

Running head: PSYCHOPATHIC PUNISHMENT

No Skin Off My Back: Retribution Deficits in Psychopathic Motives for Punishment

Eyal Aharoni\*

Lisa L. Weintraub

Alan J. Fridlund

University of California Santa Barbara

\*Corresponding author: Department of Psychology, UC Santa Barbara, Santa Barbara, California 93106; Phone: 805.893.2791; Fax: 805.893.4303; e-mail: [aharoni@psych.ucsb.edu](mailto:aharoni@psych.ucsb.edu)

Keywords: psychopathy, punishment, punitive motives, retribution, consequentialism, intent, aggression, morality, behavior control, criminality, sentencing, jurisprudence.

Acknowledgements: We thank Michael Pecchio for assistance with data collection and Scott Lilienfeld and Teng Ma for helpful comments on this paper.

Abstract

When deciding a criminal's punishment, people typically exhibit both retributive and consequentialist motives in their decision making, though retribution's role may be stronger. This study aimed to discern possible functions of retribution by examining a population predicted to be deficient in retributive drive. Participants who rated either high or low in psychopathic traits read stories about a homicide. These stories were designed to evoke both retribution and the consequentialist motive of behavior control by varying, respectively, criminal intent and likelihood of recidivism. The participants then recommended a length of confinement for the offender. Individuals high in psychopathic traits were uniquely insensitive to retributive cues, and they were particularly consequentialist in their punishment of criminal offenders. These results clarify aspects of psychopathic aggression and corroborate the hypothesis that retribution may stabilize cooperative behavior.

## No Skin Off My Back: Retribution Deficits in Psychopathic Motives for Punishment

People are frequently called upon to punish others in everyday life, from discipline by a mother to demotion by a boss. Nowhere is such punishment clearer than in a court of law, wherein it is exacted by judges and juries. Although punishment is often constrained by legislated sentencing guideline grids and jury instructions, its severity is often affected by the arbiter's implicit motives for punishment. For example, a punishment that is intended to deter an offender is likely to differ considerably from one used to express moral condemnation. Complicating the matter further is the fact that punishment goals are often left unstated and may represent the concurrent operation of multiple motives (Rappaport, 2003). Furthermore, even when the court provides explicit instructions, this does not guarantee that jurors will understand and internalize them (Haney & Lynch, 1994), and in case of jury nullification, jurors may wantonly defy them. By examining how implicit punitive motives differentially shape our punishment decisions, we can better predict how severe a punishment will be favored in a given context and why it is sought at all.

Typically, "punishment" is considered to be any cost deliberately imposed upon a target in response to a violation of a rule (see Bourke, 1977). Left unstated in this definition is what *motivates* this infliction of costs. The many motives for punishment are usually broadly divided into two categories: the desire for retribution and the consequentialist desire for behavior control. Retribution is the use of punishment to "right a wrong," or "restore a moral balance" (Spohn, 2002). Its leading proponents have described the expression of retribution as a self-justified gesture, carrying an underived moral authority (Kant, 1790/1952).

Retribution is retrospective because it concerns only crimes that have already been committed. It is typically accompanied by emotions like outrage and righteous anger (Carlsmith

et al., 2002; Daly & Wilson, 1988; Darley et al., 2000; Fessler & Haley, 2003; Lerner, Goldberg & Tetlock, 1998). Drawing on arguments by Daly and Wilson (1988), Walsh asserted that without retributive emotions, we would lack the motivation required to oppose those who violate reciprocity norms, violators would gain a reproductive advantage, and we might all have evolved into psychopaths (2000). Although the punishment may also have the effect of deterring future transgression, this is not the explicit goal of retribution. As a consequence, no matter how dangerous a person may be, his dangerousness alone is not sufficient to merit a retributive response.

In contrast to retribution, the motive of behavior control is *prospective*—to prevent future crimes from being committed (Vidmar, 2000). The more dangerous the criminal is judged, the more severe punishment is mandated to thwart a recurrence. If an offender is rendered incapable of recidivism (e.g., a burglar loses his arms in an accident), a person motivated solely by behavior control would have no reason to punish.

Studies that have investigated the motivations for punishment in typical populations have found that lay punishment is driven both by retribution and behavior control, although retribution's role may be stronger (Blumstein & Cohen, 1980; Bohm, 1992; Carlsmith et al., 2002; Darley et al., 2000; Hamilton & Rytina, 1980; McFatter, 1982; Rucker, Polifroni, Tetlock & Scott, 2004; Warr et al., 1983).

If the retributive sentiment is indifferent to crime control, what does it do for us, if anything? Like vengeance, retribution could be pathological (see Fridja, 1994). However, much debate has ensued about particular functions of retribution, with conjectures that range from a desire for equity, to costly signaling, to promoting cooperation in collective actions, to restoring welfare or reproductive fitness differentials, and more (Boyd & Richerson, 1992; Daly &

Wilson, 1988; Fehr & Fishbacher, 2004a; Fehr & Gächter, 2002; Fowler, 2005; Gintis, Bowles, Boyd & Fehr, 2003; Kurzban, DeScioli & O'Brien, in press; Price, Cosmides & Tooby, 2002; Trivers, 1971; Turillo, Folger, Lavelle, Umphress & Gee, 2002).

One clue to this puzzle might be to examine the form of a person's punitive motives when the retributive response is deficient or absent. In other words, the exception may help us to understand the rule. So in whom is retribution absent? Might certain personality types, for example, be relatively disposed against retribution? One candidate may be the psychopathic personality.

### The Psychopathic Personality

Psychopaths are individuals who seem to lack a conscience. By some estimates, they comprise approximately 2-3% of the population. Early research contended that psychopaths were deficient in the ability to distinguish right from wrong (Bushong, 1934). More recent evidence has suggested that psychopaths do understand the concepts of "right" and "wrong" abstractly, but do not use them to regulate their own behavior (Cleckley, 1976; Hare, 1993; Patrick, 1994). In support of this claim, Simon, Holzberg, and Unger (1951) found that when solutions to social conflicts were presented in a multiple-choice format, psychopathic women did not differ from female controls. The two groups did differ, however, when the requests were open-ended. This result suggests that psychopathic women can identify socially appropriate responses but then fail to apply them in concrete situations when left to their own devices<sup>1</sup>. In another study, Link, Sherer, and Byrne (1977) found that psychopathic inpatients scored higher on a morality scale than non-psychopaths, showing that moral knowledge and moral conduct may diverge in the psychopathic mind. Findings such as these have met with the peremptory objection that psychopathic accuracy in identifying moral transgressions is not genuine but

results instead from social desirability effects (Blair, 1995); this interpretation remains to be demonstrated empirically.

Along with transgressive tendencies, psychopaths often have been characterized as particularly prone to aggressive, antisocial behavior (Elliott & Gillett, 1992; Hare, 1993; Lindner, 1944; Link et al., 1977; Lykken, 1995)<sup>2</sup>. In contrast to that of non-psychopaths, however, psychopathic aggression is thought to be largely proactive rather than reactive (Blair, 2001; Cornell, Warren, Hawk, Stafford, Oram, & Pine, 1996; Raine, Dodge, Loeber, Gatzke-Kopp, Lynam, Reynolds, Stouthamer-Loeber, Liu, 2006).

This difference has implications for the emotionality of the aggressor. Reactive aggression (also known as hostile, moralistic, retaliatory, or angry aggression) is typically accompanied by righteous anger, and its psychological functions are believed to discourage antisocial behavior and facilitate cooperation (Buss, 1961; Clutton-Brock & Parker, 1995; Daly & Wilson, 1988; Frank, 1988; Gintis et al., 2003; Kingsbury, Lambert & Hendrickse, 1997; Kurzban et al., 2007; Price et al., 2002; Sell, 2005; Trivers, 1971). Proactive aggression, on the other hand, functions instrumentally, to control or manipulate the behavior of others for personal gain (Kingsbury et al., 1997). Because the psychopath may punish more proactively, righteous anger may be conspicuously absent. As Cleckley said of the classic psychopath, “mature, wholehearted anger, true or consistent indignation, honest, solid grief, sustaining pride, deep joy, and genuine despair are reactions not likely to be found within [the psychopath’s emotional] scale” (p. 348, 1976).

