

The Reasons We Punish: Creating and Validating a Measure of Utilitarian and
Retributive Punishment Orientation

by

Susan Yamamoto

A thesis submitted to
the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of

Master of Arts

in

Psychology

Carleton University
Ottawa, Canada

©2014 Susan Yamamoto

Abstract

Previous researchers have investigated the situational use of punishment, but the overall reasons why people punish have received less attention. The aim of this study was to create a measure of individual differences in punishment orientation. 200 participants completed a 30-item questionnaire designed to measure retributive and utilitarian punishment orientation. Exploratory factor analysis uncovered a ‘pro-punishment’ factor, as well as ‘ideal retributive’ and ‘ideal utilitarian’ factors. An additional sample of 200 participants completed a revised version of the scale; confirmatory factor analysis yielded acceptable model fit for harsh utilitarian, harsh retributive, and ideal retributive dimensions. The scale showed poor divergent validity, with the factors having moderate relationships with attitudes toward the legal system (Schiffhauer & Wrightsman, 1995). Predictive validity assessments showed that participants favoured one orientation dependent on the context, resulting in poor predictive utility. This scale may nonetheless contribute to a better understanding of lay punishment ethics.

Keywords: retribution, utilitarian, punishment motive, trait punishment, deontology

ACKNOWLEDGMENTS

I would like to extend my appreciation to a number of people who supported me in completing my thesis. I would like to thank my supervisor, Dr. Evelyn Maeder, whose patience, dedication, and mentorship made this study possible. Through hours of meetings and draft readings, with a lot of care and encouragement, she enabled me to reach my goals. I am also grateful to Dr. Heidi Maibom and Dr. Andrea Howard, who lent their time and helpful guidance in executing this project. Dr. Adelle Forth and Dr. Johanna Peetz have provided invaluable advice and support in shaping this thesis; I thank them for all of the time and thought they have put in to helping enhance the product. I would also like to acknowledge Dr. Ummni Khan for her time and insightful advice on this project.

I am indebted to Holly Ellingwood and Kris Merrells for all of their hard work and commitment in coding a massive amount of data. Their contribution to my project leaves me not only grateful, but in admiration of their generosity. Finally I would like to thank Laura McManus and the Jury Lab, who supported me at every step.

TABLE OF CONTENTS

INTRODUCTION.....	1
- Retributivism.....	3
- Utilitarianism.....	5
- Differentiating Between Retributive and Utilitarian Punishment.....	8
- Canadian Sentencing Principles.....	10
- Attitudes.....	12
- State Punishment.....	18
- Trait Punishment.....	21
- Behavioural Approaches.....	22
- Comparing State and Trait Punishment Orientation.....	25
- The Need for a Well-constructed Questionnaire.....	27
STUDY 1: ITEM DEVELOPMENT.....	27
- Hypotheses.....	28
METHOD.....	28
- Participants.....	28
- Materials.....	29
- Procedure.....	31
RESULTS.....	31
- Item Assessment.....	31
- Factor Analysis.....	33
- Further Scale Refinement.....	38
- Other Exploratory Measures.....	40

DISCUSSION.....	42
STUDY 2: SCALE VALIDATION.....	45
- Hypotheses.....	46
METHOD.....	48
- Participants.....	48
- Materials.....	48
- Procedure.....	54
RESULTS.....	54
- Item Assessment.....	54
- Reliability.....	56
- Confirmatory Factor Analysis.....	57
- Construct Validity.....	60
- Predictive Validity.....	63
- Exploratory Analyses.....	74
DISCUSSION.....	75
GENERAL DISCUSSION.....	83
REFERENCES.....	93

LIST OF TABLES

TABLE 1: Refined Punishment Orientation Questionnaire with factor loadings based on principal axis factoring method with varimax rotation.....	37
TABLE 2: Frequency and percentage of participant responses to question “It is better to let ___ guilty people go unpunished than to punish 1 innocent person.”.....	41
TABLE 3: Frequency and percentage of participants picturing prominent crimes.....	41
TABLE 4: Significant correlations between the Punishment Orientation Questionnaire items and Impression Management Scale of the Balanced Inventory of Desirable Responding.....	56
TABLE 5: Fit indices for initial model.....	57
TABLE 6: Correlations among dimensions of the Sentencing Goals Scale.....	60
TABLE 7: Correlations among dimensions of the POQ and the SGS.....	60
TABLE 8: Correlations between ATLS subscales and POQ subscales.....	62
TABLE 9: Two-way contingency table for familiarity with trolley dilemma and dilemma decision.....	67

