



The roles of retribution and utility in determining punishment [☆]

Kevin M. Carlsmith

Department of Psychology, Colgate University, 13 Oak Drive, Hamilton, NY 13346, USA

Received 31 October 2002; revised 1 June 2005

Available online 16 September 2005

Abstract

Three studies examined the motives underlying people's desire to punish. In previous research, participants have read hypothetical criminal scenarios and assigned "fair" sentences to the perpetrators. Systematic manipulations within these scenarios revealed high sensitivity to factors associated with motives of retribution, but low sensitivity to utilitarian motives. This research identifies the types of information that people seek when punishing criminals, and explores how different types of information affect punishments and confidence ratings. Study 1 demonstrated that retribution information is more relevant to punishment than either deterrence or incapacitation information. Study 2 traced the information that people actually seek when punishing others and found a consistent preference for retribution information. Finally, Study 3 confirmed that retribution information increases participant confidence in assigned punishments. The results thus provide converging evidence that people punish primarily on the basis of retribution.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Retribution; Just deserts; Utility; Punishment; Decision making; Justice; Psychology; Law

Punishment is a common and inescapable aspect of daily life; social behavior is often shaped by our use of punishment as well as our reaction to its use. In response to some normatively prohibited behavior, we might verbally castigate, socially ostracize, financially penalize, or physically incapacitate the perpetrator, depending on the circumstances. And to avoid such responses from others, we temper caustic comments in everyday conversations, obey speed limits, pay our taxes, and refrain from stealing a neighbor's new television. In short, punishment is a fundamental and essential component of our social life.

As a result, punishment has often been studied by psychologists. However, this research has focused largely on the recipient of the punishment; psychologists know much less about the person who inflicts the punishment. This paper explores the motives that drive us to inflict

punishment on others, and the psychological justification we use to explain our behavior afterwards.

Philosophical justifications for punishment

Classically, there are two broad justifications for the use of punishment: retribution and utility (Bentham, 1843/1962; Kant, 1790/1952; Vidmar & Miller, 1980).

Retribution

When someone harms society by violating its rules, punishment can be justified by the simple claim that wrongdoing merits punishment, and the perpetrator should suffer in proportion to his or her wrongdoing. The punishment is an end in itself, morally justifiable regardless of its subsequent consequence for either the offender or society. This approach is typically referred to as a "just deserts" or retribution perspective. The key principle of retribution is moral proportionality—a

[☆] The author thanks John Darley for his thoughtful advice, and Robyn LeBoeuf, Benoit Monin, Leif Nelson, Michael Norton, Rebecca Shiner, and Alison Mathias for helpful comments on this research.

E-mail address: kcarlsmith@colgate.edu.

punishment should mirror the degree of moral offense that the perpetrator created. In general, the most important components of an offense are those that determine its moral repugnance. These components include: (a) the magnitude of harm, (b) the perpetrator's intentions, and (c) extenuating circumstances that might mitigate or exacerbate the immorality of the act (see Carlsmith, Darley, & Robinson, 2002; McFatter, 1978; Roberts & Gebotys, 1989). For example, people assign less blame if a murder victim previously molested the murderer sexually (Greene & Darley, 1998). Retribution involves other factors as well, but these three are notable because they help to distinguish retribution from utilitarian punishment. They are the discriminant, if not canonical, components of retribution.

Utility

Another justification for punishment involves its ability to limit future transgressions. The morality of utilitarian punishment stems from the positive consequences it produces; punishment is never an end in itself. A utilitarian analysis encompasses several justifications for punishment, most notably deterrence and incapacitation.

Deterrence

Deterrence theory (see Nagin, 1998, for a review) assumes that a potential wrongdoer is a rational actor, and that punishment changes the costs and benefits of the situation, so that wrongdoing becomes an unattractive option (Bentham, 1843/1962). Most forms of punishment designed to deter people, including fines, incarceration, and corporal punishment, attempt to induce people to complete an a priori analysis that will inhibit them from doing future harm. The focus of the law is not on the degree of potential harm to others, but rather on anticipated gains for the perpetrator. When wrongdoing has a large payoff, the anticipated punishment must be correspondingly high to deter the potential perpetrator, whether the consequences for others are mild or severe. Accordingly, factors that change the expected utility of the behavior are of central importance to deterrence theory, especially if they reveal an insufficient benefit–cost ratio. These factors include: (a) the frequency of wrongdoing (high and increasing rates clearly indicate that the punishment deters insufficiently), (b) detection rates (wrongdoing with low detection or prosecution rates should have correspondingly severe punishments to generate a consistent expected value of punishment), and (c) the publicity of the punishment (deterrence can only be effective if potential wrongdoers know what the punishments are, so highly publicized punishments should be more severe than less publicized ones). Although deterrence is sensitive to other factors (e.g., the culpability of the perpetrator, the speed, and

certainty of punishment), the three factors just described are unique to deterrence.

Incapacitation

Incapacitation theory argues that the cause for wrongdoing lies primarily within the perpetrator (see Zimring & Hawkins, 1995, for a review). Past misbehavior is the best predictor of future wrongdoing, so it is important to somehow incapacitate those who have behaved badly before. In a legal context, this might require incarceration or exile. Important factors related to incapacitation include: (a) the perpetrator's prior record, (b) recidivistic characteristics of the behavior and perpetrator (e.g., employment history, residential stability, credit history, "criminogenic" beliefs, see Tangney, Baumeister, & Boone, 2004), and (c) the impulsivity of the perpetrator (see Darley, Carlsmith, & Robinson, 2000; McFatter, 1978).

Each of these theories—retribution, deterrence, and incapacitation—helps people to derive appropriate punishments, but they do so from different principles. In some cases, the theories diverge and suggest different punishments. In other cases, they converge and suggest the same punishment, but for different reasons. For example, a dishonest stockbroker who stole from pensioners to support a lavish lifestyle might be punished severely by someone seeking retribution, but someone seeking incapacitation might impose a milder punishment, arranging to keep the stockbroker from ever working again in finance (so he or she could not steal again). Someone seeking deterrence would assess whether other stockbrokers are tempted to behave similarly, and if so, would increase punishments until such temptations subside.

Research on punishment

In recent years, researchers have clarified when people will hold an offender accountable and inflict punishment (Alicke, 1992; Goldberg, Lerner, & Tetlock, 1999; Lerner, Goldberg, & Tetlock, 1998; Patterson, Kosson, & Newman, 1987; Robinson & Darley, 1997; Sommers & Ellsworth, 2000). This research has generally focused on the characteristics of situations and perpetrators that lead to greater perceived guilt and punishment. This work is clearly valuable, but it has not identified the reasons *why* people punish.

Few studies have directly addressed the motives that underlie the desire of ordinary people to punish others. Ellsworth and Ross (1983) asked respondents about their attitudes toward capital punishment, and why they favored or opposed the death penalty. They found that respondents tended to endorse any "reason" that was consistent with their position on the issue. Thus, people who supported the death penalty were likely to agree

that it is the criminal's "just deserts," that it is the only sure way to permanently incapacitate a criminal, that it deters other potential criminals, and so on. People who opposed the death penalty also endorsed reasons based on retribution, incapacitation, and deterrence to bolster their position. Overall, people did not seem to discriminate among possible reasons, choosing instead to endorse all of the reasons that supported their positions. Ellsworth and Ross concluded that the justifications for capital punishment do not drive peoples' positions on that issue. Instead, their positions drive the justifications.

