness and publicity were manipulated to be low. Fn5

Similar results were obtained for the prison sentence recommendations. A predicted main effect of intent was observed whereby higher criminal intent resulted in a several-year average increase in sentences. A small, predicted main effect of dangerousness was observed, and a small two-way interaction was observed between intent and dangerousness, showing again that offender dangerousness positively predicted punishment only when the offense was unintentional. No other main effects or interactions were observed (see Table 1). T1

Next, we reasoned that if the effect of intent on punishment is truly the result of the retributive motive, then this effect should be mediated by the desire to make the offender suffer. Moreover, the desire for suffering should not be influenced by the two consequentialist motives (dangerousness or publicity of punishment). The predicted mediation was found using a series of steps modeled after Baron and Kenny's test of mediation (1986). In this approach, mediation is demonstrated by showing that (1) the initial factor predicts the outcome variable, (2) the initial factor predicts variance in the hypothesized mediator, and (3) the mediator predicts the outcome variable, controlling for the influence of the initial factor. In accordance with Step 1, we showed above using a three-way ANOVA that our initial factor, intent, had a unique significant effect on both subjective punishment and sentence length. Second, intent independently predicted the hypothesized

[5] The original punishment measures were skewed to an excessive degree (> twice the standard error of skewness) but were successfully normalized prior to hypothesis testing, using a $\log_{10}$ and $\log_{100}$ transformation, respectively.

Table 1
Untransformed Mean Ratings of Subjective and Objective Punishment Recommendations (Range: 1–11) as a Function of Criminal Intent, Dangerousness, and Publicity

			Subjective punishment		Sentence length		
Intent	Dangerousness	Publicity	*M*	*SD*	*M*	*SD*	*N*
Low	Low	Private	5.70	2.40	6.42	2.03	33
		Public	4.94	2.05	5.79	2.48	33
		Total	5.32[b]	2.25	6.11[e]	2.27	66
	High	Private	5.97	2.14	7.03	3.46	31
		Public	6.13	2.67	6.94	3.45	31
		Total	6.05[b]	2.40	6.98[e]	3.43	62
	Total	Private	5.83	2.27	6.72	2.81	64
		Public	5.52	2.42	6.34	3.03	64
		Total	5.67[a]	2.34	6.53[c]	2.91	128
High	Low	Private	9.03	1.47	9.22	1.54	32
		Public	9.22	1.10	9.56	1.11	32
		Total	9.13	1.29	9.39	1.34	64
	High	Private	9.00	1.63	9.32	1.80	31
		Public	8.71	1.16	9.29	1.32	31
		Total	8.85	1.41	9.31	1.56	62
	Total	Private	9.02	1.54	9.27	1.66	63
		Public	8.97	1.15	9.43	1.12	63
		Total	8.99[a]	1.35	9.35[c]	1.45	126
Total	Low	Private	7.34	2.60	7.80	2.28	65
		Public	7.05	2.71	7.65	2.70	65
		Total	7.19	2.65	7.72[d]	2.49	130
	High	Private	7.48	2.43	8.18	2.97	62
		Public	7.42	2.42	8.11	2.85	62
		Total	7.45	2.41	8.15	2.90	124
	Total	Private	7.41	2.51	7.98[d]	2.63	127
		Public	7.23	2.57	7.87	2.78	127
		Total	7.32	2.53	7.93	2.70	254

Note. Significant comparisons are denoted by matching letter pairs: (A) main effect of intent, $F(1, 246) = 183.64$, $p < .001$, $\eta_p^2 = .43$; (B) Intent × Dangerousness interaction, $F(1, 246) = 3.96$, $p < .05$, $\eta_p^2 = .02$; (C) main effect of intent, $F(1, 246) = 93.73$, $p = .001$, $\eta_p^2 = .28$; (D) main effect of dangerousness, $F(1, 246) = 4.42$, $p < .05$, $\eta_p^2 = .02$; (E) Intent × Dangerousness interaction, $F(1, 246) = 5.37$, $p < .05$, $\eta_p^2 = .02$. Nonsignificant effects for subjective punishment include dangerousness: $F(1, 246) = .23$, $p = .63$, $1 - \beta = .08$; publicity: $F(1, 246) = .90$, $p = .34$, $1 - \beta = .16$; Intent × Dangerousness × Publicity interaction: $F(2, 246) = 3.32$, $p = .07$, $1 - \beta = .44$. Nonsignificant effects for objective punishment include publicity: $F(2, 246) = .21$, $p = .65$, $1 - \beta = .08$; Intent × Dangerousness × Publicity interaction: $F(2, 246) = .74$, $p = .39$, $1 - \beta = .14$.

mediator of suffering, $F(1, 246) = 136.32$, $p < .001$, $\eta_p^2 = .36$. In this step, suffering was not predicted by dangerousness, $F(1, 246) = .72$, $p = .40$, $1 - \beta = .14$, or by publicity, $F(1, 246) = .94$, $p = .33$, $1 - \beta = .16$. Third, independent linear regressions on the two punishment measures showed that the desire for suffering ratings positively predicted punishment ratings, controlling for intent, $R^2 = .65$ and $R^2 = .67$, respectively. Specifically, the desire for suffering uniquely

explained 48% of the variance in subjective punishment, $\beta = .60$, $p < .001$, and 42% of the variance in sentence, $\beta = .52$, $p < .001$. Taken together, the main effect of intent on punishment can be explained largely by the desire for the offender to suffer.