LIST OF FIGURES

FIGURE 1: Final 3 factor model with item loadings and factor covariance.....59

LIST OF APPENDICES

APPENDIX A: Recruitment Notice.....	100
APPENDIX B: Informed Consent Form	101
APPENDIX C: Punishment Orientation Questionnaire.....	103
APPENDIX D: Balanced Inventory of Desirable Responding	107
APPENDIX E: Demographics Questionnaire.....	108
APPENDIX F: Punishment Scenarios	109
APPENDIX G: Sentencing Goals Scale... ..	111
APPENDIX H: Scale of Attitudes Toward the Legal System.....	115
APPENDIX I: Need for Cognition Scale.....	118
APPENDIX J: Positive and Negative Affect Scales.....	120
APPENDIX K: Scree Plot from Initial Exploratory Factor Analysis.....	121
APPENDIX L: Revised Punishment Orientation Questionnaire	122
APPENDIX M: Final Punishment Orientation Questionnaire by Subscale.....	124
APPENDIX N: Descriptives for Study 2 Measures.....	125
APPENDIX O: Final Scenario Decision and Justification Frequencies.....	126

The Reasons We Punish: Creating and Validating a Measure of Utilitarian and Retributive Punishment Orientation

Punishment is a crucial part of social interaction; as Weiner (2006) asserted, “life is a courtroom” (p.4). We penalize others for offences ranging from tardiness to murder. Imposing a penalty on another can have grave consequences whether it be expulsion from school or a life sentence, yet a complete picture of the social aspect of punishment is wanting. Behaviourists have demonstrated how receiving punishment shapes our actions, and philosophers have debated what ought to justify punishment. Understanding why people feel compelled to punish others – their punishment orientation – is a relatively recent endeavour. Owing to the intuitive nature of such moral decision-making, there is little empirical data to contemplate this key issue (Carlsmith, 2001). Researchers have shown that one’s beliefs about how much the system treats people fairly can even predict compliance with those rules (Carlsmith, Darley, & Robinson, 2002). The Criminal Code of Canada reflects a number of distinct punishment principles, from preventing future offences through jail time, to considering mitigating circumstances to ensure proportional sanctions, as demonstrated through longer sentences for murder than for manslaughter. The recent proposal of the Not Criminally Responsible Reform Act in Canada – in which mentally ill defendants could receive a ‘high risk’ designation and the priority is public safety – demonstrates a stricter adherence to principles of utility than due process, since legal guilt has not actually been established. In order to understand the compatibility of a person’s punishment orientation with the official penal system, and the consequences of perceived unjust punishment, it is necessary to investigate not only the experience of being punished, but also the experience of punishing.

Vidmar and Miller (1980) defined punishment as: “a negative sanction intentionally applied to someone perceived to have violated a law, rule, norm, or expectation” (p. 568). Borne out of a long history of philosophical discussion, two justifications for punishment have emerged: a concern with just deserts (i.e., an eye for an eye) and a concern with preventing future transgressions – termed retribution and utilitarianism respectively. While the psychological research to date converges on these motives, there is disagreement as to what constructs underlie them. Researchers have traditionally distinguished between psychological traits – characteristics of a person that are largely enduring and relatively stable – and states, which may change according to the situation and are relatively fleeting (Nezlek, 2007). Still other researchers prefer an ‘interactionist’ perspective, according to which states and traits have a joint influence. Accordingly, some researchers contend that the motives that drive punishment are strictly state-based, and are invoked by situational variation (Weiner, 2006). Conversely, some researchers argue that punitive decisions are driven mainly by a retributive orientation (Carlsmith et al., 2002). The current literature gives a clear picture of the kinds of scenarios that lend most naturally to each punitive approach; however, more research is needed to understand the role of individual differences in this emerging state-trait debate. Since few studies have sought to measure trait-level punishment, the purpose of this study was to create and validate a punishment orientation scale that addresses the question: do certain people tend more naturally toward each respective goal? The current study aimed to demonstrate: (a) the necessity of creating and validating a new measure of punishment orientation, and (b) that punishment orientation is likely both a state and a trait. The sections to follow define retribution and utilitarianism, discuss their similarities