Even when respondents identified a single, clear justification for supporting the death penalty, Ellsworth and Ross (1983) found that their reasons were more likely to be post hoc justifications for the attitude, rather than the true cause for the attitude. For example, when people seeking deterrence were asked whether their position on the death penalty would change if the death penalty could be proven to be a weak deterrent, no better than life imprisonment, most of them continued to support the death penalty. Similarly, people who opposed the death penalty said they would not embrace it even if the death penalty could be proven to be a "much more effective" deterrent than life imprisonment. Ultimately, Ellsworth and Gross (1994) concluded that although some insights can be gained by asking people directly about punishment justifications, it is ultimately a "naïve" methodology (p. 26; see also McFatter, 1982; Nisbett & Wilson, 1977).

More recently, Reyna and Weiner (2001) and Weiner, Graham, and Reyna, 1997 have proposed an attributional account of punishment (but also see Goodstadt & Kipnis, 1970; Yukl, Falbe, & Joo, 1993). They find that people punish for different reasons based on their causal beliefs about the cause for a crime. When a crime is thought to reflect stable, internal, and uncontrollable characteristics of a perpetrator, for example, long incapacitative sentences seem appropriate. Unfortunately, Weiner and his colleagues relied on participant self-reports about their reasons for punishments, and like the participants in Ellsworth and Ross (1983), those reasons may have been post hoc justifications for the participants' attitudes, rather than the causal antecedents. For a more complete understanding of punishment, it is important to study not only the justifications that people offer for punishment, but also the more subtle factors that drive actual punishment behaviors.

To address this issue, one might systematically manipulate factors within vignettes that describe wrongdoing and then examine resulting patterns of punishment. This approach, often described as "policy capturing" (Cooksey, 1996), can yield a more accurate picture of people's motives, because people are not asked to introspect about the reasons for their behavior (cf., Nisbett & Wilson, 1977). For example, several researchers (Carlsmith et al., 2002; Darley et al., 2000) who used

this methodology have found that people react strongly to variation in factors associated with retribution, but they largely ignore variation in factors associated with deterrence. There is also a notable lack of correspondence between participants' explicitly stated support for utilitarian justifications and their actual punishment behavior. Apparently, people punish out of a desire to give perpetrators their just deserts, not a desire for future utility.

These results are broadly consistent with previous research that found support for the retributive motive in sentencing. For example, McFatter (1978) found that participants who were instructed to follow a utilitarian sentencing strategy were less sensitive to the severity of a crime than were participants who followed their own intuitions or a retribution strategy. And Roberts and Gebotys (1989) found strong correlations between sentence severity and retribution factors, but no significant correlations between sentence severity and various utilitarian factors.

There is little doubt that retribution information serves as a guide for punishing criminals. However, each study that has reached this conclusion has (logically enough) placed the retribution information squarely in front of the person assigning punishment. This is risky. Although people are responsive to variation in such information when it is clearly presented, they may not typically encounter the information in unconstrained information-gathering environments. In natural environments, where people control the flow of information, they may not routinely seek retribution information. If so, then the apparent importance of such information may be a methodological artifact. Indeed, given that people's explicit justifications for punishment are often more consistent with utilitarian theory (Ellsworth & Mauro, 1998; Nisbett & Wilson, 1977), they may intentionally focus on utilitarian information, if given the option. If an employee is caught stealing, for example, an employer may seek utilitarian information, such as the likelihood that the employee will continue to steal, or whether other employees would benefit from a lesson on the consequences of stealing. If the employer never inquires about the (retributive) factors that make an offender seem more or less blameworthy, then those factors cannot influence his or her punishment decision.

I chose to study how people make punishment decisions by looking at the *type* of information they seek, the *order* in which the information is sought, and the resulting *confidence* that people have in the appropriateness of the assigned punishment. Although these issues arise in many contexts, I continued the tradition of studying them in legal contexts. This allowed me to use previously validated materials and methodologies, and avoided the potential confounds associated with changing multiple variables simultaneously.

General strategy

Different factors determine the appropriate severity of punishment, depending on someone's justification for that punishment. By examining what types of information people seek when they are deciding how to punish an offender, one can infer their justifications for that punishment. This kind of research unfolds in several steps. First, the various factors associated with each justification for punishment must be empirically validated. Second, it is necessary to show that in different contexts, people systematically choose information related to one justification rather than another. Finally, if people do prefer certain types of information, then they should feel more confident about their punishment decisions when they are given the type of information they prefer, rather than non-preferred information.

Based on past research, I predicted that people would identify retribution information as most relevant to criminal sentencing (Study 1), that people would preferentially seek information related to retribution (Study 2), and that people with access to retribution information would be more confident about their punishment decisions than participants who only had access to utilitarian information (Study 3).

Validation of materials

A central assumption of my research is that each justification for punishment (retribution, incapacitation, and deterrence) is associated with unique aspects of an infraction. The logic of this assumption and some supporting data are laid out in Carlsmith et al. (2002). However, that research focused exclusively on deterrence and retribution; the validation results for incapacitation theory were not reported. Those results are included in Table 1, which is expanded from the original table presented in Carlsmith et al. (2002, p. 288).

Carlsmith and his colleagues first derived three uniquely relevant items for each justification from standard philosophy texts, and then tested whether ordinary people could accurately connect the individual items back to the proper justification. To do this, 117 college participants (from the same population as the present research) were presented with a sheet of paper containing short paragraphs describing retribution, deterrence, and incapacitation on the left, and listing nine corresponding items (plus additional filler items) on the right. The paragraphs were taken directly from Darley et al. (2000), and were modifications of those used by McFatter (1978). The participants' task was to draw a line between each paragraph and the three most relevant items. For example, participants might have linked "extenuating circumstance" with retribution, and "publicity of the crime and punishment" with deterrence. All

Table 1
Participants' perceptions of individual items

	Justification		
	Retribution (%)	Incapacitation (%)	Deterrence (%)
Retribution items			
Magnitude of harm	76*	09	16
Perpetrator intent	63*	27	11
Extenuating circumstance	69*	17	14
Incapacitation items			
Likelihood of violence	01	81*	18
Prior record	09	74*	17
Self-control	14	70*	16
Deterrence items			
General frequency	06	25	69*
Detection rate	10	23	68*
Publicity	11	14	76*

* Indicates that a row frequency is significantly different from other row frequencies by χ^2 analysis ($p < .05$, $N = 117$).

items and justifications were written in an abstract manner and did not involve any particular crime.

In general, participant choices matched the rational classifications from philosophy texts (see Table 1). On average, 72% of the choices were "correct," indicating that the participants generally understood how the various items related to each of the justifications. The largest proportion in each row of Table 1 is significantly greater than both of the smaller proportions. Moreover, the correct classification percentages did not vary significantly across the three justifications, $\chi^2(2, N = 117) = 3.4$, $p > .15$. On average, the retribution items were correctly classified 70% of the time, the incapacitation items were correctly classified 75% of the time, and the deterrence items were correctly classified 71% of the time. The most misperceived item was "perpetrator intent." It was properly linked to retribution by 63% of the sample, but to incapacitation by 27% of the sample, and to deterrence by 11% of the sample.