Additional Analyses

To test the prediction that more politically liberal participants will be more lenient overall in their punishment decisions, the correlation between political conservatism and punishment was examined. Indeed, more liberal participants recommended punishments that were less severe than those more conservative, $r = .13$, $p < .05$. A similar trend was observed with respect to sentence length; more liberal participants tended to recommend shorter sentences, $r = .10$, $p = .097$.

Study 1 Conclusion and Limitations

Study 1 confirmed the hypothesis that punitive sentiments would emerge in reaction to a retributive cue (intent) even when the common consequentialist cues of incapacitation and deterrence were minimized, suggesting evidence of retribution. This pattern of results is consistent with findings by Baron and Ritov (1994). In that study, participants fined a negligent company more than what would be needed for full compensation even when circumstances precluded the need for specific or general deterrence. The present study extends this pattern of results to the criminal domain.

With regard to criminal punishment, similar conclusions have been drawn by many other researchers. As noted, Darley, Carlsmith and others have attempted to elicit retribution independently of consequentialist motives, and they observed greater punishment in response to what they believed to be purely retributive cues.

The present results provide purer evidence of retribution as a motive for punishment than previous studies because those studies did not orthogonally manipulate indices of retribution, incapacitation, and deterrence simultaneously, or did not do so sufficiently. Consequently, variation associated with one motive could be explained by its overlap with another. The present study attempted to lessen the influence of these common confounds, and still found evidence strongly suggestive of the retributive motive. Thus, although the Study 1 conclusion is not novel on its face, it strengthens the validity of that conclusion by ruling out critical but previously neglected confounds.

In Study 1, we also observed a small but significant effect of dangerousness on sentence length recommendations. Similar effects have been observed in other studies (e.g., Aharoni, 2009; Darley et al., 2000; for an exception see Sunstein, Schkade, & Kahneman, 2000). However, in the present case, this main effect was trumped by an interaction whereby dangerousness was only relevant when intent was low. Surprisingly, we found no independent effect of publicity on punishment recommendations—a result that is inconsistent with previous literature (e.g., Kurzban et al., 2007; Piazza, 2008). One possible explanation for this null effect is that our between-subjects design may have lacked the sensitivity required to observe true effects of publicity on punishment. However, further research would be needed to explain this inconsistency.

As with any scientific experiment, Study 1 is not without its limitations. As we noted, the fact that participants rated the "high-intent" offender as somewhat more dangerous than the "low-intent" offender could leave open the possibility that punishments inspired by the intent manipulation are really tracking dangerousness instead. However, it is difficult to defend this interpretation in light of the two-way interaction and the small effect size associated with the dangerousness main effect. Nonetheless, future efforts to fully dissociate intent from dangerousness will be needed to strengthen our confidence in the unique role of intent in such punishment recommendations.

Some might take issue with the assumption that the brain tumor manipulation is a pure way of operationally defining low intent. As noted, this choice reflected a compromise in crossing intent with dangerousness and a desire to replicate past research. To its credit, belief in the manipulation was supported by the results of the manipulation checks. However, an alternative method could be to manipulate low intent in the context of a crime that was committed accidentally. Although this alternative can make it challenging to cross intent with future dangerousness factorially, it has nonetheless yielded similar effects (see Aharoni, 2009), suggesting that the predicted effect of intent on punishment is robust.

The study was conducted in the sterility of a psychology laboratory and should be extended to more realistic settings. Hypothetical scenarios like those we used have been validated in previous research (Bornstein, 1999; Carlsmith et al., 2002; Rose & Ogloff, 2001; Wilson & O'Gorman, 2002), but richer, more consequential and personally relevant settings, such as mock trials, may produce more naturalistic punishment behavior. In true courts of law, jury decisions are deliberated upon in groups. The present study sought to understand individual punishment psychology before examining how this may be moderated by group factors. Future studies may include true audiences, especially given the demonstrated power of group dynamics on jurors' decision-making (see Janis, 1972). The presence of a true audience is likely to increase the punitive motive of general deterrence.

Finally, our sample of undergraduate participants is unrepresentative of the diversity found in typical courtroom populations. As shown, our university students tended to identify most with democratic, liberal, political attitudes. It was also found that participants who were more politically liberal were slightly less punitive than their more conservative counterparts. Therefore, it is plausible that typical effects of retribution in a population with a more representative proportion of conservative attitudes could be even stronger than we presently observed. Certainly, a broader spectrum of participant ages, backgrounds, and legal expertise would potentially extend the generality of our findings.

Study 2: Are We Rational Punishers? Punishment Without Reasons

Method

Participants

Participants consisted of 49 psychology students from the UCSB, who received introductory psychology course credit for their participation in a study

entitled "Social Opinions."[6] We excluded responses from two subjects who reported consequentialist reasons that were unanticipated because the goal of the interview was not to measure all possible practical motives but to simply rule out those that are common.[7] What remained was an *N* of 47 (36 females, 11 males). All participants were age-qualified for U.S. jury selection where ages ranged from 18–26. Other demographic information collected included ethnicity (Caucasian: 32.7%; Hispanic: 20.4%; Asian: 32.7%; Other: 2%), religious affiliation (Christianity: 18.4%; Catholicism: 16.3%; Buddhism: 4.1%; Islam: 0%; Other: 4.1%; None: 18.4%), church attendance ($M = 10.79$ times per year; $SD = 19.53$), political party affiliation (Democratic: 57.1%; Republican: 2.0%; Independent: 4.1%; Libertarian: 2.0%; Other: 8.2%), and political spectrum affiliation ($M = 4.14$; $SD = 1.97$) where 1 = "very liberal" and 11 = "very conservative." Two participants reported having served jury duty, and two reported having a family member who had served as a judge or lawyer.[8] Fn6 Fn7 Fn8

Design

The design of this study was a semistructured oral interview with a pretest/posttest measurement. The participants' task was to decide, according to their own standards, the nature and magnitude of punishment that the offender should have received. The vignette was designed to rouse strong moralistic responses. However, as suggested, it was also designed such that the most common practical (i.e., consequentialist) reasons for desiring punishment were moot. In this context, participants were asked to either (1) ascribe punishment or not; and among the former, to either (2) retract or sustain that punishment. By some conceptualizations, any tendency to sustain punishment at rates statistically greater than zero percent of the total sample would be theoretically interesting. However, we upheld a stricter criterion whereby chance performance was based on the assumption of an equal likelihood of obtaining one of the four possible outcomes above. Thus, the specific hypothesis to be evaluated is that more than 25% of the total sample will sustain punishment without explicit reasons.