and differences, discuss Canada's sentencing principles as they relate to punitive goals, and present both state and trait based perspectives to date. Finally, the creation of a punishment orientation scale is described.

Retributivism

Retributive theory has roots in a philosophical stance holding that a moral act should be good in and of itself (Schedler, 1980). Most closely associated with the philosopher Kant (1785/2006), this theory prescribes a world in which people act only in a way that would be permissible for all others to do the same in those circumstances, a principle known as the 'categorical imperative': "Act only on that maxim through which you can at the same time will that it should become a universal law" (p. 286). A second, related dictum suggests that people must never be used as a means to an end; rather, they are to be treated as ends unto themselves. Kant's theory is referred to as a theory of 'the right,' in which consequences cannot dictate morality, rather only a 'good will' can do so (Kant, 1785/2006). Moral acts are considered duties; if one, for example, receives too much change from a store clerk, he or she must return it not from fear of consequences, but because it is a moral imperative (Kant, 1785/2006). Even if an action will lead to a positive result, some acts are simply incompatible with moral duties. In essence, retributive theorists believe a punishment must be given purely because it is right and not as a means to an end; innocent people cannot be sacrificed to achieve some desired goal. Therefore, they must focus on the past wrongdoing in order to balance out the scales of justice (Weiner, 2006). The retributivist, though not explicitly opposed to concepts such as deterrence, is unconcerned with the future benefit of the punishment. Rather, the age old wisdom: 'an eye for an eye' most readily comes to mind. It is of paramount concern

that a punishment be proportional to the harm done (Schedler, 1980). A person does not have the right to enact a punishment that exceeds the offence, but may only correct a privilege unjustly held by the offender (Goldman, 1979). It is important to note that proportionality does not necessarily require that the offence be matched in kind, but only that it is matched in magnitude (Goldman, 1979). The current study will define retributive punishment orientation as the propensity to seek a proportional punishment for the purpose of righting a past wrong regardless of the usefulness of the punishment.

The theory behind this punishment orientation is sometimes criticized for problems with its application (Heubel & Biller-Andorno, 2005). For instance, the excessive focus on the individual may, at times, create difficulty in appreciating the person in relation to others, where a theory focusing on maximizing the overall good might be better equipped. It can be difficult to gauge how universal moral duties should be applied to unique, morally ambiguous situations (Heubel & Biller-Andorno, 2005). After all, consequences do seem to matter in many cases, and there is considerable nuance from one situation to another that may require more flexibility than is afforded by a strict definition of 'right'. Furthermore it can be challenging to define what is 'proportional' to the harm done. However, retributivism can be commended for its refusal to use people as means to an end, thereby supporting individual rights more than its counterpart, utilitarianism. Interestingly, some researchers assert that people prefer a retributive standpoint when evaluating individual cases, while not being in favour of such principles abstractly (Carlsmith et al., 2002). In this way, a theory that accounts for how a group may function in harmony might operate at an overarching level, while a retributive, individual based theory is mobilized when undermining individual rights is deemed

intolerable.

Utilitarianism

Commonly associated with the philosopher Bentham (1789/2006), and later built upon by Mill (1859/2008), the central tenet of utilitarian philosophy is that an act must maximize the aggregate good for those affected by that act. Therefore, the utilitarian is more concerned with the overall well-being of the group than with individual rights; for a punishment to be justified, it must produce some positive effect. In fact, one important distinction between retributive and utilitarian theory is that the latter is ‘consequentialist’, where questions of deservedness do not factor in to the punitive decision.