The distinction between retroactive and proactive aggression can be seen as mapping onto the twin motives of punishment. Conceived in this way, punishments motivated by behavior control serve as examples of proactive aggression, whereas those motivated by retribution are

characteristic of reactive aggression. Assuming these stipulations, it might be predicted that the punishment enacted by psychopaths would be motivated primarily by behavior control and would lack the retributive influence operating in typical, more prosocial, populations. If so, then the relative absence of the retributive response from the psychopath's mental toolkit would be consistent with the evidence that retribution functions to facilitate cooperative, prosocial behavior.

### Measuring Retribution

How can it be determined whether someone is motivated reactively, to exact retribution, or proactively, to control behavior? We discuss separately how each motive might be surmised.

Several properties of retribution permit the discerning of its operation in punishment decisions. First, retribution should be relatively indifferent to cues diagnostic of recidivism because, as the argument goes, people should be punished for what they have done, not necessarily for what they *might* do.

Additionally, according to the principle of proportionality, exactly how retributive a person is toward an offender will depend upon the seriousness of the offense. For this reason, researchers who have wanted to observe changes in retributive sentiments have typically manipulated offense seriousness (Carlsmith et al., 2002; Darley et al., 2000; McFatter, 1982; White, 1975). Despite its reasonableness, this method carries the limitations that crime seriousness does not exclude the operation of behavior control. This is because crime seriousness often implies crime *costliness* as well—and the non-retributive person motivated by behavior control should be especially motivated to reduce the likelihood of costly crimes, whether the costs are physical, emotional, or financial (Warr et al., 1983). As such, crime seriousness does not help to distill the relative contributions of the two motives.

How retributive a person will be should also depend on attributions of *intent*. Research has shown that punishments are more severe for offenses that are judged more intentional, even when the intended consequences of the criminal act are thwarted (Baron & Richardson, 1994: p. 145, 210; Geen, 1968; Greenwell & Dengerink, 1973; White, 1975). The reason intent has proven more important than consequence may be because it signals more reliably just how highly the perpetrator regards our welfare (Epstein & Taylor, 1967; Nickel, 1974; Ohbuchi & Kambara, 1985; Sell, 2005). Criminal intent also can lead people to infer how likely an offender is to recidivate, but this relationship can be neutralized by convincing judges that an offender who transgressed intentionally is, nonetheless, not at risk of recidivating. In the present experiment, we posit that retributive sentiments will be particularly sensitive to criminal intent (see also Horai & Bartek, 1978; Horan & Kaplan, 1983; Vidmar & Crinklaw, 1974).

#### Measuring Behavior Control

Ascertaining the operation of behavior control requires a different strategy. Experimentally, this motive is usually manipulated by adjusting the putative offender's criminal history to range from nil to extensive (Darley et al., 2000). This method is imperfect, partly due to the fact that criminals with longer histories are not necessarily believed to be more dangerous. Importantly, rates of recidivism tend to decrease in late adulthood, so longer criminal histories can actually become predictive of *decreased* likelihood of recidivism (Hoffman & Beck, 1984). This effectively lowers the ceiling on any bipolar scale designed to measure variation in recidivism estimates. There is another problem with manipulating criminal history: Whether the offender has a criminal history or not, he has in each case been convicted at least once, leading judges to perceive both as relatively dangerous, thus forcing a high floor on any punishment

scale. Finally, criminal history can provoke retribution in and of itself, should the punisher believe that the target offender has not atoned for his past misdeeds.

One way to maximize scale variability and avoid conflating behavior control with retribution might be to dispense with the criminal history manipulation entirely and, instead, vary likelihood of recidivism more directly—by appealing to expert testimony. Participants can be informed that an offender was given a series of valid psychological assessments and found by behavioral experts to be either highly likely or unlikely to recidivate. Of course, the success of this type of manipulation would rely upon participants' confidence in the experts' findings. If such confidence could be achieved, then expert testimony would be a direct and economical way to circumvent problems associated with manipulating criminal history.

When varied orthogonally, cues of criminal intent should evoke punishments motivated more by retribution, and expert testimony on recidivism should evoke punishments motivated more by behavior control.

### Study Overview and Hypotheses

Previous research led us to predict that both high psychopathy individuals (HPIs) and low psychopathy individuals (LPIs) will increase punishment when a threat of recidivism is present compared to when it is absent (Blair, 2001; Cornell et al., 1996; Darley et al., 2000; Raine et al., 2006; Rucker et al., 2004). This is because both groups should be nominally motivated to control dangerous behavior. However, controlling for threats of recidivism, HPIs should be less reactive to intent than LPIs. That is, compared to HPIs, LPIs should punish more severely if the crimes are intentional. In statistical terms, we thus expect both a main effect for threat and an interaction between intent and psychopathy.

Modeled after research by Darley and colleagues (2000), participants read vignettes about a homicide committed by an adult male and were given varying information about the *criminal intent* and the *likelihood of recidivism*. These constructs were designed to engage the motives of retribution and behavior control, respectively. Participants were separated into *high vs. low psychopathy* as measured by their scores on the Psychopathic Personality Inventory (PPI). The PPI has been established in a number of studies to be a valid and reliable indicator of psychopathic personality traits (Cale & Lilienfeld, 2006; Lilienfeld, 1990; Lilienfeld & Andrews, 1996).

The experiment used a 2 (high/low psychopathy) x 2 (high/low intentionality) x 2 (high/low likelihood of recidivism) between-subjects factorial design. This is a variant of what has been termed an “extreme-groups approach” (see Preacher, Rucker, MacCallum & Nicewander, 2005). We discuss our choice of this design and its benefits and limitations later in this article.

One challenge we faced was how to obtain our participants’ punishment decisions in all their complexity. A classic measure of criminal punishment, of course, has been to solicit univocal recommendations for prison sentence length. However, we have come to believe that such traditional prison-sentencing measures may actually introduce bias. One instance may be in cases of unintentional offenders who are nonetheless likely to recidivate, such as some mentally ill offenders. Here, a prison sentence measure would be insensitive to behavior-control punishments simply because this type of punishment is considered inappropriate for unintentional offenders. Commitment to a psychiatric hospital, in contrast, may be more indicative of behavior-control responses toward these offenders.

In attempt to capture a more equal probability of detecting both punitive motives, we designed three related questions. First, participants indicated how strongly they felt the offender should be placed in a *prison*, then, how strongly the offender should be placed in a *mental institution*. Then, independently of these two ratings, they assigned our primary dependent variable of interest, a *confinement length* for the offender. This three-question structure permitted us to examine attitudes toward confinement length by itself, and also in relation to particular confinement sites. Although this approach forsakes the simplicity of the traditional one-question approach, we believe that it captures more of the complexity of real-life punishment decisions.

We used the vignette method in this study because the crime descriptions can be rendered roughly equivalent to the hearing of actual court cases. Although participants are fully aware that their task is purely hypothetical, this knowledge does not appear to have a significant impact on participants' responses (Wilson & O'Gorman, 2003).

## Methods

### Participants

Participants were selected based on their scores on the Psychopathic Personality Inventory (PPI). The PPI was administered to 718 undergraduate males and females from an introductory psychology course at the University of California, Santa Barbara (UCSB). They were given course credit in exchange for their involvement. All experimental procedures met accepted ethical guidelines and were approved by the UCSB Committee on Research.