These results suggest that people generally understand how various factors are related to different justifications for punishment, but their understanding is imperfect. People occasionally consider the wrong information when using a particular justification for punishment. Thus, any attempt to cleanly classify people as "pure" retributionists or utilitarians would necessarily be limited by their interpretations of the theories.

Maybe the "mistakes" that people make are not due to their misunderstanding of different justifications for punishment, but rather to overlaps in the justifications themselves. Even though "magnitude of harm" is highly relevant to retribution, for example, it may actually be relevant to incapacitation as well. As a logical matter, this overlap makes little sense. To incorporate both factors simultaneously requires one to jump from one theoretical track to another. As a practical matter, however, it is certainly possible for people to operate this way, and it is

thus important to explore this possibility in future research.

Study 1

Study 1 tested whether people believe that retribution information is more relevant than deterrence and incapacitation information to criminal sentencing. It also tested people's ability to correctly apply different theories of punishment to actual cases. Participants rated the relevance of each type of information shown in Table 1. They did this first using their intuitive or "natural" views, and then an additional three times using the retribution, incapacitation, and deterrence perspectives.

Method

Participants

One hundred and thirty-two Princeton University undergraduates (52% female) participated in a mass testing session. They were recruited through e-mail and flyers, and offered \$8.00 for completing approximately 1 h of surveys assembled by various researchers in the Psychology Department. One participant was excluded from the analyses for not completing all the dependent measures.

Materials

The nine information items evaluated by participants came from the validation study and are presented in Table 1. Each item corresponded to retribution, deterrence, or incapacitation, and in this study, an additional explanatory sentence accompanied each item to clarify its meaning.¹ Participants rated the relevance of each item to sentencing on a scale ranging from 1 *not relevant* to 7 *extremely relevant*. Participants also rated their agreement with each of three brief paragraphs describing different justifications for punishment, again taken from the validation study. These ratings were made on a scale ranging from 1 *not at all* to 7 *very much*. Participants were also asked to select "the one philosophy that best represents your personal view of justice."

Procedure

Participants completed a survey in which they imagined that they were sentencing a convicted criminal. The instructions read: "In any criminal proceeding, there is a collection of facts or 'elements' to the case. As you can imagine, any given case might have hundreds of elements. An important task for a prosecutor then is to figure out which facts a jury would be most interested in knowing before they sentence a defendant. We are interested in your opinion of the following elements. Tell us how

important each element would be if you were on a jury and had to assign a sentence to the guilty perpetrator."

The initial, "default" ratings of the items were followed by a request for participants to indicate their personal endorsement of the three justifications for punishment. They were then asked to sequentially adopt each of the three perspectives and to rate the items for relevance again. Each participant thus rated the items four times. The materials were fully counterbalanced across participants, including the order of the adopted perspectives, and the order of the nine items within each perspective.

At the end of the study, participants were given a printed description of the research and an opportunity to ask questions. There was no deception in this study, so individual debriefings were not conducted.

Results

Order and gender

Order of adopted perspective, order of item within perspective, and participant gender did not affect any of the relevance ratings (all interaction F s < 1.0). These variables will not be discussed further.

Data reduction

To simplify the analyses, ratings of individual items associated with each justification were averaged into a single index, in line with the results of the validation study. Thus, ratings of the three retribution items (harm, intent, and circumstances) were averaged, as were ratings of the three deterrence items (frequency, detection rate, and publicity), and then ratings of the three incapacitation items (violence, prior record, and self-control).

Default ratings

What type of information did people think was most relevant to their sentencing decision? A one-way repeated measures ANOVA on the initial relevance ratings, using the three types of information as the repeated factor, was significant $F(2, 258) = 391.24$, $p < .001$, $\eta^2 = .75$. The results, summarized in the upper part of Table 2, showed that retribution information was

Table 2
Relevance of information by category and perspective

	Mean relevance ratings (<i>SD</i>)		
	Retribution	Incapacitation	Deterrence
Default perspective	5.69 ^a (0.84)	4.97 ^b (1.01)	2.91 ^c (1.09)
Adopted perspectives			
Retribution	5.73 ^a (1.11)	4.09 ^b (1.50)	2.64 ^c (1.22)
Incapacitation	4.72 ^a (1.40)	6.06 ^b (1.07)	3.27 ^c (1.37)
Deterrence	4.17 ^a (1.39)	4.02 ^a (1.68)	4.76 ^b (1.17)

Note. Means with different superscripts are significantly different from other means within the same row using a Bonferroni correction ($p < .05$, $N = 130$).

¹ Complete materials are available from the author.

perceived as most relevant to the sentencing task ($M=5.69$). Incapacitation information ($M=4.97$) was perceived as significantly less relevant ($p<.001$ with a Bonferroni correction), and deterrence information ($M=2.91$) was perceived as significantly less relevant than both of the other information types, $t(129)=659.85, p<.001, t(129)=352.76, p<.001$.²

Adopted perspectives

The ratings people made under the three adopted perspectives provided additional information about how they think about punishment. Such data can reveal whether the participants sufficiently understood the different justifications to use them successfully. For each adopted perspective, the three corresponding items should have been rated as more relevant than the other six items. Thus, ratings of the retribution items should have been highest under the retribution perspective; ratings of the incapacitation items should have been highest under the incapacitation perspective; and ratings of the deterrence items should have been highest under the deterrence perspective. The data can also reveal which of the three justifications produced ratings that most closely resembled the participants' default ratings, and thus best represented their "natural" views.

A two-way repeated measures ANOVA showed a significant main effect for Adopted Perspective, $F(2,258)=20.18, p<.001$, and for Item Category, $F(2,258)=128.44, p<.001$. These effects, however, were qualified by the two-way interaction between the factors, $F(4,516)=143.21, p<.001$. The relevance ratings for a given item category were highest when the participant had adopted the relevant perspective. The results in the lower part of Table 2 addressed participants' understanding of the different justifications for punishment. When asked to adopt a retribution perspective, participants rated the retribution items as more relevant than the other types of items ($M=5.73$ vs. 4.09, 2.64). A one-way repeated measures ANOVA was significant, $F(2,260)=278.16, p<.001$, and follow-up contrasts with a Bonferroni correction, indicated that the means for retribution, incapacitation, and deterrence were all different from each other. Similarly, when asked to adopt

an incapacitation perspective, participants rated the incapacitation items as more relevant than the other types of items ($M=6.06$ vs. 4.72, 3.27). A one-way repeated measures ANOVA was significant, $F(2,258)=162.76, p<.001$, and follow-up contrasts with a Bonferroni correction, again indicated that all three means were different from each other. Finally, when asked to adopt a deterrence perspective, participants rated the deterrence items as more relevant than the other types of items ($M=4.76$ vs. 4.17, 4.02). A one-way repeated measures ANOVA was significant, $F(2,260)=12.95, p<.001$, and follow-up contrasts with a Bonferroni correction, indicated that the deterrence items were rated as more relevant than the other two sets of items, which were not rated differently from each other.

Note, however, that there was some intrusion of retribution into ratings of incapacitation and deterrence items. The relevance of retribution items never dropped below the midpoint (4.0) of the rating scale, suggesting that participants could focus on incapacitation or deterrence information when necessary, but they were unwilling to disregard retribution information.