Utilitarian punishment can be broken down in to three subsets: incapacitation, rehabilitation, and deterrence. While each is conceptualized under utilitarian theory, these goals have important distinguishing features that emerge through a psychological lens. First, incapacitation seeks simply to prevent a particular offender from committing another crime. For example, the punishment of exile would satisfy this goal; a person is no longer physically present in order to commit another crime (Carlsmith, 2001). Weiner (1985) likens this motive to a “time out” often used in operant conditioning, where the goal is to stop the behaviour (p.433). Second, the rehabilitative approach is marked by an attempt to not only incapacitate but also to reform the offender. Simple exile would be insufficient to satisfy the goal of rehabilitation, because the offender is not assisted in creating change. The rehabilitationist may wish to morally educate the offender and encourage him or her to internalize the values of the group (Weiner, 1985). According to Weiner, this pro-social helping behaviour is associated with feelings of sympathy for the transgressor, noting that it is necessary to believe that a person is capable of change in

order to seek this goal. Third, the deterrence approach seeks to prevent future similar offences. Deterrence can be of either specific or general form. In line with a behaviourist perspective of punishment, specific deterrence aims to reduce future offences by creating in the transgressor an aversion to the act. General deterrence aims to discourage other would-be transgressors from committing similar acts such that when contemplating a cost benefit analysis, the possibility of a severe punishment outweighs the smaller benefit of the crime. While a rehabilitationist will likely have sympathy for an offender, it may not be a prerequisite for deterrence. Preventing others from committing similar crimes may require a punishment far harsher than the crime itself. For instance, a society may wish to publicize a punishment and ensure it exceeds the severity of the crime in order to make the act unappealing to others (Carlsmith et al., 2002). Hence, sympathy will sometimes be superseded by a need to protect society from danger.

It seems clear that these utilitarian goals indeed share a view toward promoting future good through preventing future offences. For a rehabilitative goal, an attempt to change the offender for the better simultaneously aims at removing the potential for harm from society. Incapacitating the offender and deterring others serves the same aim. Indeed, unlike for retributivists, a utilitarian punitive act does not appear to satisfy a psychological need independently of its usefulness; what it may later accomplish is foremost. The current study will therefore define utilitarian punishment motive as a propensity to punish others to reduce future offenses without regard for proportionality. Weiner, Graham, and Reyna (1997) contended that rehabilitation and deterrence should not be conceptualized as one orientation as each has a distinct relationship with attributional constructs, such as whether the offence was a fleeting occurrence.

Consequently, it is possible that a rehabilitative orientation will emerge as a separate orientation rather than a component of utilitarianism; items specifically addressing rehabilitation will therefore be created.

One of the most well known demonstrations of a utilitarian dilemma concerns a thought experiment describing a hypothetical trolley that is barrelling towards five people (Foot, 1967). The question posed is whether or not it is permissible to derail the trolley to save the five people at the expense of killing one person in the other path. Through a utilitarian cost-benefit analysis, one can conclude that saving the five will likely result in a greater amount of happiness. However, to a retributive theorist it is not morally permissible to undermine individual rights in this manner; to undermine one person's rights is to undermine every person's rights. Thomson (1985) later drew attention to the complete reversal of these intuitions in seemingly similar dilemmas. For example, those willing to derail the trolley are unwilling to take five organs from a healthy person and provide them to five patients in need. This distinction raises an interesting question regarding the application of punishment to an innocent person. On one hand, we might expect that driven by a desire to protect society, utilitarian punishers are willing to *permit* harm to an innocent provided the benefit is great. This is especially plausible if the lay-punisher considers the inevitable punishing of innocents to be a naturally occurring event rather than a deliberate harm. On the other hand, as might be true for jurors, meting out punishment could be viewed as the direct imposition of harm on an innocent. It is possible that lay-punishers will intuitively draw this distinction, and this notion also helps account for the preference of utilitarian theory abstractly rather than in actual application.