Materials

The interview instructions asked participants to imagine they are the judge in the penalty phase of a criminal trial. The trial, depicted in a three-paragraph vignette, described an offender's physical and sexual assault and homicide of a minor. Guilt was positively determined by DNA evidence. General deterrence concerns were minimized by stipulating that no information about the sentence would be released to the public in order to protect the victim's identity. Concerns

[6] All experimental procedures were approved by the UCSB Committee on Human Subjects Research.

[7] The two consequentialist reasons for punishment that were unanticipated were (1) that the offender's release could set a legal precedent whereby other more culpable offenders might derive a legal excuse, and (2) that the offender's release could lead to future crime by passing on "criminal genes" to future offspring. The former reason inspired a rebuttal in a subsequent version of the interview, shown below. The latter reason was noted, but never reported by any other participant.

[8] Responses to demographic questions represent approximately 70% of the total sample. These items were not collected in the remaining ~30% of participants.

about specific deterrence and incapacitation of the offender were minimized by noting that before the penalty phase of his trial, the offender developed symptoms of an autoimmune disease that permanently paralyzed his limbs, rendering him not dangerous to society. Finally, to dispel the desire to punish for the sake of the victim's family, the vignette stated that the family did not desire punishment. (See Appendix B for full interview script.)

Procedure

Participants were asked to read a vignette about a man who committed a malicious criminal offense. Next, they were asked in a face-to-face interview the nature and magnitude of punishment that the offender should receive (pretest). The participants who recommended punishment were asked to justify their reasons for punishment. When they expressed justifications that were irrelevant to the story, the interviewer reminded them why each justification was inapplicable to the scenario. For instance, if they justified their desire for punishment for the purpose of preventing the offender from recidivating, they were reminded that the court accepted evidence that he will no longer be a danger to society on account of his paralysis. They then were probed for their acceptance of this argument. Similarly, if they sought punishment for the purpose of deterring others, they were reminded that the punishment was to be withheld from public record. If they upheld the punishment as a means to preserve consistency of the law, they were reminded that, because the penalty would be confidential, no one else would know whether the ruling had been consistent in this case. If they presented a retributive justification for punishment (e.g., "justice" or "an eye for an eye"), or favored consistency of a moral rule for its own sake, they were challenged to explain what purpose this would serve. This process continued for several minutes until participants either (a) expressed a reason for punishment that was unanticipated or (b) ran out of reasons for the punishment. When the latter occurred, participants once again were asked (posttest) whether they still recommended the punishment and, if so, whether they had any other reasons for doing so.

Responses were coded by the interviewer. To meet criteria for a particular code, note that participants were not expected to demonstrate their philosophical knowledge of punishment theory, but instead, to simply verbalize any lay practical reason for punishing. The interviewer then judged those responses as fulfilling a particular punishment motive.

Results

Provided the option to either (1) ascribe punishment or not, and among the former, to (2) retract or sustain that punishment, mere chance would predict that 25% of the total sample would sustain punishment. In fact, three of the 47 qualified participants, or 6.4%, desired no punishment at all. Ten more (21.3%) were willing to retract punishment after their arguments were contradicted during the interview. The remaining 34 participants (72.3%) were unwilling to retract punishment, and, when prompted, did not generate any consequentialist reasons for their resolve. A Chi-Square test was employed to examine whether these latter two proportions statistically differed from the chance-level criterion. Indeed they did, $\chi^2(1) = 13.09$, $p < .001$.

Notably, in response to the question "How bad was the crime on a 1–10 scale?" the original three participants who did not wish to punish the offender nonetheless rated the offense just as bad ($M = 9.0$, $SD = 1.0$) as did the other participants ($M = 8.93$, $SD = 1.05$). Indeed, no single interviewee rated the moral seriousness of the crime less than 7 points out of 10.

The uneven gender distribution of the participants, though similar to the gender base rates in our college population, left our effects vulnerable to a gender confound. However, we found no evidence that gender affected the three possible punishment recommendations (Mann–Whitney $U(45) = 175.00$, $Z = -.74$, $p = .577$).

Study 2 Conclusion and Limitations

As predicted, a substantial proportion of participants recommended punishment even after being disarmed of their explicit reasons for doing so. As noted, any persistence of punishment statistically greater than zero would have been interesting. However, the observed proportion far exceeded even the more stringent chance-level estimate of 25% and prevailed as the supermajority response. Moreover, given that our participants, like most university students (see Intercollegiate Studies Institute report, 2009), describe themselves as more liberal than conservative, we might expect that these university students are disproportionately "soft on crime" compared to the general U.S. population, and as such would be less inclined, if anything, toward moralistic punishment. Thus, a random sample might predict even greater persistence in such punitive attitudes.