Consistent with an extreme-groups design, participants whose scores were greater than one standard deviation above the population mean ( $M = 137.57$ ,  $SD = 12.67$ ) were classified as high psychopathy individuals (HPIs), whereas those whose scores were less than one standard

deviation below the mean were classified as low psychopathy individuals (LPIs). Our final sample consisted of 24 HPIs (9 F, 15 M) and 29 LPIs (23 F, 6 M). Mean age was 18.6 yr. ( $SD = .95$ ).

### Materials

Participants read one of four vignettes that differed in terms of the two manipulated variables, the intentionality of the offense and the offender's likelihood of recidivism. "Intentionality" was manipulated by stipulating the presence or absence of a brain tumor that caused the offender's violent behavior. While the brain tumor manipulation is not the most common representation of low intent seen in criminal courts, it has been used in experimental research of this kind because it is a relatively unambiguous way of manipulating the alleged "cause" of the violent act, and it permits more credible portrayals of high likelihood of recidivism than do other forms of low intent (see Darley et al., 2000)<sup>3</sup>.

Likelihood of recidivism was manipulated by a purported psychological assessment which rendered an authoritative judgment on whether or not the offender was likely to recidivate. For example, participants in the low-intentionality, low-likelihood-of-recidivism condition read the following passage:

Oren is 56 years old. He has retired from his job as a customer service representative for a small telephone company. Recently, while walking in his neighborhood, he ranted at a passerby and hit the man with a crowbar that he grabbed from a construction site. The victim was killed.

A medical examination after Oren's arrest reveals that he had a growing brain tumor in an area in which tumors consistently lead to violence. Medical experts all agree that the tumor, pressing on a critical brain area, was the cause of the violent incident.

Oren had hallucinated that the passerby was a devil who was attacking him, and so he acted in "self defense."

New surgical developments allow doctors to operate to remove the tumor and this is done. A reliable neuropsychological assessment test has determined that Oren is extremely unlikely to commit more crimes in the future.

After reading the passage, participants were asked to rate "How strongly do you feel the criminal should be freed? detained? executed?" 1 (not at all) to 11 (extremely); "Assuming the criminal is detained, how strongly do you feel he should reside in a prison? mental institution?" 1 (not at all) to 11 (extremely); and "How long do you think the criminal should be confined?" (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). As checks on our manipulations, we asked participants how intentional they considered each offense and how likely the offender was to recidivate.

### Procedure

Participants were run in separate cubicles. After first providing written informed consent, they were asked to complete the vignette questionnaire on the computer. Next, participants completed the manipulation checks. (See Appendix.) In order to verify that there were no inadvertent, overriding experimental demand characteristics, participants were also asked to surmise the purpose(s) of the study. Before the participants were dismissed, they were given a debriefing statement which informed them of the actual purposes of the study.

### Results

*Manipulation checks.* First, we sought to verify that the offenders who were stipulated to commit their crime intentionally were in fact rated as acting more intentionally than those stipulated to transgress unintentionally. A three-way analysis of variance (ANOVA) was

conducted with intent, likelihood of recidivism, and psychopathy as independent variables, and the results showed that this was the case,  $F(1, 53) = 86.73, p < .001, \eta_p^2 = .65$ . (See Table 1.) We also observed an intent x psychopathy interaction,  $F(1, 53) = 4.33, p < .05, \eta_p^2 = .08$ . Pairwise comparisons, using Fisher's Least Significant Difference test, revealed that HPIs' ratings of intent were slightly more extreme than those of the LPIs. This interaction weighs against the hypothesis that intent would show a comparatively lower influence on the variance in the HPI's punishments. There were no other interactions.

The same test was performed to gauge participants' estimates of the offenders' likelihood of recidivism. Again, the offenders stipulated to be highly likely to recidivate were indeed perceived as much likelier to recidivate than those stipulated to be unlikely to recidivate,  $F(1, 53) = 49.59, p < .001, \eta_p^2 = .51$ . We also found a smaller independent effect of intent,  $F(1, 53) = 6.24, p < .05, \eta_p^2 = .12$ , whereby low-intent offenders were seen as more likely to recidivate than high-intent offenders. That is, offenders whose offense was alleged to be caused by a brain tumor were believed to be at some greater risk of recidivism than those whose offense ostensibly resulted from free will. Although unexpected, this finding verifies that increases in punishment due to increased intent cannot be caused by the likelihood of recidivism manipulation<sup>4</sup>. There were no interactions.

[Insert Table 1 here]

According to our first hypothesis, both HPIs and LPIs would assign longer confinements to high-threat offenders compared to low-threat offenders. To test this, we conducted a three-way, between-subjects ANOVA (intent x likely x psychopathy) on our dependent variable, confinement length. Consistent with our hypothesis, both groups gave longer confinements for

higher-threat offenders,  $F(1,46) = 72.02, p < .001, \eta_p^2 = .61$ . We did not find a main effect for psychopathy,  $F(1,46) = 2.06, p = .16, \eta_p^2 = .04, 1-\beta = .29$ . (See Table 2 & Figure 1.)

[Insert Table 2 here]

[Insert Figure 1 here]

Interestingly, we also observed an interaction in which HPIs were even more lenient than LPIs when it came to the confinement of criminals who were unlikely to recidivate,  $F(1,46) = 6.71, p < .05, \eta_p^2 = .13$ . This result is consistent with our prediction, but inconsistent with the operation of reactive or opportunistic aggression.

Our second hypothesis was that psychopathy would interact with intentionality. Specifically, HPIs would assign similar confinement lengths for intentional and unintentional crimes, whereas LPIs would assign longer confinements for intentional crimes than for unintentional crimes. Consistent with our hypothesis, we found a marginal interaction between psychopathy and intentionality for confinement length,  $F(1,46) = 3.17, p = .082, \eta_p^2 = .07, 1-\beta$  (power) = .415. HPIs gave similar confinement lengths for both intentional and unintentional crimes, whereas LPIs tended to give longer confinements for intentional crimes than for unintentional ones.

Pairwise comparisons also showed that, with respect to intentional crimes, the HPIs' confinements were shorter than those of the LPIs,  $p < .05$ . This excludes the interpretations that HPIs were insensitive to intent simply because they were more punitive across these conditions. We also found a main effect for intent, suggesting that both groups are to some degree sensitive to intent,  $F(1,46) = 5.38, p < .05, \eta_p^2 = .11$ ; this finding is admittedly qualified by the psychopathy x intent interaction (above). (See Table 2 & Figure 2.)

[Insert Figure 2 here]

We conjectured that the psychopathy x intent interaction may have only achieved marginal significance because the low intent cases that were crossed with high likelihood of recidivism were assigned relatively lengthy confinements. In other words, the demand to assign a behavior-control punishment for the dangerous mentally ill offender may have masked the desire to withhold retributive punishment from the low-intent offenders generally. If so, we would expect confinement site to moderate the effect of confinement length among LPs, but not HPs. That is, for LPs, the effect of intent on confinement length should differ from zero when factoring in support for prison, whereas, for HPs, it should disappear.

We tested this hypothesis by regressing intent on confinement length for LPs and HPs with *confinement site* as a moderating variable. The confinement-site variable was computed by subtracting support for an institution from support for prison such that positive values represented support for prison, and negative numbers represented support for institutionalization. As predicted, the regression yielded a significant effect of intent among the LPs,  $\beta = .595$ , partial  $r = .459$ ,  $p < .05$ , but not for the HPs,  $\beta = .310$ ,  $p = .346$ . This result supports the hypothesis that intent was indeed relevant to the LP's punishment recommendations, but not to those of the HPs.

### Discussion

Our data produced three important results. First, confinements assigned by both HPs and LPs were longer for those criminals with a high likelihood of recidivism than for those with a low likelihood of recidivism. This suggests that both groups are instrumentally motivated to regulate threatening behavior. Our second major finding was a clear divergence in the responses of LPs and HPs toward variations in criminal intent: Unlike the LPs, the HPs consistently gave similar punishments regardless of variations in intent, suggesting that they are

comparatively insensitive to retributive motives for punishment. Third, we found that the HPIs were less punitive than the LPIs toward both intentional offenders and low-recidivism offenders. These results confirm our hypothesis that HPI punishment is predominantly motivated by behavior control.