Table 2 also suggests that the default perspective was retribution. Relevance ratings for the retribution items did not differ significantly by paired t test when participants were instructed to adopt that perspective ($M=5.73$) and when they used their own judgment ($M=5.69$), $t(129)=.33, ns$. In the default condition, however, participants did give slightly more weight to incapacitation information ($M=4.97$ vs. 4.09), $t(130)=8.35, p<.001$, and to deterrence information ($M=2.91$ vs. 2.64), $t(129)=2.93, p<.01$. The similarity between the ratings made under the default and retribution perspectives contrasts with marked differences between the default ratings and ratings made from the other two perspectives. The relevance ratings in the default perspective for retribution, incapacitation, and deterrence items were significantly different by paired t test from every corresponding rating in both the incapacitation and deterrence perspectives (all $ps<.01$). That is, for retribution items, the default perspective rating (5.69) is different from the incapacitation perspective (4.72) and the deterrence perspective (4.17). The same pattern holds for incapacitation items (4.97 vs. 6.06, 4.02) and deterrence items (2.91 vs. 3.24, 4.76). These analyses indicate that the default relevance ratings were quite similar to ratings made from the adopted retribution perspective, but dissimilar to ratings made from the adopted incapacitation and deterrence perspectives.

The analyses above are supported in both a relative and an absolute sense. I compared each of the mean relevance ratings in Table 2 to the neutral scale midpoint of 4.0, using one-sample t tests. In the default condition, participants rated the retribution and incapacitation items significantly above the midpoint, $t(129)=22.93, p<.001$ and $t(130)=11.12, p<.001$, and they rated the

² Some readers might be troubled by the inclusion of the item "magnitude of harm" to the retribution category, because it is too fundamental to the decision-making process to be fairly included in any one category. As noted by Carlsmith et al. (2002), however, the seriousness of the offense is integral to a theory of retribution and does not have a place within the utilitarian positions described by Bentham (1843/1962) and others (Zimring & Hawkins, 1995). When this item was removed from the retribution index, the composite remained significantly higher than both the deterrence, $t(129)=21.0, p<.001$, and incapacitation indexes, $t(129)=3.8, p<.001$. Each of the subsequent analyses was conducted with "magnitude of harm" removed from the retribution index and the results remained significant and the interpretation unchanged.

deterrence items significantly below the midpoint, $t(129) = -11.43, p < .001$. Among the adopted perspectives, the relevance ratings along the diagonal (e.g., retribution items within retribution perspective) were all significantly above the scale midpoint (all $t_s > 7.0$, all $p_s < .001$). In contrast, the off-diagonal ratings were not significantly above the scale midpoint (all $t_s < 1.4$, all $p_s > .16$). The only exception involved retribution. Participants rated the retribution items higher (4.72) than the scale midpoint when they adopted the incapacitation perspective, $t(129) = 5.90, p < .001$.

The similarity in ratings between the default condition and the adopted retribution condition was also apparent in correlational analyses. For example, the correlation between ratings of the retribution items made in the default and adopted retribution conditions was .43. For the incapacitation items, the correlation between these conditions was .59, and for the deterrence items, it was .49 (all $p_s < .001, N = 130$). In contrast, the correlations between ratings made in the default condition and the adopted incapacitation condition were notably smaller: .21 for retribution items, .25 for incapacitation items, and .26 for deterrence items (all $p_s < .02, N = 130$). Likewise, the correlations between ratings made in the default and adopted deterrence conditions were small: .09 for retribution items, .23 for incapacitation items, and .24 for deterrence items. Only the latter two correlations were significant ($p < .01, N = 130$). The default-retribution correlations were tested against the default-incapacitation and default-deterrence correlations in six separate z tests, after conducting Fischer's r to z transformations. In all cases, the comparisons were significant ($p < .05$). This pattern of correlations suggests that asking participants to adopt a retribution perspective produces responses quite similar to those they would make on their own, but asking participants to adopt either an incapacitation or deterrence perspective produces responses that are quite different.

Individual differences

In general, participants endorsed all three justifications for punishment and did not strongly differentiate among them. The mean endorsement ratings were 5.0 for retribution, 4.6 for incapacitation, and 4.3 for deterrence. A one-way repeated measures ANOVA yielded $F(2, 260) = 4.13, p = .02$, and post hoc tests with a Bonferroni correction indicated that deterrence was rated significantly lower than the other two punishment justifications, which did not differ from each other. These results replicate previous research indicating that at an abstract level, people are happy with all three justifications and do not feel a strong need to choose among them (Carlsmith et al., 2002; Darley et al., 2000; Ellsworth & Gross, 1994; Ellsworth & Ross, 1983).

When forced to choose among the three perspectives, however, 48% of the participants said that retribution best reflected their own views; 26% chose deterrence, and 26% chose incapacitation. These percentages are significantly different from what would be expected by chance, $\chi^2(2, N = 128) = 11.92, p < .003$. It is interesting to note, however, that participants' choices were not related to any of the other dependent measures. Indeed, the default relevance ratings did not vary as a function of the participants' self-described orientations. This was tested with a mixed-model 3×3 ANOVA in which the retribution, incapacitation, and deterrence ratings were a within-subjects factor and participant orientation was a between-subjects factor. The two-way interaction was not significant ($F = .24$), nor was the main effect of participant orientation ($F = .08$). This illustrates the danger of relying on attitudinal measures when studying punishment.

Discussion

The results of Study 1 confirmed that information related to a retribution justification for punishment is most relevant to people who are sentencing the perpetrators of crimes. Incapacitation information was perceived as relevant, but less so than retribution information. Deterrence information was not perceived as very relevant at all.

It is important to note, however, that the acquisition of information by participants in this study had no associated "cost." Participants did not have to read lengthy crime reports or pour through reams of statistical data to gather the different types of information. It was thus strategic for them to consider all of the available information and then discard any information that seemed irrelevant later on. This probably created a low threshold for concluding erroneously that information *is* useful, and a very high threshold for concluding erroneously that information is *not* useful. With this in mind, the relatively high default ratings in Table 2 make more sense. Participants may have been wary of labeling any information as irrelevant when: (a) it was described only in abstract terms, (b) the task seemed important, and (c) the cost of information was negligible.

Another problem is that the significant differences among retribution, incapacitation, and deterrence ratings were relative differences that might not replicate in a between-subjects design. Participants evaluated the relevance of each type of information in comparison to the other types of information, rather than on their absolute merits. If participants had evaluated deterrence items in isolation, for example, then they might have given them higher relevance ratings.

Finally, I did not investigate which information items people looked at first. Given the importance of first impressions and the construction of explanatory narratives in decision-making (Pennington & Hastie, 1993), it

is useful to know *when* information is processed. People tend to limit the amount of information they seek, and they often stop gathering information once they know enough to make a decision (i.e., people are “cognitive misers”). Even if a particular item of information has some relevance, people might thus bypass it in their pursuit of expediency.

Study 2

Study 1 showed that people see retribution information as the most relevant type of information for determining punishments. From this it follows that people should preferentially seek retribution information before assigning punishments. Study 2 assessed the order in which participants recruited different types of information before sentencing a criminal, and tested the hypothesis that they would seek retribution information before incapacitation or deterrence information. Study 2 also tested whether the acquisition of retribution information increased participants’ confidence in the appropriateness of their sentence assignments more than did the acquisition of incapacitation or deterrence information. This kind of research—described as behavioral process tracing—has been widely used in the areas of social judgment and decision-making (Jacoby, Jaccard, Kuss, Troutman, & Mazursky, 1987). Typically, participants are confronted with a problem and then given access to a database that contains information relevant for solving that problem. They can acquire as much or as little of that information as they wish. The depth, content, and sequence of their information seeking can then be assessed.