Similar results have been obtained by other researchers. As mentioned above, Carlsmith and colleagues showed that participants' punitive sentiments were not satisfied when they learned that the offender was unlikely to recidivate due to a disabling medical condition (Carlsmith, Darley, & Robinson, 2002). The present study expanded on this finding by showing that punitive sentiments will still persist when other practical concerns beyond specific deterrence are addressed.

Similarly, Greene and colleagues (Greene, Nystrom, Engell, Darley, & Cohen, 2004; Greene et al., 2001) found that their study participants were far less likely to support killing an innocent bystander to save five others when doing so would have required personal contact rather than use of a mechanical switch. In addition, participants were unable to provide a rationale for this inconsistency, suggesting their responses were shaped more by heuristic than rational processes. The present study complements these findings by suggesting that within the context of criminal punishment, some lay punishment decisions may also be formulated by such heuristics more than by practical reason.

Finally, the observed pattern of results closely parallels research by Haidt et al. (2000/2001), which found that the majority of participants had great difficulty defending their persistent objections to ostensibly immoral acts with little negative consequence. Again, the present study extends this work by suggesting a strong role for heuristics judgment, not just as an internal objection to immoral acts, but also as an obligation to punish putatively immoral actors by way of legal sanction.

Before we accept this conclusion, several other factors must be considered. First, this exploratory study was engineered to discover whether unexamined

heuristic processes played any role in moral decision making, not to exclude any role for practical reasoning. To the contrary, we would argue that practical reason can and does guide punitive decisions. Both may operate hand-in-hand in the course of moral development. For example, the rationalizations that occur following a moral decision may constitute the reasons that underlie the next moral decision. As the responses of 21% of our participants suggest, punitive heuristics can be managed.

Second, the interview was not double-blind. Because the interviewer was privy to the study's hypothesis, it is possible that, through unintended experimental demands, the interviewer could have biased participants to sustain their punishment recommendations when they might have otherwise retracted. We upheld two checks against this possibility. First, any ambiguous or ambivalent punishment judgments were coded as counter to the study's hypothesis (i.e., instrumental). Second, the interview was specifically scripted, via an onslaught of suggestions, to gently, but repeatedly, encourage punitive participants to retract their punishments on practical grounds—again, counter to the study's hypothesis. In this manner, the nature of the questioning places the interviewer in an amiable but adversarial relationship with the participants, and we surmise that a questioner naïve to the experiment's hypotheses at the outset would probably not remain so for long.

Because the interview was designed to encourage punishment retraction, it is possible that some punitive judgments persisted out of stubbornness, or psychological reactivity against the interviewer's appeals. Indeed, such reactance patterns have been observed in studies of the effects of persuasion on conformity (e.g., Brehm & Himelick Cole, 1966; Brehm & Mann, 1975). However, such patterns are usually associated with attitudes of annoyance or defiant anger (Dillard & Shen, 2005). While we cannot exclude the operation of reactance, we witnessed no testiness or irritability among any of the subjects engaged in our interview.

It is also possible that the participants who were classified as persistent punishers in fact had specific reasons for their apparent persistence that they did not communicate. One such possibility might be that, despite having valid practical reasons for punishment, they were unable to articulate these reasons. However, in our view, it is the very fact that such individuals were unable to articulate practical reasons that defines them as nonrational noninstrumental punishers. Another possibility is that participants had valid practical reasons for punishment that they chose not to articulate due to social desirability pressures. However, if their responses were truly driven by social desirability, they should have been expected to retract their punishments as well, but they did not.

As with all vignette-based studies, it is unclear the extent to which participants' punishment recommendations would be comparable to their real-world analogues. One challenge to the ecological validity of such studies is that decisions are inconsequential, accruing no actual costs or benefits to the punisher. However, there is reason to believe that under real-world circumstances, where reputations might be at stake, even more respondents would punish in a heuristic noninstrumental manner. For instance, in one-shot experimental economics games with real monetary payoffs, researchers have found that reductions in punishers' perceived anonymity increase the degree to which they punish those who have

violated a cooperation norm, even though participants were told they would not play more than one round (Kurzban, Descioli, & O'Brien, 2007; Piazza, 2008). Similarly, other studies have shown that visual depictions of disembodied eyespots increase the amount of money participants are willing to share with "responders" in a dictator game, suggesting that our sense of social obligation may be cued automatically by nonrational stimuli (Haley & Fessler, 2005). Still others have observed increases in participants' donations, as well as their moralistic rationales for these donations, when such donation decisions had payoffs that were real as opposed to merely hypothetical (Aguiar, Brañas–Garza, & Miller, 2008). The authors attribute some of these differences to increases in information relevant to the recipient's welfare. To the extent that information about the offender, victim, and other observers increases moralistic punishment, these cues should be more potent in actual criminal trials than in vignette-based studies.

Another challenge to the ecological validity of the present study is that lay punishment behavior might not generalize to that of trial court judges. Their knowledge and expertise in legal policy and practice might of course strongly influence how much they punish a given offender. However, even judges are not immune to common biases that guide human decision making. For instance, it has been consistently observed that judges tend to grant more lenient sentences to female defendants than to male defendants, controlling for other legally relevant variables (Bickle & Peterson, 1991; Daly & Bordt, 1995; Steffensmeier, Kramer, & Streifel, 1993). Even holding such extralegal influences aside, the practice of formulating a punishment decision is complicated by the fact that most Western jurisdictions offer extremely wide discretion to the punisher regarding adherence to retributive and consequentialist punishment motives (Rappaport, 2003). With no legal guidance on how punishers ought to weigh the relative importance of the various punishment goals, the question of the punisher's motives may be, at heart, a lay issue. That is, jurors in capital trials and judges alike must largely resort to their own devices when making punishment decisions in this nebulous domain. As such, examining lay decision making may provide a meaningful contribution to questions of the psychology of criminal punishment. In support of this contention, justice research has shown that undergraduate samples can provide valid models of typical jury behavior (Rose & Ogloff, 2001).