That HPIs punished mainly out of behavior control, not retribution, supports the argument that psychopaths are rationalistic in their thinking (Deigh, 1996) but are impaired with respect to moral emotions (Blair et al., 1995). Folk notions of psychopathy would predict that this impairment leads to increased hostile or opportunistic aggression, and research has suggested that these forms of aggression are indeed associated with psychopathy (Lindner, 1944; Widiger & Lynam, 1998). For example, Book and Quinsey (2004) argued that psychopaths cannot draw on empathy displays in order to signal a sense of fairness, and so they compensate by increasing displays of indignation, which often produce excessive aggression. Their data, however, may better describe unsuccessful psychopaths in personal exchanges. In contrast, our finding that HPIs were less punitive toward both high-intent and low-recidivism offenders may reflect instead a non-clinical sample making third-party judgments. We will discuss this topic in greater detail below.

Some have theorized that consequentialist punishment is, in principle, unconstrained in its magnitude (McClosky, 1965; Parfit, 1984). Importantly, this study was not designed to discover when consequentialists maximize aggression, but holding those conditions aside, it does demonstrate that consequentialist strategies do not necessarily produce hostile or opportunistic aggression. Instead, consequentialist strategies, including psychopathic ones, may lead to attenuated punishments under certain conditions. Specifically, such strategies should be conservative to the extent that the dispensation of punishment is costly and the putative offender exhibits no real or imagined threat. Where psychopathic punishment and conventional

consequentialist punishment may diverge is that conventional consequentialism is usually concerned with threats against society in general, whereas psychopaths should only be motivated to reduce threats to their own self-interest (see Cale & Lilienfeld, 2006). The present experiment did not control how costly it was to dispense punishment, nor did it distinguish personal from societal threats. Future research should test our hunch that conservative psychopathic punishment is tailored largely to serve self-interest.

If punishment is normally perceived as costly, why didn't HPIs rate the intent values even lower? One reason may be that despite our experimental efforts to cross intent with likelihood of recidivism, intent normally is often a reliable *predictor* of recidivism. In other words, offenders who have transgressed intentionally are likelier on average to recidivate than those who have transgressed without intent. Also, our participants rated the low-intent offender as still somewhat likely to recidivate. For these reasons, we made no predictions about the absolute values of the HPIs' punishment.

#### *Functions of Retribution.*

We argued previously against the notion that consequentialist strategies would necessarily lead to greater aggression than moralistic ones. In fact, our data suggest that the opposite may be true, at least in conditions of high intent and low recidivism. This result forces us to question exactly why the LPIs are prone to inflict punishment upon intentional offenders despite the assurances that these offenders are not dangerous. In other words, perhaps what demands explanation is not why psychopaths are consequential, but why non-psychopaths so predictably render punishments that prove futile in the face of unwanted behavior. Kant would invoke a fundamental moral sense; we, like many others, seek a functionalist explanation.

What do the present results tell us about the possible functions of retribution? One defining feature of the psychopath is his pursuit of self-interest at the expense of others' welfare. As long as there are opportunities to exploit the aid of others, psychopaths should be willing to aggress, undoubtedly, to levels far beyond those measured in the present study. Once identified, non-psychopaths should perceive psychopaths as terrible candidates for cooperation.

The fact that psychopaths are low in retributive sentiments is consistent with the hypothesis that retribution facilitates prosocial, cooperative behavior (see Boyd & Richerson, 1992; Clutton-Brock & Parker, 1995; Daly & Wilson, 1988; Fehr & Gächter, 2000, 2002; Fowler, 2005; Frank, 1988; Gintis et al., 2003; Kurzban et al., 2007; Price et al., 2002; Trivers, 1971; Walsh, 2000). If so, these results would complement data by Fehr and Gächter (2000, 2002), among others, which suggest that punishment facilitates cooperation in iterated economic relationships. Seen in this light, psychopaths may be characterized as prone not only to cheating on collective actions, but also as prone to free-riding on the collective enforcement of punishment of other cheaters. This is not to say that psychopathic behavior is less functional – only that it may occupy a different niche in environments rich in social competition and cooperation. Certainly, in terms of population dynamics, it can be argued that these purely instrumental strategies would be unstable in large proportions, yet they have the potential to confer obvious advantages when they occur sporadically amid the context of largely prosocial strategies.

If the retributive motive is a solution to an adaptive problem of cooperation, then there must be certain conditions prerequisite for its activation. Our data suggest that signals of intent may provide reliable input conditions for a retributive response given that they may have been reliable predictors of recidivism in our evolutionary past. Although a complete theory of intentionality does

not presently exist, research in agency and Theory of Mind has had success in determining various cues that signal intent (Heider and Simmel, 1944; Leslie, 1994b).

Ironically, the goals served by retribution may ultimately be instrumental—with respect not to momentary self-interest per se, but instead to the long-term reproductive advantage it may have produced. Exactly how a retribution adaptation may be designed, and whether its absence from psychopathic behavior is itself an adaptation, are ongoing questions critical to our understanding of both prosocial and antisocial aggression and punishment. Their answers, by allowing fuller descriptions of human aggression and punishment, may inform legislation and public policy relating to crime prevention, intervention, jury selection, and jury instruction. Such detailed models of retribution may provide a means to make laws that both consider our natural dispositions and integrate checks against them.

#### *Limitations and Future Directions.*

##### Use of Extreme-Groups Design

Our conclusions, of course, are necessarily limited by our procedures. Because of the exploratory nature of our hypotheses, our primary strategy for this experiment was simply to identify a group of individuals in our population who most closely resemble classic psychopaths and then compare them to a group that least resembles them. This “extreme-groups design” allowed a cost-effective way to test our hypothesis, but these designs all carry notable drawbacks: They are known to violate assumptions of normality and linearity, and they tend to inflate variances and effect sizes (Preacher et al., 2005).

The effects reported in this experiment are subject to all of these concerns and so must be interpreted conservatively. Accordingly, we emphasize that the findings from our *F*-statistics

alone likely represent a real—albeit nuanced—cluster of individuals in the larger population. In light of recent findings using a typical sample in an identical design, we note that LPIs punished similarly to typical respondents: When confinement site was controlled, they both increased punishment as a function of increasing intent (Aharoni & Fridlund, under review). This adds to our confidence that the unique pattern of punishment observed in our HPIs was not spurious. Future attempts to extend these findings should be done with much larger Ns, permitting a regression design so that individuals across the psychopathy scale can be compared parametrically.

#### Assumption of Diagnostic Homogeneity

In our analyses, we have characterized psychopathic punishment as a unitary phenomenon. This is an acknowledged simplification, given the heterogeneity among different construals of “psychopathy.” Murphy and Vess (2003) found evidence of many subtypes of psychopaths, some of which exhibit hostile traits like impulsivity and sadism that seem to diverge from the instrumental aggression observed in our sample. Thus, it could be that the decreased aggression observed in psychopathic punishment resulted from a measurement failure in identifying hostile psychopaths. We note, however, that the PPI was designed specifically to capture traits commonly associated with hostile aggression (impulsivity, nonconformity, blame externalization, etc.) as well as instrumental aggression (social potency, stress immunity, etc.) and has been used successfully to assess violent, criminal populations (Cale & Lilienfeld, 2006; Lilienfeld, 1990; Lilienfeld & Andrews, 1996). Furthermore, others have validated these trait clusters using a two-factor model (Benning, Patrick, Hicks, Blonigen & Krueger, 2003). Taken together, it appears that hostile traits like impulsivity and sadism are not definitive of psychopathy although they may be present in some

psychopaths (see also Millon & Davis, 1998). A large sample size would permit future analyses to account for subtle differences in psychopathic aggression as a function of such trait clusters.