Method

Participants

Forty-two Princeton University undergraduates (60% female) participated in this study as part of a course requirement.

Procedure

Working alone, participants completed a “judicial decision-making” experiment using laboratory computers. The experimenter began by stating that the purpose of the study was to learn about the cognitive processes underlying jury decision-making. Participants learned that their task was to assign prison sentences to hypothetical criminals who were known to be guilty of their crimes. All that participants initially knew was that a crime had occurred. To assign a sentence, they first had to assemble facts about the case, one fact at a time. To do this, they viewed a request screen on the computer that contained nine information items, and then selected the type of information that would be most useful for sentencing the criminal

in one particular case. The computer then provided the requested information (including a cumulative record of any previously requested information). Participants then sentenced the perpetrator (using a scale ranging from 1 *not at all severe* to 7 *extremely severe*), and made a rating of their confidence in the sentence (using a scale ranging from 1 *not at all confident* to 7 *extremely confident*). This procedure was repeated five times. The order in which the information categories were listed on the screen remained constant for a given individual, but was counterbalanced across participants. At the end, participants were debriefed about the purpose of the study. No one expressed any suspicions or could correctly guess the research hypothesis.

Materials

The specific information shown to participants can be found in Appendix A. Each information category contained items of both low severity and high severity (low severity versions are always listed first in Appendix A). For example, the three retribution items (magnitude of harm, intent, and extenuating circumstance) had low severity versions that should generate less moral outrage, and high severity versions that should generate more moral outrage. All participants were randomly assigned to receive either the high or low severity items for each type of information (retribution, incapacitation, and deterrence). Thus participants could have received three high severity retribution items, three high severity incapacitation items, and three low severity deterrence items (HHL), or some other combination. In all cases, participants only interacted with a single criminal case. The primary purpose of this manipulation was to replicate previous studies that showed a heightened sensitivity to variation in retribution items compared to incapacitation and deterrence items (Carlsmith et al., 2002; Darley et al., 2000).

Results

Data reduction

As in Study 1, the individual items requested by participants were grouped by category for statistical analysis.

Information choices

During the first trial, participants could select any of the nine information items. On each subsequent trial, that number necessarily decreased by one. Because each information category (retribution, incapacitation, and deterrence) contained only three items, even the most determined “retributionist” was thus forced to select items from one of the other categories on the fourth and fifth trials.

Table 3 shows that participants preferred the retribution category. This trend was particularly clear during

Table 3
Type of information chosen on each trial

	Type of information chosen		
	Retribution (%)	Incapacitation (%)	Deterrence (%)
Trial 1	97	03	00
Trial 2	64	36	00
Trial 3	57	36	07
Trial 4	29	60	12
Trial 5	15	56	28

the first three trials, before a participant could exhaust a given category. After three trials, for example, 31% of the sample had chosen retribution items exclusively, and 57% of the information they chose on that trial was related to retribution. By contrast, 36% of the information chosen on that trial was related to incapacitation and 7% was related to deterrence. After five trials, when 41% of participants had already selected all of the available retribution information, participants turned to incapacitation rather than deterrence information. These data were further analyzed using two non-parametric approaches.

Portfolio analysis

After each trial, I examined the total portfolio of information amassed by each participant and then tested whether it was consistent with a model of chance selection. At Trial 1, for example, each participant had a portfolio containing one type of information (retribution, incapacitation, or deterrence). By chance, each type of information should have been chosen equally often across participants. By the end of the experiment, each participant had a portfolio containing five information choices, and again it was possible to compare the patterns of choices that actually occurred with what should

have occurred by chance. To simplify these analyses, incapacitation and deterrence information were combined into a single “utility” category.

Table 4 shows each of the observed and expected portfolios for each trial, omitting (for simplicity and better statistical reliability) any portfolios chosen by three or fewer persons. A strong pattern of results emerged. Consider, for example, the portfolios for participants on Trial 3. There were eight possible portfolios on that trial, only three of which were selected by more than three participants. Thirty-one percent of the participants had portfolios that contained *only* retribution information (RRR), and all the rest had portfolios dominated by retribution information (RRU, RUR). The binomial probability of the RRR portfolio being so popular is less than .001 (chance predicts 12.5%). Clearly, retribution information was selected often and at levels quite different from chance.

Table 4 also shows that participants chose retribution information early, as well as often. On Trial 1, their first opportunity to get information, 97% chose retribution. On Trial 2, 64% chose retribution, and on Trial 3, 54% chose retribution. Only 20% chose retribution on Trial 4, but that is partly because 31% had already “closed out” the category. So among the 69% of the participants that remained, nearly a third chose retribution. Nobody chose retribution information on Trial 5, but that is partly because 41% had already “closed out” the category.

Analysis of ranks

My second analytical approach involved giving each participant a rank-preference score for each of the nine items. If a given item were chosen first, for example, then it received 5 points. If it were chosen second, then it

Table 4
Portfolio analysis of chosen information

	Information type	Observed frequency	Expected frequency	Binomial probability	
		Raw (%)	Raw (%)	<i>z</i>	<i>p</i>
Trial 1	R	38 (97)	20 (50)	5.76	.001
Trial 2	RR	25 (64)	10 (25)	5.55	.001
	RU	13 (33)	10 (25)	1.11	.267
Trial 3	RRU	13 (33)	5 (12.5)	3.87	.001
	RRR	12 (31)	5 (12.5)	3.39	.001
	RUR	9 (23)	5 (12.5)	1.94	.052
Trial 4	RRRU	12 (31)	3 (7.1)	5.47	.001
	RRUU	9 (23)	3 (7.1)	3.65	.001
	RRUR	4 (10)	3 (7.1)	0.61	.542
	RURR	4 (10)	3 (7.1)	0.61	.542
	RURU	5 (13)	3 (7.1)	1.22	.222
Trial 5	RRRUU	12 (31)	2 (4.5)	7.72	.001
	RRUUU	7 (18)	2 (4.5)	3.86	.001
	RRURU	4 (10)	2 (4.5)	1.54	.124

Note. *N* = 39. Portfolios with three or fewer observations are omitted. R = retribution information, U = utilitarian information (incapacitation + deterrence). Order within each portfolio represents order in which participant chose items.

received 4 points, and so on until just 1 point if the item were chosen last. Non-selected items received 0 points. These item scores were summed for each participant within each information category (retribution, incapacitation, and deterrence), to produce three category rank-preference scores. For example, consider a person who chose retribution information on the first and second trials and incapacitation information on the remaining trials. That person received a total retribution rank-preference score of 9 (5 + 4), a total incapacitation rank-preference score of 6 (3 + 2 + 1), and a total deterrence rank-preference score of 0.

The mean rank-preference score was 9.88 ($SD = 1.97$) for retribution information, 4.33 ($SD = 1.27$) for incapacitation information, and 0.71 ($SD = 1.83$) for deterrence information. An omnibus repeated-measures analysis using Friedman's test of ranks was significant, $\chi^2(2, 42) = 70.1$, $p < .001$, indicating that the order in which people chose different categories of information was not random. Follow-up pairwise comparisons using the Wilcoxon signed ranks test were all significant ($p < .001$). These analyses clearly showed that retribution information was sought first, followed by incapacitation information. Deterrence information was a distant third.