It is also evident that this consequentialist theory has shortcomings, which

retributivism can answer to. One major criticism of utilitarianism is that it is counter-intuitive. For example, it is apparent where the previously described trolley case becomes problematic if the person to be sacrificed for the greater good is one's mother. Like retributivism, though, it can be difficult if not impractical to calculate the cost-benefit analysis, especially without lengthy consideration as when a trolley is fast approaching. Finally, utilitarianism can be said to undermine individual rights; for the greater good, it may be necessary to forfeit one's rights, considering the aggregate benefits each person might enjoy from the sacrifices of others. Retributivism, however, contends that some acts are simply wrong, thereby preventing excessive harms to individuals. While neither theory is without criticism, Weiner, Graham, and Reyna (1997) cautioned that the distinction between the theories is blurred at times; therefore, it is imperative to consider what key features separate these theories and how they can potentially converge.

Differentiating Between Retributive and Utilitarian Punishment

Key differences. In understanding the origin of these motives, a question necessarily arises as to whether they are mutually exclusive both in terms of psychological function and logical premises. In fact, there have been some attempts to answer the latter question in the philosophical literature by means of a hybrid theory. For example, Goldman (1979) discusses the "paradox of punishment"; namely, punishment seems to be both necessary and unjustifiable (p.42). In other words, a punitive system serves the function of keeping society safe and orderly, while inevitably 'misfiring' from time to time (Schedler, 1980). As Lessnoff (1971) pointed out, under a hybrid conceptualization of punishment, we could follow a system in which the goal is deterrence with the caveat that it is applied only to rights forfeited through the offence.

The presence of some flaws in individual punishments would fail to negate the system at large. However, Goldman cautioned that such a theory is at best as problematic as each individual theory. For instance, the retributive principle of proportionality is difficult to coalesce with the excessive punishment that may accompany a utilitarian viewpoint; in order to be deterrent, a punishment must sometimes unfairly violate rights beyond those forfeited through a crime.

Based on the different requisite conditions for each motive, it is likely that those who show strong support for one goal may need to violate certain morals in order to endorse the other. For instance, retributivists require that a punishment be deserved in order to be justified; a focus on the specific wrongdoing is key to the morality of the act. Yet for a utilitarian, the preferences of the individual are often necessarily ignored. What is required for one to be justified is potentially intolerable to the other. As Rawls (1955) contended, “no justification is without those who detest it” (p.4). A further distinction, as Rawls argued, is that utilitarian decisions concern the overall system whereas retributive decisions focus on specific rules and cases. In effect, the retributivist responds to the question of why person X was punished whereas the utilitarian responds more readily to the question of why people are punished (Rawls, 1955). While the retributive outlook is backward-looking and concerned with righting a past wrong, the utilitarian outlook is forward-looking with a view to improving society (Weiner, 2006). The distinction is essentially: utilitarians punish because they have to (as a means to an end), whereas retributives punish because they want to (Carlsmith et al., 2002). It is therefore of interest to identify whether lay persons will generally divide among these punishment ideals, to the aim of illuminating the types of punishments that may naturally conflict with their

overall orientation. To illustrate, it is useful to consider the principles behind Canada's sentencing policies.

Canadian sentencing principles. The hybrid nature of Canada's sentencing principles as stated in the Criminal Code warrants discussion. That is, all three subsets of utilitarianism as well as retributive principles are highlighted as fundamental objectives. Deterrence is both explicitly stated and theoretically alluded to. One objective aims to "deter the offender and others from committing the same crime" while endeavouring to "denounce illegal conduct" (718). Similarly, the criminal code aims to "separate offenders from society" and "assist in rehabilitating the offender" (718). Within the same list of objectives, retributive principles are encompassed by the goal of promoting "a sense of responsibility in the offender" (718). Further, section 718.1 requires that a sentence be "proportionate" to both the seriousness of the crime and the responsibility of the offender. Likewise, section 718.2 outlines the consideration of aggravating and mitigating circumstances in imposing a sentence.