Another reason we may not have found evidence of reactive or opportunistic aggression is because of the population we sampled. Many psychologists have distinguished between the more intelligent and controlled “white-collar” psychopath and the less cerebral and more impulsive “blue-collar” psychopath (Hare, 1970; Lykken, 1995). White-collar psychopaths are probably more likely to end up in college than blue-collar psychopaths, thus biasing our sample.

Given the relatively small point-prevalence of psychopathy, we chose to examine individuals who rate high in psychopathic personality traits as opposed to individuals expressly diagnosed as psychopaths. Nonetheless, our strict criterion for qualifying as an HPI significantly restricted our sample size. For this reason, some of our analyses suffered from low statistical power. Using a larger sample size, future research should address how well these results generalize to clinical psychopaths.

#### Use of Single Homicide Offense

Another limitation in this study was that homicide was the only offense evaluated in each condition. We focused on homicide for purposes of experimental control and because homicide normally incites moral condemnation and strong associated emotions. It is nonetheless possible that different crimes, as well as non-criminal offenses, may provoke different psychopathic motives. Given the present findings, we would expect psychopathic punishment to vary with the perceived costliness of the offense and its likelihood of recurrence, but not with respect to uniquely retributive cues like moral seriousness. To test this prediction, future research should select crimes in which costliness, recidivism, and moral seriousness can be manipulated independently.

### Possible Sex Confound

There is some evidence that psychopathy may manifest differently in men and women. For example, Strand and Belfrage (2005) found that psychopathy predicted antisocial behavior in men, and lying, deceit, and lack of control in women (see also Grann, 2000). Other data suggest that psychopathy may “hide” among females diagnosed with conditions such as somatization disorder and Munchausen’s-by-proxy (see Cale and Lilienfeld, 2002). In our experiment, 79.3% of our LPI sample was female while 62.5% of our HPI sample was male. Thus, our results are vulnerable to a sex confound. However, this prediction that male psychopaths are more antisocial works against our hypothesis, which correctly predicted that psychopaths will exhibit relative leniency any time the offender is not a threat. With equal numbers of male and female psychopaths, Strand and Belfrage’s findings compel us to predict perhaps even greater leniency than we observed. Moreover, the gender inequality in our sample is consistent with the usual diagnosed prevalences in the literature. A larger sample would be needed to investigate whether psychopathic males and females show similar differences in their punishment behavior.

### Construct Validity of Mental Illness Manipulation

It is possible that punitive motives toward the mentally ill differ from punitive motives toward more typical cases of low intent, like offenses resulting from negligence. People may not necessarily treat the mentally ill as free of intent, although our manipulation checks showed that they were indeed perceived as less intentional. We also reported evidence above that our mentally ill offender was perceived as likelier to recidivate than the intentional offender. These points may explain why our marginal psychopathy x intent interaction did not achieve significance before factoring in the confinement site moderator.

Importantly, our primary goal for the low-intent condition was simply to characterize an offense whose cause can be attributed to a source outside the offender's rational free will<sup>5</sup>. As mentioned, we chose this particular manipulation to increase the credibility of our likelihood manipulation. Thus, we cannot make definitive claims that the behavior-control punishments directed at our low-intent offender will generalize to other types of low-intent cases. However, we do note that mental illness has long been judged as a mitigating factor in legally assigning responsibility for violent crimes (Cooke & Pogany, 1975; Howells, McEwan, Jones, & Mathews, 1983; Monahan & Hood, 1976). Also, behavior control is a prominent response to accidents caused by mental illness (Quinsey, Cyr & Lavalley, 1986). The present experiment confirmed these findings: Both HPIs and LPIs rated the mentally ill offender as less intentional, and HPI punishment was less influenced by intent. Taken together, we expect that these findings will generalize to other unintentional offenses.

#### Experimental Control of Assessment Type

The effects of intent upon punishment decisions in our study may have been affected by small confounds in the language used to describe the high- and low-intent offenders. For example, the assessment test in the high-intent condition was described as a "psychological" assessment, while the assessment test in the low-intent condition was described as a "neuropsychological" assessment. We chose this wording specifically to make each assessment test ecologically appropriate to its context, namely, that neuropsychological assessments are appropriate for brain tumor patients but are inappropriate for intentional offenders. Conceivably, if our participants believed that psychological predictions were less credible than neuropsychological predictions, they might have punished the high-intent offender more leniently. This could also explain why our marginal intent interaction did not achieve significance before factoring in confinement site.

Scenarios that control such variables without compromising ecological validity will serve to purify any effects of intent on punishment.

None of these limitations mitigates the importance of our central finding that a significant proportion of punishment strategies commonly associated with the psychopathic personality are uniquely unresponsive to cues that typically evoke retribution. Furthermore, this insensitivity contributes to the argument that typical expressions of retributive punishment may play a regulatory role against exploitative, antisocial behavior.

## References

- Aharoni, E. & Fridlund, A. J. (under review). *Not just retribution: The role of behavior control in third-party punishment*.
- American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders*, (4<sup>th</sup> ed.). Washington, DC: American Psychiatric Association.
- Baird, S.A. (2002). The links between primary and secondary psychopathy and social adaptation. *Colgate University Journal of the Sciences* (34), 61-82.
- Barlow, D.H., & Durand, V.M. (2002). *Abnormal Psychology*. Pacific Grove: Brooks/Cole.
- Baron R. & D. Richardson (1994). *Human Aggression*, New York: Plenum.
- Blair, R. J. R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition*, 57, 1-29.
- Blair, R. J. R. (2001). Neurocognitive models of aggression, the antisocial personality disorders, and psychopathy. *Journal of Neurology, Neurosurgery and Psychiatry*, 71, 727-731.
- Benning, S. D., Patrick, C. J., Hicks, B. M., Blonigen, D. M. & Krueger, R. F. (2003). Factor structure of the psychopathic personality inventory: Validity and implications for clinical assessment. *Psychological Assessment*, 15(3), 340-350.
- Blumstein, A., Cohen, J. (1980). Sentencing of convicted offenders: An analysis of the public's view. *Law and Society Review*, 14(2), 223-261.
- Bohm, R. (1992). Retribution and capital punishment: Toward a better understanding of death penalty opinion. *Journal of Criminal Justice*, 20, 227-236.
- Book, A. S. & Quinsey, V. L. (2004). Psychopaths: Cheaters or warrior-hawks? *Personality and Individual Differences*, 36, 33-45.
- Bourke, V. J. (1977). The ethical justification of legal punishment. *American Journal of*