Last, as in Study 1, attempts to measure noninstrumental punishment should be extended to more ecologically realistic settings and to samples rich in demographic and political diversity. It will be vital to establish that heuristics-driven punishment is not simply an artifact of our laboratory environment or our relatively small homogenous university sample.

These important limitations notwithstanding, the overwhelming persistence of punishment without explicit, practical reasons suggests that at least some of our lay punitive decisions may be driven by nonrational noninstrumental heuristics.

General Discussion

Early research in motives for criminal punishment made great strides in demonstrating the complex nature of these motives, but it did not adequately demonstrate the independent operation of retribution. As suggested, this difficulty arises from the fact that retributive and consequentialist motives for punishment

are often entangled in practice. However, understanding their unique contributions may be necessary to decide when they assist or hinder social goals such as crime prevention. This places a continued burden on experimentalists to demonstrate the unique contributions of each motive to the various contexts in which punishment decisions are made. The present studies attempted to overcome some of the prior limitations by mitigating the most common consequentialist reasons for punishment, and evaluating whether punishment in this context remains responsive to purely retributive cues.

The present studies showed that participants persisted in punishment of a violent offender even when the practical consequences of punishment were removed. This pattern was evoked largely on the basis of criminal intent (Study 1) and was hard to verbally justify (Study 2). These findings lend needed support to the theory that such punishment is driven primarily by retribution; however, this stance we call retribution may be better explained by heuristic processes rather than by abstract moral principles. These conclusions converge with and extend prior research, as we have outlined in the Study 1 and II conclusions sections above.

If human punishment is driven largely by retribution, and retribution is arrived at by way of heuristic judgments, we are forced to ask whether we can trust such judgments. Sometimes heuristic decision making serves us well. Indeed, researchers have suggested that retributive motives may in fact have reliably deterred aggressors over our species' phylogenetic development (Aharoni, 2009; Daly & Wilson, 1988; McCullough, 2008; Price, Cosmides, & Tooby, 2002; Trivers, 1971), which may explain why these theoretical constructs can be so difficult to tease apart. At other times, however, even the most common heuristic inclinations can be plainly wrong, as research in psychology has long warned (Denes–Raj & Epstein, 1994; Gilovich & Douglas, 1986). For example, research on infanticide behavior (Daly & Wilson, 1988) shows how impulses that might have benefitted our ancestors' individual fitness may not be useful to societal groups. Likewise, it remains a topic of continued debate as to whether and when retribution serves prosocial goals (e.g., Greene & Cohen, 2004; Moore, 1997; Robinson & Darley, 1997). AQ: 1

When research suggests that incarceration is not just ineffective at deterring crime but even criminogenic, the persistent desire for retribution without practical reasons becomes difficult to justify. As such, policies that enable retribution, such as those found in the Model Penal Code, may warrant deeper scrutiny. Ultimately, a justice policy that is critical of retributive punishment could motivate the development of more effective deterrence practices within courtrooms, correctional facilities, and therapeutic settings. Such a change will depend on a thorough scientific understanding of the psychology of punitive decision making.

References

Aguiar, F., Brañas–Garza, P., & Miller, L. M. (2008). Moral distance in dictator games. *Judgment and Decision Making, 3,* 344–354. http://journal.sjdm.org/

Aharoni, E. (2009). *Why do we punish? Studies of lay judgments against criminal offenders*. Doctoral dissertation, Department of Psychology, UCSB. Santa Barbara, CA.

Baron, J., & Ritov, I. (1994). Reference points and omission bias. *Organizational Behavior and Human Decision Processes, 59,* 475–498. doi:10.1006/obhd.1994.1070

Baron, R., & Richardson, D. (1994). *Human aggression*. New York, NY: Plenum Press.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182. doi:10.1037/0022-3514.51.6.1173

Beccaria, C. (1986). *On crimes and punishments,* (D. Young, Trans.) Indianapolis, IN: Hackett. (Original work published in 1764)

Bentham, J. (1789/1988). *An introduction to the principles of morals and legislation,* New York, NY: Prometheus Books.

Bickle, G., & Peterson, R. (1991). The impact of gender-based family roles on criminal sentencing. *Social Problems, 38,* 372–394. doi:10.1525/sp.1991.38.3.03a00050

Blum West, S. R. (1985). The seriousness of crime: A study of popular morality. *Deviant Behavior, 6,* 83–98. doi:10.1080/01639625.1985.9967661

Bohm, R. (1992). Retribution and capital punishment: Toward a better understanding of death penalty opinion. *Journal of Criminal Justice, 20,* 227–236. doi:10.1016/0047-2352(92)90047-D

Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior, 23,* 75–91. doi:10.1023/A:1022326807441

Brehm, J. W., & Himelick Cole, A. (1966). Effect of a favor which reduces freedom. *Journal of Personality and Social Psychology, 3,* 420–426. doi:10.1037/h0023034

Brehm, J. W., & Mann, M. (1975). Effect importance of freedom and attraction to group members on influence produced by group pressure. *Journal of Personality and Social Psychology, 31,* 816–824. doi:10.1037/h0076691

Carlsmith, K. M. (2006). The roles of retribution and utility in determining punishment. *Journal of Experimental Social Psychology, 42,* 437–451. doi:10.1016/j.jesp.2005.06.007

Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish? Deterrence and Just Deserts as motives for punishment. *Journal of Personality and Social Psychology, 83,* 284–299. doi:10.1037/0022-3514.83.2.284

Chen, M. K., & Shapiro, J. M. (2007). Do harsher prison conditions reduce recidivism? A discontinuity-based approach. *American Law and Economics Review, 9,* 1–29. doi:10.1093/aler/ahm006

Cid, J. (2009). Is imprisonment criminogenic? A comparative study of recidivism rates between prison and suspended prison sanctions. *European Journal of Criminology, 6,* 459–480. doi:10.1177/1477370809341128

Costanzo, P. R., Coie, J. D., Grumet, J. F., & Farnill, D. A. (1973). Reexamination of the effects of intent and consequences of children's moral judgments. *Child Development, 44,* 154–161. doi:10.2307/1127693

Cottingham, J. (1979). Varieties of retribution. *The Philosophical Quarterly, 29,* 238–246. doi:10.2307/2218820, doi: 10.2307/2218820

Daly, K., & Bordt, R. (1995). Sex effects and sentencing: An analysis of the statistical literature. *Justice Quarterly, 12,* 43–177. doi:10.1080/07418829500092601

Daly, M., & Wilson, M. (1988). *Homicide.* New York, NY: Aldine de Gruyter.

Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2000). Incapacitation and just deserts as motives for punishment. *Law and Human Behavior, 24,* 659–683. doi:10.1023/A:1005552203727

Darley, J. M., Klosson, E. C., & Zanna, M. P. (1978). Intentions and their contexts in the moral judgments of children and adults. *Child Development, 49,* 66–74. doi:10.2307/1128594

Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology, 66,* 819–829. doi:10.1037/0022–3514.66.5.819

Dillard, J., & Shen, L. (2005). On the nature of reactance and its role in persuasive health communication. *Communication Monographs, 72,* 144–168. doi:10.1080/03637750500111815

Garner, B. A. (Ed.). (2004). *Black's law dictionary, 8.* St. Paul, MN: West.

Gazzaniga, M. S. (1998). *The mind's past.* Berkeley, CA: University of California Press.

Gazzaniga, M. S. (2000). Cerebral specialization and interhemispheric communication: Does the corpus callosum enable the human condition? *Brain, 123,* 1293–1326. doi:10.1093/brain/123.7.1293

Geen, R. (1968). Effects of frustration, attack and prior training in aggressiveness upon aggressive behavior. *Journal of Personality and Social Psychology, 9,* 316–321. doi:10.1037/h0026054

Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart.* New York, NY: Oxford University Press.

Gilovich, T., & Douglas, C. (1986). Biased evaluations of randomly determined gambling outcomes. *Journal of Experimental Social Psychology, 22,* 228–241. doi:10.1016/0022–1031(86)90026–0

Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society of London B, 359,* 1775–1785.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron, 44,* 389–400. doi:10.1016/j.neuron. 2004.09.027

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral Judgment. *Science, 293,* 2105–2108. doi:10.1126/science.1062872

Greenwell, J., & Dengerink. H. (1973). The role of perceived versus actual attack in human physical aggression. *Journal of Personality and Social Psychology, 26,* 66–71. doi:10.1037/h0034223

Grogger, J. (1991). Certainty vs. severity of punishment. *Economic Inquiry, 29,* 297–309. doi:10.1111/j. 1465–7295.1991.tb01272.x

Haggard, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Sciences, 9,* 290–295. doi:10.1016/j.tics. 2005.04.012

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108,* 814–834. doi:10.1037/0033-295X.108.4.814

Haidt, J., Bjorklund, F., & Murphy, S. (2000). *Moral dumbfounding: When intuition finds no reason*. (Unpublished manuscript, University of Virginia)

Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior, 26,* 245–256. doi:10.1016/j.evolhumbehav.2005.01.002

Hamilton, V., & Rytina, S. (1980). Social consensus on norms of justice: Should the punishment fit the crime? *American Journal of Sociology, 85,* 1117–1144. doi:10.1086/227127

Harlow, R. E., Darley, J. M., & Robinson, P. H. (1995). The severity of intermediate penal sanctions: A psychophysical scaling approach for obtaining community perceptions. *Journal of Quantitative Criminology, 11,* 71–95. doi:10.1007/BF02221301

Intercollegiate Studies Institute. (2009). *The shaping of the American mind: The diverging influences of the college degree & civic learning on American beliefs.* Wilmington, DE: American Civil Literacy Program.

Janis, I. L. (1972). *Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes.* Oxford, England, U.K.: Houghton Mifflin.

Kant, I. (1952). *Kant's critique of judgment,* (J. C. Meredith, Trans.) Oxford, U.K.: Clarendon Press. (Original work published in 1790)

Kant, I. (1998). *Kant's groundwork of the metaphysics of morals: Critical essays.* P. Guyer (Ed.). Lantham, MD: Rowman and Littlefield. (Original work published in 1785)

Karniol, R. (1978). Children's use of intention cues in evaluating behavior. *Psychology Bulletin, 85,* 76–85. doi:10.1037/0033–2909.85.1.76

Keller, L. B., Oswald, M. E., Stucki, I., & Gollwitzer, M. (2010). A closer look at an eye for an eye: Laypersons' punishment decisions are primarily driven by retributive motives. *Social Justice Research, 23,* 99–116. doi:10.1007/s11211-010–0113-4

Kurzban, R., DeScioli, P., & O'Brien, E. (2007). Audience effects on moralistic punishment. *Evolution and Human Behavior, 28,* 75–84. doi:10.1016/j.evolhumbehav.2006.06.001

Lepper, C. A., Anders, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology, 39,* 1037–1049. doi:10.1037/h0077720

McCullough, M. E. (2008). *Beyond revenge: The evolution of the forgiveness instinct.* San Francisco, CA: Jossey-Bass.