Confidence about sentences

Across participants and trials, how much incremental confidence did each type of information generate? To answer this question, I calculated the average change in reported confidence after receiving each type of information. For someone who selected retribution information on the first three trials and incapacitation information on the remaining trials, there were thus two confidence-change scores, one for retribution (corresponding to the average change from Trial 1 to Trial 2, and from Trial 2 to Trial 3), and one for incapacitation (corresponding to the average change from Trial 3 to Trial 4, and from Trial 4 to Trial 5).

All three types of information improved confidence in sentences—each mean was at least marginally greater than zero by one-sample t tests: $M_{jd} = .54$, $t(39) = 3.41$, $p = .001$; $M_{incap} = .17$, $t(41) = 1.86$, $p = .07$, $M_{det} = .28$, $t(14) = 1.89$, $p = .08$. However, the effects of different types of information on confidence were not equal. Planned contrasts revealed that retribution information improved confidence more than did incapacitation information, $t(39) = 2.22$, $p = .03$. Neither pairwise comparison involving deterrence information was significant, despite the relatively large mean differences. The low N (13) for the latter analyses reflects the fact that few participants requested deterrence information on any trial. This probably contributed to the lack of statistical significance.

This analysis could not disentangle the effects of information type from the effects of information order. For instance, it is unclear whether the confidence boost

associated with retribution was due to the presence of retribution information in a portfolio, or to the fact that retribution information was selected earlier than other types of information. In this experiment, information order and type were not orthogonal because participants were necessarily allowed to determine the order in which information entered their portfolios. This issue was explored in Study 3.

Discussion

The results of Study 2 reinforced those of Study 1—when people sentence criminals, they do so from a retributive rather than utilitarian stance. Study 2 showed that participants requested information related to retribution sooner and more frequently than utilitarian information. If people feel that retribution information is especially relevant for sentencing, then acquiring such information should also make them more confident about their sentences. This hypothesis was confirmed—confidence increased more from one trial to another when retribution information was requested than when either type of utilitarian information was requested.

It should be noted that each category contained only three items of information, and that participants were “sampling without replacement.” This design had both strengths and weaknesses. Obviously, I could not get an accurate description of participants' true preferences after the third trial, because at that point it was possible for them to have exhausted a given category. It would have been useful, perhaps, to have “deeper” categories. However, by limiting participants' choices on the fourth and fifth trials, I gained insight into their secondary preferences.

It may seem surprising that different types of information had different effects on confidence. Each participant was free to choose the information he or she received, so one might expect uniformly high confidence ratings. Presumably, people can recognize the types of information that are most useful for the task at hand, and one would rationally expect them to seek out that information. Yet that was not the case. Retribution information provided the greatest boost in confidence—nearly half a point on a seven-point scale for each retribution item. One implication of this result is that people may not have chosen the correct (most useful) information for the task they were given. Participants who chose utilitarian information actually made poor choices, as indicated by their subsequent confidence ratings.

Study 3

The purpose of Study 3 was to confirm that the acquisition of retribution information—in contrast to utilitarian information—increases confidence in sentencing decisions. This study was a conceptual replication of

Study 2, but the order in which participants received different pieces of information was controlled, thereby disentangling the effects of information order and type.

Method

Participants

Thirty-five Princeton University undergraduates (57% female) took part in a 15-min study on judicial decision-making. They were recruited through e-mail and participated remotely using an interactive web page. A small monetary incentive (\$3.00) was offered for participation.

Materials

The materials and procedures were generally the same as in Study 2. Deviations are noted below. To simplify the procedure and analyses, the “high severity” version of each item from Appendix A was used in all conditions. The punishment severity ratings and the confidence ratings were identical to those in Study 2, but measured on 15-point (1–15) scales to minimize potential floor and ceiling effects.

Procedure

As before, participants learned that a crime had been committed, and that they would have an opportunity to learn more about it. Afterwards, they could give the perpetrator whatever sentence seemed fair. To create a more reliable measure of confidence change, participants completed nine trials rather than five.

Participants viewed all nine pieces of information, one per trial, blocked according to information type (retribution, deterrence, and incapacitation). The order of these blocks, and the order of the items within each block, was counterbalanced across participants. The resulting 3×3 between-subjects design ensured that order of presentation was balanced across participants. Thus, the impact of order was “spread out” across participants, so that each information item occurred in all possible positions, and equally across positions, for the sample. As a result, any effect of information type was independent of order.

At the end of the study, participants were given an opportunity to express their thoughts about the study, debriefed about its purpose, and encouraged to contact the researcher for additional information.

Results and discussion

My primary hypothesis was again that participants who received information relevant to retribution would feel more confident about their sentencing decisions. As before, I calculated a confidence-change score for each participant, for each type of information. All three types of information again improved confidence in sen-

tences. Retribution information yielded the largest confidence boost; a one-sample t test showed that it was significantly greater than zero ($M = 1.54$, $SD = 2.33$), $t(34) = 3.92$, $p = .001$. The increase in confidence from incapacitation information was marginally greater than zero ($M = .63$, $SD = 1.93$), $t(34) = 1.90$, $p = .07$, and deterrence information produced a non-significant increase in confidence ($M = .48$, $SD = 1.96$), $t(34) = 1.40$, $p = .17$.

A one-way repeated measures ANOVA, with information type as the within-subjects factor, was significant, $F(2, 68) = 4.04$, $p = .02$. Follow-up contrasts revealed that the increased confidence associated with retribution information was significantly higher than the increases associated with either incapacitation ($p = .03$) or deterrence ($p = .02$) information. The latter two increases were not significantly different from each other.

The results of this study replicated those of Study 2, but removed the confound created by allowing the participants to select the order of information. In Study 2, it was unclear whether the increased confidence associated with retribution stemmed from the retribution information itself, or from the fact that people chose to receive retribution information earlier. In Study 3, it was unambiguous that retribution information increased participant confidence ratings, whenever it was received.

General discussion

My research shows that when people punish criminals, their sentencing decisions are affected primarily by retribution, rather than by incapacitation or deterrence. Although people say they value utilitarian goals, when it comes to actually seeking information and assigning sentences, their behavior reveals that they care most about retribution.

The classification of information into retribution, incapacitation, or deterrence categories was derived from the writings of moral philosophers, but received empirical support in the initial validation study. Study 1 demonstrated that people generally view retribution information to be more relevant to sentencing than utilitarian information. It also demonstrated that people could link different information items with the philosophies from which they were derived. When asked to adopt each perspective, participants could do so appropriately, albeit with some error. Finally, comparisons among the participants' responses revealed that their natural (default) approach to sentencing probably involved retribution.

Study 2 showed that when participants sought information about crimes in order to sentence perpetrators, they searched most often for retribution information. Incapacitation information was sought less often and only later in the process, and deterrence information was rarely sought. Consequently, the impact of

utilitarian information is probably weaker when people control the flow of information than when it is controlled by others (e.g., researchers). Study 2 also showed that people feel more confident in the accuracy of their sentences when they have gathered retribution rather than incapacitation or deterrence information. These results, however, were qualified by the fact that participants sought retribution information earlier than the other types.