- Jurisprudence*, 22, 1-19.
- Boyd, R. & Richerson, P. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, 13, 171-195.
- Bushong, R. E. (1934). Report of the committee on psychopathic personality. *Managing Officers Association of the State Institutions of Ohio. Quarterly Bulletin*. 11 (4), 1934, 7-12.
- Buss, A. H. (1961). *The psychology of aggression*. New York: Wylie.
- Cale, E.M., & Lilienfeld, S. O. (2002). Sex differences in psychopathy and antisocial personality disorder. A review and integration. *Clinical Psychology Review*, 22, 1179-1207.
- Cale, E. M. & Lilienfeld, S. O. (2006). Psychopathy factors and risk for aggressive behavior: A test of the “threatened egotism” hypothesis. *Law & Human Behavior*, 30(1), 51-74.
- 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(2), 284-299.
- Cleckley, H. (1976). *The mask of sanity* (5<sup>th</sup> Ed.). Saint Louis: The C. V. Mosby Company.
- Clutton-Brock, T. & Parker, G. (1995). Punishment in animal societies. *Nature*, 373, 209-216.
- Cooke, G. & Pogany, E. (1975). The influence on judges' sentencing practices of a mental evaluation. *Bulletin of the American Academy of Psychiatry and the Law*, 3, 245-251.
- Cornell, D. G., Warren, J., Hawk, G., Stafford, E., Oram, G. & Pine, D. (1996). Psychopathy in Instrumental and Reactive Violent Offenders. *Journal of Consulting and Clinical Psychology*. 64 (4), 783-790.
- Daly, M. & Wilson, M. (1988). *Homicide*. New York: 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*(6), 659-683.
- Deigh, D. (1996). Empathy and universalizability. In L. May, M. Friedman, A. Clark (Eds.), *Mind and morals: essays on cognitive science and ethics*, (pp. 743-763). Cambridge: The MIT Press.
- Elliott, C. & Gillett, G. (1992). Moral insanity and practical reason. *Philosophical Psychology, 5*(1), 53-67.
- Epstein, S. & Taylor, S. (1967). Instigation to aggression as a function of degree of defeat and perceived aggressive intent of the opponent. *Journal of Personality, 35*(2), 265-289.
- Fehr, E. & Fischbacher, U. (2004a). Third-party punishment and social norms. *Evolution and Human Behavior, 25*, 63-87.
- Fehr, E. & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review, 90* (4), 980-994.
- Fehr, E. & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415* (10), 137-140.
- Fessler, D. & Haley, K. (2003). The strategy of affect: emotions in human cooperation. In Hammerstein, P. (Ed.) *Genetic and cultural evolution of cooperation* (p. 7-36). Cambridge: MIT Press.
- Fowler, J. (2005). Altruistic punishment and the origin of cooperation. *Proceedings of the National Academy of Sciences, 102* (19), 7027-7049.
- Frank, R. H. (1988). *Passions Within Reasons*, New York: W. W. Norton & Company.
- Fridja, N. (1994). Lex talionis: On vengeance. In Goozen, S., Van de Poll, N., & Sergeant, J. (Eds.) *Emotions: Essays on emotion theory*. Lawrence Erlbaum Associates, Hillsdale, NJ.

- Geen, R. (1968). Effects of frustration, attack and prior training in aggressiveness upon aggressive behavior. *Journal of Personality and Social Psychology*, 9(4), 316-321.
- Gintis, H., Bowles, S., Boyd, R. & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, 24, 153-172.
- Grann, M. (2000). The PCL-R and gender. *European Journal of Psychological Assessment*, 16(3), 147-149.
- Greenwell, J. & Dengerink, H. (1973). The role of perceived versus actual attack in human physical aggression. *Journal of Personality and Social Psychology*, 26(1), 66-71.
- Hamilton, V. & Rytina, S. (1980). Social consensus on norms of justice: Should the punishment fit the crime?" *American Journal of Sociology*, 85, 1117-44.
- Haney, C. & Lynch, M. (1994). Comprehending life and death matters: A preliminary study of California's capital penalty instructions. *Law and Human Behavior*, 18(4), 411-436.
- Hare, R. D. (1970). *Psychopathy: theory and research*. New York: John Wiley & Sons, Inc.
- Hare, R. D., Hart, S. D. & Harpur, T. J. (1991). Psychopathy and the DSM-IV criteria for antisocial personality disorder. *Journal of Abnormal Psychology*, 100(3), 391-398.
- Hare, R. D. (1993). *Without conscience: the disturbing world of the psychopaths among us*. New York: The Guilford Press.
- Heider, F. & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, 57(2), 243-259.
- Hoffman, P. & Beck, J. (1984). Burnout: Age at release from prison and recidivism. *Journal of Criminal Justice*, 12 (6), 617-623.
- Horai, J. & Bartek, M. (1978). Recommended punishment as a function of injurious intent, actual

- harm done, and intended consequences. *Personality & Social Psychology Bulletin*. 4(4), 575-578.
- Horan, H. & Kaplan, M. (1983). Criminal intent and consequence severity: Effects of moral reasoning on punishment. *Personality & Social Psychology Bulletin*. 9(4), 638-645.
- Howells, K., McEwan, M., Jones, B., & Mathews, C. (1983). Social evaluations of mental illness (schizophrenia) in relation to criminal behavior. *British Journal of Social Psychology*, 22, 165-166.
- Kant, I. (1952). *Kant's Critique of Judgment*, (J.C. Meredith, Trans.) Oxford: Clarendon Press. (Original work published in 1790)
- Kingsbury, S. J. Lambert, M. T. & Hendrickse, W. (1997). A two-factor model of aggression. *Psychiatry*, 60, 224-232.
- Kurzban, R., DeScioli, P., & O'Brien, E. (2007). Audience effects on moralistic punishment. *Evolution and Human Behavior*.
- Lerner, J. S., Goldberg, J. H., & Tetlock, P. E. (1998). Sober second thought: the effects of accountability, anger, and authoritarianism on attributions of responsibility. *Personality and Social Psychology Bulletin*, 24, 563-574.
- Leslie, A. M. (1994b). ToMM, ToBy, and agency: Core architecture and domain specificity. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*. (pp. 119-148). Massachusetts: Cambridge University Press.
- Levenson, M. R. (1992). Rethinking psychopathy. *Theory & Psychology*, 2(1), 51-71.
- Lindner, R. M. (1944). A formulation of psychopathic personality. *Psychiatry*, 7, p. 59-63.
- Link, N. F., Sherer, S. E., & Byrne, N. P. (1977). Moral judgment and moral conduction in the psychopath. *Canadian psychiatric associations journal*, 22(7), 341-346.

- Lilienfeld, S. O. (1990). Development and preliminary validation of a self-report measure of psychopathic personality. (Doctoral dissertation, University of Minnesota, Minneapolis, 1990).
- Lilienfeld, S. O. (1994). Conceptual problems in the assessment of psychopathy. *Clinical Psychology Review, 14*(1), 17-38.
- Lilienfeld, S. O. & Andrews B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *The journal of personality assessment, 66*(3), 488-524.
- Lynam, D. R., Whiteside, S., & Jones, S. (1999). Self-reported psychopathy: A validation study. *Journal of Personality Assessment, 73*(1), 110-132.
- Lykken, D. T. (1995). *The antisocial personalities*. Hillsdale: Lawrence Erlbau, Associates.
- McCloskey, H. J. (1965). A non-utilitarian approach to punishment, *Inquiry, 8*, p. 239-55.
- McFatter, R. (1982). Purposes of punishment: Effects of utilities of criminal sanctions on perceived appropriateness. *Journal of Applied Psychology, 67*(3), 255-267.
- Millon, T. & Davis, R. D. (1998). Ten subtypes of psychopathy. In T. Millon, E. Simonsen, M. Birket-Smith, R. Davis (Eds.), *Psychopathy: antisocial, criminal, and violent behavior*. (pp. 161-170). New York: The Guilford Press.
- Monahan, J. & Hood, G. L. (1976). Psychologically disordered and criminal offenders: Perceptions of their volition and responsibility. *Criminal Justice and Behavior, 3*, 123-133.
- Murphy, C. & Vess, J. (2003). Subtypes of psychopathy: proposed differences between narcissistic, borderline, sadistic, and antisocial psychopaths. *Psychiatric Quarterly, 74*(1), 11-29.