McFatter, R. (1982). Purposes of punishment: Effects of utilities of criminal sanctions on perceived appropriateness. *Journal of Applied Psychology, 67,* 255–267. doi:10.1037/0021–9010.67.3.255

Miller, D. T., & Vidmar, N. (1981). The social psychology of punishment reactions. In M. Lerner and S. Lerner (Eds.), *The justice motive in social behavior.* Model Penal Code § 1.02(2) (1962).

Monahan, J., & Hood, G. L. (1976). Psychologically disordered and criminal offenders: Perceptions of their volition and responsibility. *Criminal Justice and Behavior, 3,* 123–133. doi:10.1177/009385487600300202

Moore, M. (1997). *Placing blame.* New York, NY: Clarendon Press.

Packer, H. L. (1968). *The limits of the criminal sanction* (pp. 19–23). Stanford, CA: Stanford University Press.

Piaget, J. (1948). *The moral judgment of the child.* New York, NY: Free Press.

Piazza, J. (2008). The effects of perceived anonymity on altruistic punishment. *Evolutionary Psychology, 6,* 487–501.

Price, M., Cosmides, L., & Tooby, J. (2002). Punitive sentiment as an anti-free rider psychological device. *Evolution and Human Behavior, 23,* 203–231. doi:10.1016/s1090-5138(01)00093–9

Rappaport, A. (2003). Rationalizing the commission: The philosophical premises of the U.S. sentencing guidelines. *Emory Law Journal, 52,* 557–652. http://www.law.emory.edu/index.php?id=4158

Robinson, P. H., & Darley, J. M. (1995). *Justice, liability, and blame* (pp. 14–28). Boulder, CO: Westview Press, Inc.

Robinson, P. H., & Darley, J. M. (1997). The utility of desert. *Northwestern University Law Review, 91,* 453–499. http://www.law.northwestern.edu/lawreview/

Rose, G., & Ogloff, J. (2001). Evaluating the comprehensibility of jury instructions: A method and an example. *Law and Human Behavior, 25,* 409–431. doi:10.1023/A:1010659703309

Rucker, D., Polifroni, M., Tetlock, P., & Scott, A. (2004). On the assignment of punishment: The impact of general-societal threat and the moderating role of severity. *Personality and Social Psychology Bulletin, 30,* 673–684. doi:10.1177/0146167203262849

Steffensmeier, D., Kramer, J. H., & Streifel, C. (1993). Gender and imprisonment decisions. *Criminology, 31,* 411–446. doi:10.1111/j.1745–9125.1993.tb01136.x

Sunstein, C. R., Schkade, D., & Kahneman, D. (2000). Do people want optimal deterrence? *The Journal of Legal Studies, 29,* 237–253. doi:10.1086/468069

Surber, C. F. (1977). Developmental processes in social inference: Averaging of inten-

tions and consequences in moral judgment. *Developmental Psychology, 13,* 654–665. doi:10.1037/0012–1649.13.6.654

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

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Tyler, T. R., & Boeckmann, R. J. (1997). Three strikes and you are out, but why? The psychology of public support for punishing rule breakers. *Law and Society Review, 31,* 237–265. doi:10.2307/3053926

Vidmar, N. (2000). Retribution and revenge. In J. Sanders J. & V. L. Hamilton (Eds.), *Handbook of justice research in law*. New York, NY: Kluwer/Plenum Press.

Viney, W., Parker–Martin, P., & Dotten, S. D. H. (1988). Beliefs in free will and determinism and lack of relation to punishment rationale and magnitude, *The Journal of General Psychology, 115,* 15–23. doi:10.1080/00221309.1988.9711084

Warr, M., Meier, R., & Erickson, M. (1983). Norms, theories of punishment, and publicly preferred penalties for crimes. *The Sociological Quarterly, 24,* 75–91. doi:10.1111/j.1533–8525.1983.tb02229.x

White, G. F. (1975). Public responses to hypothetical crimes: Effect of offender and victim status and seriousness of the offense on punitive reactions. *Social Forces, 53,* 411–419. doi:10.2307/2576583

Wilson, D., & O'Gorman, R. (2003). Emotions and actions associated with norm-breaking events. *Human Nature, 14,* 277–304. doi:10.1007/s12110-003–1007-z

Appendix A

Publicity Manipulation:

[PRIVATE] To protect your privacy, all of your answers will be confidential and will not leave this experiment unless you give us permission. The results will be coded, so even the experimenters won't be able to see your particular answers.

[OR]

[PUBLIC] After your session, the experimenters will review your answers and evaluate your performance as a judge.

Case Stimuli:

Greg is 26 years old. He is a customer service representative for a small cell phone retailer. Recently, while walking in his neighborhood, he ranted at 17-year-old male passing by, and hit the victim with a crowbar that he grabbed from a nearby construction site. The victim died of massive head injuries shortly afterward.

A police investigation after Greg's arrest revealed that the victim had been a coworker of Greg's at the cell phone company. Apparently, they had gotten along fine at the start, except the co-worker was proving to be a much better employee

(*Appendices continue*)

than Greg, and so Greg blamed him for “”showing him up,“” which blocked a promotion that Greg had been hoping for. After missing the promotion, Greg, who had previously been mild mannered on the job, had become quite angry. On what became his last day of employment, he was discovered to be trying to sabotage the co-worker’s accounts and files, and when he was caught, he yelled and lunged at the security guard. He was fired on the spot and led out the door.