Study 3 confirmed that retribution information makes people more confident about their sentences than does utilitarian information, regardless of the order in which information is received. This provided additional, converging evidence that despite their favorable attitudes toward utilitarian information, people depend primarily on retribution information for sentencing.

Limitations

Before exploring the implications of these results, it is important to consider some limits on their generalizability. The participants that I studied tended to be similar in age, race, and political outlook. This may be a substantial liability; my results should be replicated using broader samples. However, prior research (see Hamilton & Rytina, 1980; Rossi, Berk, & Campbell, 1997; Rossi, Waite, Bose, & Berk, 1974; Tyler & Weber, 1983), suggests that demographic variables and personal experience (e.g., victim status, perceived risk of crime) do not increase support for utilitarian punishments. Thus, it is unlikely that the strong support I found for retributive punishment can be attributed solely to sample limitations.

Generalizing from a limited set of operationalized variables is always problematic. One can easily generate examples of crimes (and related factors) not used in my studies. Study 2, for example, involved only one vignette (partitioned into nine “pieces” of information), and thus is hardly representative of all possible cases in which crimes must be punished. Yet the design of my studies does permit some generalization. In the first two studies, participants responded to general categories of information, rather than specific instantiations of those categories. Thus the proper question about generalization is whether the nine information items I offered are a representative sample of the types of information that a person might seek. So even though my studies did not utilize an exhaustive set of scenarios, they did side-step the issue of idiosyncratic stimuli by focusing on general stimulus categories for testing the hypotheses.

Most critically, my research focused entirely on reactions to hypothetical crimes and criminals, and the judgments made by participants had no real consequences. Social psychologists have repeatedly demonstrated that there is a substantial gulf between

predicted behavior and actual behavior. Thus, one should be cautious about generalizing these results to different types of punishment situations—particularly personal interactions—without further research. But my approach is not without merit, because there are many real situations in which decisions are made using similar procedures. For example, the “punishers” in hierarchical organizations (e.g., manager, principals, and legislators) often make or enforce policies in the abstract, without necessarily coming into contact with perpetrators or victims. Indeed, some believe that making such decisions using cool logic, rather than hot emotions, is desirable. My methods probably speak to such situations better than research featuring real interactions and opportunities to retaliate. Such research, of course, should still be carried out at some point, because punishment is often a spontaneous act.

Several problems were not fully addressed by my studies. My starting point was the answers moral philosophers have offered to the question “why do we punish?”. This approach may have restricted the kinds of questions, and therefore the kinds of answers, open for investigation. For example, there are more motives for punishment than were tested in my studies. And there is no reason to suppose that these motives are necessarily discrete; people might punish others for multiple reasons that reflect different views of punishment. If so, then I limited unnecessarily the range of independent and dependent variables that I studied, which limits the validity and generalizability of my findings.

What this criticism suggests for the future is an entirely empirical approach to studying justifications for punishment. One might begin with open-ended questions about why people punish, along with a careful review of all the factors that previous research has identified as relevant to attributions of responsibility and punitive intent. Armed with the “universe” of possible motives, one could then begin a systematic analysis of which factors are relevant and which factors cluster together. Such an approach might reveal new “amalgamated” theories of punishment not described by moral philosophers (e.g., an incapacitation approach that takes moral outrage into account), or it might reveal distinct clusters that do, in fact, map neatly onto the major theories I have already identified.

Connections with other research

The results of my studies are broadly compatible with the existing literature in this area (Finkel, Maloney, Valbuena, & Groscup, 1996; Hogarth, 1974; Rossi et al., 1997), and dovetail with the findings of McFatter (1978, 1982), Darley et al. (2000), and Carlsmith et al. (2002). Nonetheless, there are also studies that find support for more utilitarian motives in punishment. Most notably,

Goldberg et al. (1999) and Rucker, Polifroni, and Tetlock (2004) have developed an “intuitive prosecutor” metaphor that describes individuals as societal watchdogs who are concerned about the breakdown of civil society, and who punish from a utilitarian desire to maintain societal order. Their results show an important moderating role for offense severity, and reveal that the retribution motive dominates as offense severity increases. These researchers might argue that although utilitarian motives are the ultimate cause for people’s punitiveness, retribution is often the proximate cause.

There is also evidence that people differ meaningfully in their support for different punishment justifications (Darley et al., 2000). Moreover, ordinary people and legislative bodies have justified punishments in radically different ways throughout history. This suggests that justifications for punishment are a learned response, rather than an innate human characteristic. This contradicts the strong claim that there is a universal belief in a just world (Lerner & Miller, 1978). The Just World Hypothesis (Lerner & Simmons, 1966) is largely consistent with a retributive model of punishment: people get what they deserve and deserve what they get. Punishments are, by definition, thus proportional to the offenses. Although my results are generally consistent with the Just World Hypothesis, I would not necessarily predict similar results in cultures that have developed more utilitarian justifications for punishment.

Future directions

The psychology of retribution is complex. My research represents some small steps towards a more comprehensive theory of punishment. Although retribution seems to be a powerful motive, there must certainly be situations where utilitarian motives are important to people. Identifying these situations is the next logical step in this research area. For example, certain offenses (perhaps those that generate irreparable harm) may evoke more intense deterrent motives. Likewise, such perpetrator characteristics as age and openness to rehabilitation may change the perceived effectiveness of various utilitarian responses, thus making some of them more attractive. To this end, work that seeks moderating and mediating variables that increase the influence of retributive and utilitarian factors is a logical next step.

The most promising avenue for future research on these motives comes from the very foundation of the different philosophical justifications for punishment. Retribution focuses on the completed act, and does not concern itself with the likely consequences of the punishment. By contrast, utility expressly and exclusively focuses on the future benefits that can result from punishment. It seems logical, then, to expect that a person’s “temporal orientation”—a focusing on the future or the past—might activate either a future-based or past-based theory of punishment. If true,

retribution may be the default approach to punishment only when punishment is assigned after the deed is done. The default approach might be more utilitarian, however, when the punishment is determined in advance, before any harm has occurred. In the latter situation, one might expect critical factors relevant to the retribution approach, such as mitigating circumstances or intent, to be relatively unimportant. If someone is trying to prevent future harm, it makes little sense to allow “loopholes” that would excuse or permit the offending behavior.

Finally, the nature of the relationship between the perpetrator and the punisher might influence the purpose of the punishment. In my research so far, this relationship has been abstract, impersonal, and ephemeral. It is plausible, and perhaps probable, that punishment among friends, lovers, or family members is driven by a different set of psychological processes.

Appendix A

A.1. Magnitude of harm (retribution)

- The perpetrator managed a middle sized independently owned store near where he lived, and had responsibility for accounting for money at the close of business each day. He has been prosecuted for embezzling \$5600 over an 8 month period.
- The perpetrator managed a chain of dry cleaning plants that used quite toxic cancer-causing chemicals. He is being prosecuted for dumping the toxic chemicals into fields on the edge of town over an 8 month period.

A.2. Perpetrator intent (retribution)

- Based on the records kept by the perpetrator, it appears he did not intentionally break the law.
- Based on the records kept by the perpetrator, it appears that he intentionally broke the law.

A.3. Extenuating circumstance (retribution)

- The perpetrator was attempting to ensure that funds would be available for the annual year-end party the company traditionally throws for the patients of the local children’s hospital.
- The perpetrator was attempting to ensure that funds would be available to support his habit of going out to bars and placing bets on sporting events.