- Nickel, T. (1974). The attribution of intention as a critical factor in the relation between frustration and aggression. *Journal of Personality, 42*, 482-492.
- Ohbuchi, K. & Kambara, T. (1985). Attacker's intent and awareness of outcome, impression management, and retaliation. *Journal of Experimental Social Psychology, 21*, 321-330.
- Parfit, D. (1984). *Reasons and Persons*, Oxford: Clarendon Press.
- Patrick, C. (1994). Emotions and psychopathy: startling new insights. *Psychophysiology, 31*, 319-330.
- Preacher, K. J., Rucker, D. D., MacCallum, R. C. & Nicewander, W. A. (2005). Use of the extreme groups approach: A critical reexamination and new recommendations. *Psychological Methods, 10*(2), 178-192.
- Price, M., Cosmides, L. & Tooby, J. (2002). Punitive sentiment as an anti-free rider psychological device. *Evolution and Human Behavior, 23*, 203-231.
- Quinsey, V. L. & Cyr, M., Lavalley, Y. (1986). Treatment opportunities in a maximum security psychiatric hospital: A problem survey. *International Journal of Law & Psychiatry, 11*, 179-194.
- Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., Stouthamer-Loeber, M. & Liu, J. (2006). The reactive-proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior, 32* (2), 159-171.
- Rappaport, A. (2003). Rationalizing the Commission: The Philosophical Premises of the U.S. Sentencing Guidelines, *Emory Law Journal, 52*, 557-652.
- 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(6), 673-684.
- Sell, A. (2005). Regulating welfare tradeoff ratios: three tests of an evolutionary computational model of human anger. (Doctoral dissertation, University of California at Santa Barbara, 2005).
- Simon, B., Holzenberg, J. D., & Unger, J. F. (1951). A study of judgment in the psychopathic personality. *Psychiatric quarterly*, 25, 132-150.
- Spohn, C. C. (2002). *How do judges decide?: the search for fairness and justice in punishment*. Thousand Oaks: Sage Publications, Inc.
- Strand, S. & Belfrage, H. (2005). Gender differences in psychopathy in a Swedish offender sample. *Behavioral Sciences and the Law*, 23, 837-850.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46(1), 35-57.
- Turillo, C., Folger, R., Lavelle, J., Umphress, E. & Gee, J. (2002). Is virtue its own reward? Self-sacrificial decisions for the sake of fairness. *Organizational Behavior and Human Decision Processes*, 89, 839-865.
- Vidmar, N. (2000). retribution and revenge. In J. Sanders & V. L. Hamilton (Eds.), *Handbook of justice research in law* (pp. 31-63). New York: Kluwer.
- Vidmar, N. & Crinklaw, L. D. (1974). Attributing responsibility for an accident: A methodological and conceptual critique. *Canadian Journal of Behavioural Science*. 6(2), 112-130.
- Walsh, A. (2000). Evolutionary psychology and the origins of justice. *Justice Quarterly*, 17(4), 841-864.

- Warr, M., Meier, R. & Erickson, M. (1983). Norms, theories of punishment, and publicly preferred penalties for crimes. *The Sociological Quarterly*, 24, 75-91.
- 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(3), 411-419.
- Widiger, T. A. & Lynam, D. R. (1998). Psychopathy and the five-factor model of personality. In T. Millon, E. Simonsen, M. Birket-Smith, R. Davis (Eds.), *Psychopathy: antisocial, criminal, and violent behavior*. (pp. 171-187). New York: The Guilford Press.
- Widom, C. S. (1977). A methodology for studying noninstitutionalized psychopaths. *Journal of Consulting and Clinical Psychology*, 45(4), 674-683.
- Wilson, D. & O'Gorman, R. (2003). Emotions and actions associated with norm-breaking events. *Human Nature*, 14(3), 277-304.

## Footnotes

---

<sup>1</sup> The design of Simon et al.'s study does not allow us to determine whether psychopathic women were incapable of arriving at socially appropriate responses to the open-ended questions or simply uninterested in doing so. However, this question is independent of our broader argument that, whatever explains their reduced tendency to spontaneously produce socially appropriate responses, psychopaths do seem capable of recognizing these behaviors.

<sup>2</sup> Key researchers (Barlow and Durand, 2002; Hare, Hart & Harpur, 1991) have noted the confusion, rife in the field, between the classic "psychopath" and the Antisocial Personality Disorder diagnosis in DSM-IV (4<sup>th</sup> ed., American Psychiatric Association, 1994). Historically, the psychopathy construct emphasized the psychopath's emotional coolness and manipulateness; the DSM-IV refers to a subclass of those psychopaths whose behavior is frankly antisocial and often criminal. The DSM-IV definition may thus exclude "successful psychopaths" who insinuate themselves, often successfully, in general society (see Baird, 2002; Levenson, 1992; Lilienfeld, 1994; Lynam, Whiteside & Jones, 1999; Widom, 1977). The questionnaire we used to screen psychopathic traits surveys the range of classic psychopathic traits rather than the narrowly constrained criteria that form the DSM-IV Antisocial Personality.

<sup>3</sup> Some legal theorists might contend that that our scenario describing a mentally ill offender did not reflect low intent. On this account, although this offender was portrayed as delusional, he fully *intended* to kill the character in his delusion, and so some other factor(s) besides intent must be invoked to explain any resulting reductions in punishment. We agree in principle that this definition of intent might not explain such any reductions sufficiently. Nonetheless, it remains uncertain exactly what other aspects of mental illness *would* provoke reductions in punishment

---

and responsibility. Our aim in this study was not to answer these questions but to vary cues that independently elicit retributive and consequentialist sentiments, so any manipulation of these motives that achieve this should be sufficient. Our manipulation checks attest to the fact that participants did indeed attribute lower intent to the mental illness condition than to its counterpart, validating our usage of the term. Perhaps one reason this manipulation was successful is that lay conceptions of the intentionality of an offense, unlike many legal conceptions, require, not only that the act be deliberate, but that the intentional object be properly identified. This was not the case when our mentally ill offender mistook the bystander for an attacker.

<sup>4</sup> Precautionary reasoning may explain why low-intent offenders were believed to be at greater risk of recidivism regardless of tumor removal; in the absence of total certainty, it is better to be wary of people who have suffered from brain trauma than to be sorry. Nonetheless, our experimental data show that the recidivism manipulation produces the strongest effect on perceived likelihood of recidivism, indicating that participants did believe the expert assessment despite the effect of intent.

<sup>5</sup> Our use of the term “free will” is intended not as a deep psychological or metaphysical contention or presumption, but as a restatement of our participants’ intuitive folk notions of the causes of behavior that appears to arise *sui generis* without an obvious chain of causal precursors.

## Captions

Table 1. Mean ratings of how intentional each offender was perceived, and how likely each offender was believed to recidivate, as a function of the intent and recidivism conditions (range: 1 = not at all; 11 = extremely). Matching letters denote significant comparisons.

Table 2. Mean confinement lengths for high and low psychopathy individuals as a function of the offender's likelihood of recidivism and intent (range: 1 = 0 days; 11 = life). Matching letters denote significant comparisons, except letter pair "c," which reflects a marginal interaction with respect to HPI ratings of intent.

Figure 1. Mean confinement lengths for high and low psychopathy individuals as a function of likelihood of recidivism (range: 1 = 0 days; 11 = life).

Figure 2. Mean confinement lengths for high and low psychopathy individuals as a function of criminal intent (range: 1 = 0 days; 11 = life).

## Appendix

Vignettes

*Low Intentionality, Low Likelihood* - See above.

*Low Intentionality, High Likelihood*

Oren is 56 years old. He has retired from his job as a customer service representative for a small telephone company. Recently, while walking in his neighborhood, he ranted at a passerby and hit the man with a crowbar that he grabbed from a construction site. The victim was killed.

A medical examination after Oren's arrest reveals that he had a growing brain tumor in an area in which tumors consistently lead to violence. Medical experts all agree that the tumor, pressing on a critical brain area, was the cause of the violent incident. Oren had hallucinated that the passerby was a devil who was attacking him, and so he acted in "self defense."

The tumor is inoperable, it cannot be removed. Medication proves ineffective at reducing the violent disposition the tumor creates. A reliable neuropsychological assessment test has determined that Oren is at an extremely high risk for committing more crimes in the future.