Intent Manipulation:

[HIGH] Everyone at the cell phone company was shocked that Greg would ever do anything like this. As is standard procedure for this type of offense, Greg then had to submit to an array of psychological and medical tests, which included an MRI scan, blood and endocrine tests, assessments of his life history, temperament, and attitudes, and general cardiovascular, and endocrine functioning.

[OR]

[LOW] Everyone at the cell phone company was shocked that Greg would ever do anything like this. As is standard procedure for this type of offense, Greg then had to submit to an array of psychological and medical tests, which included an MRI scan, blood and endocrine tests, assessments of his life history, temperament, and attitudes, and general cardiovascular, and endocrine functioning. These tests found that Greg had a brain tumor in an area associated with violent behavior. The medical and psychological experts assigned to this case all agree that the tumor, by pressing on this critical brain area, was the cause both of Greg’s decline at work and his subsequent sabotage and violence. Greg had a delusion that his co-worker was an evil demon who was attacking him, and so he believed he had to act in self defense.

Dangerousness Manipulation:

[HIGH] The medical and psychological experts assigned to this case all agree on the basis of his evaluation profile that Greg is *extremely dangerous*. [OR] The tumor is inoperable, it cannot be removed. Medication proves ineffective at reducing the violent disposition the tumor creates. On this basis, the experts determined that Greg is *extremely dangerous.*

[OR]

[LOW] The medical and psychological experts assigned to this case all agree on the basis of his evaluation profile that Greg is *not at all dangerous*. [OR] The tumor proves to be easily operable, and medication is successful at restoring Greg to a healthy state. On this basis, the experts determined that, after his surgery and follow-up treatment, Greg is *not at all dangerous*.

Publicity Manipulation:

[PRIVATE] Because the victim was a legal minor at the time the case was filed, state law upheld the Victim Protection Act, which requires that all cases

(*Appendices continue*)

involving minors under the age of 18 are closed to the public, as are any hearings, transcripts, or sentences resulting from the trial.

[OR]

[PUBLIC] The courtroom contained a full audience, including local media coverage; and state law requires that the trial hearing, transcript, and sentence decision be made publicly available.

Appendix B

Interviewer. Now I'd like to ask you several questions about your reactions to the story. I have to ask you some questions designed to generate a discussion. If it ever sounds like I'm trying to argue with you, that's not my intention. I'm just required to ask. As always, you don't have to answer any questions that you don't want to. Okay?

How did you feel about the offender's crime? (free response)

How bad was it? (scale: 1-10)

How many years in prison would you give him? (free response)

How would you feel if the judge decided to let him off scot free?

___ If okay: So you don't think he deserves any kind of punishment for what he did?

- If no: Why not? _____
- Okay. Thanks for your participation.

___ If community service: What if you learned that he already does community service for fun. Would you then let him go, or would you require him to commit additional hours of service?

Is it important to you that the community service causes him to reflect on what he did and feel bad about it (just like punishment)? ___ If not okay: Why?

___ **Retribution.** He brutally attacked somebody, so he deserves some punishment.

Interviewer. Why does he deserve punishment? Why does doing something wrong warrant the punishment you gave him? What's wrong with just letting him go?

___ **Specific deterrence.** He might hurt someone again.

Interviewer. Well, remember, there was extensive evidence that he is not dangerous. If that's true, what's wrong with letting him go?

___ **Nonphysical dangerousness.** He might still be able to instigate crime by nonphysical means.

Interviewer. Okay, but let's suppose that he's truly not dangerous in any way. If not, would you still want to punish him?

___ **General deterrence.** Punishing him will deter others.

(*Appendices continue*)

Interviewer. Remember, in this case, his punishment will not be publicized anyway because of a state law protecting minors. So, if no one will know about his punishment, what's wrong with letting him go?

___ **Contract violation.** He voluntarily broke the rules, knowing that there's a penalty for that. Anyone who breaks a contract is basically volunteering to receive the consequences for that.

Interviewer. Well, just because he broke a contract doesn't necessarily mean he deserves punishment. People break rules all the time (like speeding), but we don't usually give them a criminal sentence. So, why would we punish them here?

___ **Consistency.** It's important that the law is consistent.

Interviewer. Okay, I can see the importance of sending a message to the community that the law is consistent. But since this punishment outcome will be private, what's the point of being consistent this time? What's wrong with letting him go?

___ **Family.** He should be punished for the victim's family's sake.

Interviewer. Okay. But remember, the victim's family said that they forgave him and didn't desire to seek punishment. Knowing that, would you then excuse the offender?

___ **Culture.** Punishment is part of our cultural/societal practice.

Interviewer. So if the offender committed this rape and murder in another culture where rape and murder is considered okay, would you be comfortable with that? Or would you still feel that he actually deserves punishment?

___ **Legal precedent.** If the offender is excused, other offenders who might be more dangerous might have an easier time getting excused in the future because of this legal precedent.

Interviewer. Okay, but let's suppose that when the court ruled that no information about the crime would be disclosed, this meant it wouldn't even be used as a precedent by the court. Knowing that, would you still want to punish the offender?

___ **Authority.** The bible says "an eye for an eye."

Interviewer. That's true, but the bible also says "love your neighbor as yourself." So, what's wrong with forgiving him and letting him go?

Interviewer. Okay, do you have any other reasons for punishing the offender that we have not already discussed?

___ **Persistent.** [If Retribution 3×]: Well, it's hard to explain, but I just know that he doesn't deserve to be freed.

Received June 13, 2011
Revision received September 6, 2011
Accepted September 9, 2011 ■