A.4. Likelihood of violence (incapacitation)

- There is little evidence that the perpetrator would commit violent acts in the future. Typically, people who commit this sort of crime do not later become violent offenders.

- There is evidence that the perpetrator would commit violent acts in the future. Typically, people who commit this sort of crime later become violent offenders.

A.5. *Prior record (incapacitation)*

- A standard review of court records reveals that the perpetrator has a clean history and has never before been involved in any criminal activity of this sort.
- A standard review of court records reveals that the perpetrator has a criminal history and has been involved in numerous activities of this sort.

A.6. *Self-control (incapacitation)*

- The perpetrator is not an impulsive individual. He is described as even-tempered and disciplined in his behavior.
- The perpetrator tends to be an impulsive individual. He is described as wild-tempered and undisciplined in his behavior.

A.7. *General frequency (deterrence)*

- Crimes of this sort are rarely committed, and have been declining over the past decade.
- Crimes of this sort are frequently committed, and have been increasing over the past decade.

A.8. *Detection rate (deterrence)*

- Although most people are not aware of it, crimes of this sort are eventually very easy to detect, and prosecutors have a very high conviction rate in this area.
- Although most people are not aware of it, crimes of this sort are actually very hard to detect, and prosecutors have a very low conviction rate in this area.

A.9. *Publicity (deterrence)*

- The crime received little attention when it occurred, and the trial received no media coverage at all. It is unlikely whether the proceedings and outcome will become known to anyone except those directly involved.
- The crime received a great deal of attention when it occurred, and the trial received extensive media exposure throughout. As a result, many people in the community will be aware of the proceedings and outcome.

References

Alicke, M. D. (1992). Culpable causation. *Journal of Personality and Social Psychology*, 63, 368–378.

- Bentham, J. (1962). Principles of penal law. In J. Bowring (Ed.), *The works of Jeremy Bentham* (p. 396). Edinburgh: W. Tait (original work published 1843).
- 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.
- Cooksey, R. W. (1996). *Judgment analysis: Theory, methods, and applications*. San Diego, CA: Academic.
- 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.
- Ellsworth, P. C., & Gross, S. R. (1994). Hardening of the attitudes: Americans' views on the death penalty. *Journal of Social Issues*, 50, 19–52.
- Ellsworth, P. C., & Mauro, R. (1998). Psychology and law. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 684–732). New York: McGraw-Hill.
- Ellsworth, P. C., & Ross, L. (1983). Public opinion and capital punishment: A close examination of the views of abolitionists and retentionists. *Crime and Delinquency*, 29, 116–169.
- Finkel, N. J., Maloney, S. T., Valbuena, M. Z., & Groscup, J. (1996). Recidivism, proportionalism, and individualized punishment. *American Behavioral Scientist*, 39, 474–487.
- Goldberg, J. H., Lerner, J. S., & Tetlock, P. E. (1999). Rage and reason: The psychology of the intuitive prosecutor. *European Journal of Social Psychology*, 29, 781–795.
- Goodstadt, B., & Kipnis, D. (1970). Situational influences on the use of power. *Journal of Applied Psychology*, 54, 201–207.
- Greene, E. J., & Darley, J. M. (1998). Effects of necessary, sufficient, and indirect causation on judgments of criminal liability. *Law and Human Behavior*, 22, 429–451.
- Hamilton, L. V., & Rytina, S. (1980). Social consensus on norms of justice: Should the punishment fit the crime? *American Journal of Sociology*, 85, 1117–1144.
- Hogarth, J. (1974). *Sentencing as a human process*. Toronto, Canada: University of Toronto Press.
- Jacoby, J., Jaccard, J. J., Kuss, A., Troutman, T., & Mazursky (1987). New directions in behavioral process research: Implications for social psychology. *Journal of Experimental Social Psychology*, 23, 146–175.
- Kant, I. (1952). The science of right. (W. Hastie, Trans.). In R. Hutchins (Ed.), *Great books of the western world* (pp. 397–446). Edinburgh: T. & T. Clark (original work published 1790).
- 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.
- Lerner, M. J., & Simmons, C. H. (1966). Observer's reaction to the "innocent victim": Compassion or rejection? *Journal of Personality and Social Psychology*, 4, 203–210.
- Lerner, M. J., & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin*, 85, 1030–1051.
- McFatter, R. M. (1978). Sentencing strategies and justice: Effects of punishment philosophy on sentencing decisions. *Journal of Personality and Social Psychology*, 36, 1490–1500.
- McFatter, R. M. (1982). Purposes of punishment: Effects of utilities of criminal sanctions on perceived appropriateness. *Journal of Applied Psychology*, 67, 255–267.
- Nagin, D. S. (1998). Deterrence and incapacitation. In M. Tonry (Ed.), *The handbook of crime and punishment* (pp. 345–368). New York: Oxford University Press.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.
- Patterson, C. M., Kosson, D. S., & Newman, J. P. (1987). Reaction to punishment, reflectivity, and passive avoidance learning in extraverts. *Journal of Personality and Social Psychology*, 52, 565–575.

- Pennington, N., & Hastie, R. (1993). The story model for juror decision making. In R. Hastie (Ed.), *Inside the juror: The psychology of juror decision making* (pp. 192–221). NY: Cambridge University Press.
- Reyna, C., & Weiner, B. (2001). Attributional analysis of the goals of teachers' punishment and intervention strategies. *Journal of Educational Psychology*, 93, 309–319.
- Roberts, J. V., & Gebotys, R. J. (1989). The purposes of sentencing: Public support of competing aims. *Behavioral Sciences and the Law*, 7, 387–402.
- Robinson, P. H., & Darley, J. M. (1997). The utility of desert. *Northwestern University Law Review*, 91, 453–499.
- Rossi, P. H., Berk, R. E., & Campbell, A. (1997). Just punishment: Guideline sentences and normative consensus. *Journal of Quantitative Criminology*, 13, 267–290.
- Rossi, P. H., Waite, E., Bose, C. E., & Berk, R. E. (1974). The seriousness of crimes: Normative structure and individual differences. *American Sociological Review*, 39, 224–237.
- Rucker, D., Polifroni, M., & Tetlock, P. E. (2004). On the assignment of punishment: The impact of general-societal threat and the moderating role of severity. *Personality & Social Psychology Bulletin*, 30, 673–684.
- Sommers, S. R., & Ellsworth, P. C. (2000). White juror bias: An investigation of prejudice against Black defendants in the American courtroom. *Psychology, Public Policy, and Law*, 7, 201–229.
- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72, 271–322.
- Tyler, T. R., & Weber, R. (1983). Support for the death penalty. *Law and Society Review*, 17, 201–224.
- Vidmar, N., & Miller, D. (1980). Social psychological processes underlying attitudes toward legal punishment. *Law and Society Review*, 14, 565–602.
- Weiner, B., Graham, S., & Reyna, C. (1997). An attributional examination of retributive versus utilitarian philosophies of punishment. *Social Justice Research*, 10, 431–452.
- Yukl, G., Falbe, C. M., & Joo, Y. Y. (1993). Patterns of influence behavior for managers. *Group & Organization Management*, 18, 5–28.
- Zimring, F. E., & Hawkins, G. (1995). *Incapacitation*. NY: Oxford University Press.