*High Intentionality, Low Likelihood*

Oren is 56 years old. He has retired from his job as a customer service representative for a small telephone company. Recently, while walking in his neighborhood, he ranted at a passerby and hit the man with a crowbar that he grabbed from a nearby construction site. The victim was killed.

A police investigation after Oren's arrest reveals that the victim had been a coworker of Oren's at the telephone company. Allegedly, he had been a much better worker than Oren, and Oren blamed him for "showing him up," which blocked a promotion that Oren had been hoping for. After missing the promotion, Oren, previously mild mannered on the job, had become quite angry. He was detected attempting to sabotage the work of the coworker, and finally asked to take a somewhat early retirement.

A reliable psychological assessment test required by the court has determined that Oren is extremely unlikely to commit more crimes in the future.

*High Intentionality, High Likelihood*

Oren is 56 years old. He has retired from his job as a customer service representative for a small telephone company. Recently, while walking in his neighborhood, he ranted at a passerby and hit the man with a crowbar that he grabbed from a nearby construction site. The victim was killed.

A police investigation after Oren's arrest reveals that the victim had been a coworker of Oren's at the telephone company. Allegedly, he had been a much better worker than Oren, and Oren blamed him for "showing him up," which blocked a promotion that Oren had been hoping for. After missing the promotion, Oren, previously mild mannered on the job, had become quite angry. He was detected attempting to sabotage the work of the coworker, and finally asked to take a somewhat early retirement.

A reliable psychological assessment test required by the court has determined that Oren is at an extremely high risk for committing more crimes in the future.

Table 1.

Intent	Likelihood of Recidivism	Psychopathy	Perceived Intent		Perceived Likelihood of Recidivism		N
			Mean	SD	Mean	SD	
Low	Low	LPI	2.78	2.11	3.78	2.44	9
		HPI	2.20	1.64	3.40	2.51	5
		Total	2.57	1.91	3.64	2.37	14
	High	LPI	4.00	2.97	9.00	0.89	6
		HPI	2.00	1.27	8.00	1.10	6
		Total	3.00	2.41	8.50	1.09	12
	Total	LPI	3.27 <sup>c</sup>	2.46	5.87	3.27	15
		HPI	2.09 <sup>c</sup>	1.38	5.91	2.98	11
		Total	2.77 <sup>a</sup>	2.12	5.88 <sup>b</sup>	3.09	26
High	Low	LPI	6.88	3.31	3.50	1.69	8
		HPI	8.60	2.07	2.80	1.64	5
		Total	7.54	2.93	3.23	1.64	13
	High	LPI	8.29	0.95	5.71	2.50	7
		HPI	8.88	1.36	6.38	2.20	8
		Total	8.60	1.18	6.07	2.28	15
	Total	LPI	7.53 <sup>b</sup>	2.53	4.53	2.33	15
		HPI	8.77 <sup>b</sup>	1.59	5.00	2.65	13
		Total	8.11 <sup>a</sup>	2.20	4.75 <sup>b</sup>	2.44	28
Total	Low	LPI	4.71	3.39	3.65	2.06	17
		HPI	5.40	3.81	3.10	2.03	10
		Total	4.96	3.49	3.44 <sup>a</sup>	2.03	27
	High	LPI	6.31	3.01	7.23	2.52	13
		HPI	5.93	3.75	7.07	1.94	14
		Total	6.11	3.36	7.15 <sup>a</sup>	2.20	27
	Total	LPI	5.40	3.28	5.20	2.87	30
		HPI	5.71	3.70	5.42	2.78	24
		Total	5.54	3.44	5.30	2.81	54

Table 2.

Intent	Likelihood of Recidivism	Psychopathy	Confinement Length		N
			Mean	SD	
Low	Low	LPI	6.11	1.90	9
		HPI	5.40	2.88	5
		Total	5.86	2.21	14
	High	LPI	10.00	1.27	6
		HPI	11.00	0.00	6
		Total	10.50	1.00	12
	Total	LPI	7.67 <sup>c</sup>	2.55	15
		HPI	8.45	3.45	11
		Total	8.00 <sup>e</sup>	2.93	26
High	Low	LPI	9.25	1.67	8
		HPI	6.60	0.89	5
		Total	8.23	1.92	13
	High	LPI	10.29	0.76	7
		HPI	10.25	1.04	8
		Total	10.27	0.88	15
	Total	LPI	9.73 <sup>c, d</sup>	1.39	15
		HPI	8.85 <sup>d</sup>	2.08	13
		Total	9.32 <sup>e</sup>	1.77	28
Total	Low	LPI	7.59 <sup>b</sup>	2.37	17
		HPI	6.00 <sup>b</sup>	2.11	10
		Total	7.00 <sup>a</sup>	2.37	27
	High	LPI	10.15	0.99	13
		HPI	10.57	0.85	14
		Total	10.37 <sup>a</sup>	0.93	27
	Total	LPI	8.70	2.28	30
		HPI	8.67	2.73	24
		Total	8.69	2.46	54

Figure 1.

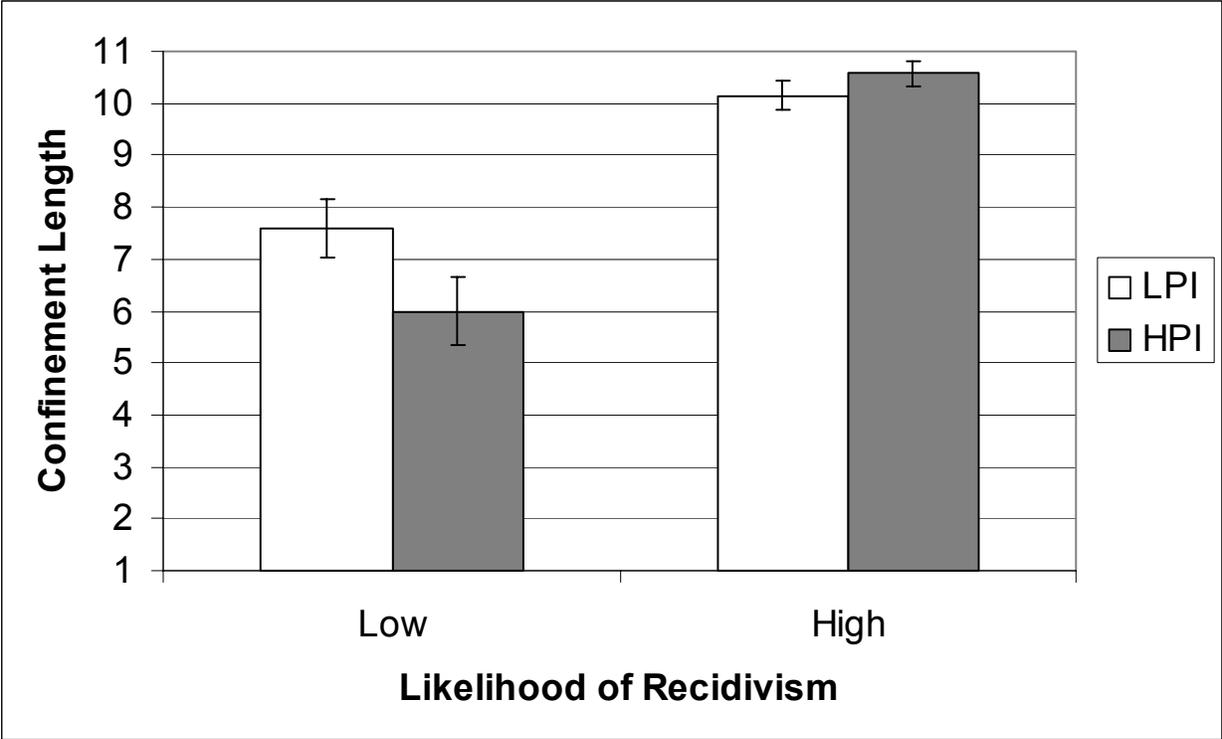


Figure 2.