Two principles, taken together, illustrate the difficulties with simultaneous adherence to strict retribution and utilitarian systems. First, section 718.2 (b) states that sentences should be alike for "similar offenders for similar offences committed in similar circumstances." This objective underscores the importance of consistency in sentencing, encompassing the retributive concern with deservedness in assigning punishment. In other words, the system cannot simply punish at will based on utility. Second, the assertion that "an offender should not be deprived of liberty, if less restrictive sanctions may be appropriate in the circumstances" is difficult to accommodate in considering punishment on a macro-level. This retributive focused objective is best understood in

reference to justice on a case by case basis. Recalling, as Schedler (1980) pointed out, that human systems of justice will most certainly fall through at times, it seems all but impossible to merge these objectives with the nature of a flawed system, as well as the need to deter others from similar offences. As a result we might conclude, as advocated by hybrid theorists, that the goal of deterrence is not permissible should individual rights be unfairly violated, unless that punishment is construed as “necessary evil” (Goldman, 1979, p.52). This embodies the pillars of a hybrid theory; an ideal system can ignore the fact of individual injustices since it is purely theoretical. After all, a utilitarian account is quite appealing until one is confronted with a specific miscarriage of justice.

While these punishment goals are in many ways logically incompatible, the possibility that people do not necessarily distinguish between the two should be considered. While they did not have the benefit of a well-validated trait punishment measure, Carlsmith et al. (2002) reported high agreement with both retributive and utilitarian focused items, whereas participants appeared behaviourally retributive. That is, when participants were asked to select the most appealing punishment theory from descriptions of utilitarianism and retributivism, many endorsed a utilitarian goal. However, when seeking information in order to make a punitive judgment, participants disproportionately used information that the researchers deemed retributive (i.e. crime severity, mitigating circumstances). As Carlsmith (2006) pointed out, these two goals converge on certain appropriate punishments – although for different reasons – and may also completely disagree as to what is fair. Rather than asking participants to categorize their orientation outright, it will be helpful to study the aspects of each theory that create reluctance in some to endorse either motive. While Carlsmith et al. (2002) importantly

questioned the distinction between an ideal punishment system and actual behaviour when evaluating specific cases, Canada's sentencing principles clearly advocate both systems as overarching objectives. It is possible that ideal personal views will similarly encompass both systems. For instance, people might follow utilitarian principles until those values come into conflict with underlying retributive needs. It may be that there are also those for whom there is no retributive threshold at which excessive punishment becomes intolerable. The present study will also consider the question of whether there is a distinction between retributive and utilitarian viewpoints in terms of macro and micro punishment (Carlsmith, Darley, & Robinson, 2002).

Attitudes

As the goal of this study was to create a more precise measure of punitive attitudes, it is helpful to consider what type of attitudes the questionnaire might capture. Banaji and Heiphetz (2010) described attitudes as "the fundamental orientation to evaluate people, other living beings, things, events, and ideas along a good-bad dimension" (p. 382). As Millar and Tesser (1986) asserted, attitudes are an interplay between affect, cognition, and general evaluations. As researchers have found a disconnect between punitive attitudes and behaviour (see Carlsmith, Darley, & Robinson, 2002; and Carlsmith, 2006), it is essential to consider what type of attitude is under study so as to better predict behaviour. The research to date concerning punishment has focused heavily on more general evaluations, tending to measure attitudes toward the effectiveness of the actual penal system.

To measure attitudes toward criminal sanctioning, Cullen, Clark, Cullen, and Mathers (1985) utilized a six item measure of retribution, deterrence, and incapacitation,

which demonstrated moderate reliability, for example: “criminals deserve to be punished because they have harmed society” (p.316). In particular, they were interested in the relationship between punishment and attributions about the nature of crime, hypothesizing that those who view crime as a rational undertaking will be more punitive than those who view crime as arising from social strain (i.e., classical versus positivist belief systems). Among these six items, while three items appear to address retribution and utilitarianism directly, the remaining items do not distinguish between the efficacy of criminal sanction and the desire for punishment, for example “putting people in prisons does not make much sense since it will only increase crime because prisons are schools for crime” (p. 316). While this item appears more utilitarian, it may be equally attractive to a retributive, who is not per se against the goal of deterrence and may not believe that prison is an appropriate, proportional punishment for certain crimes. Cullen et al. also included nine items, with high reliability, measuring the extent to which participants believe that rehabilitation is an effective sanction, for example: “the rehabilitation of adult criminals just does not work” (p. 316). They reported that those with a positivist viewpoint tended to endorse the rehabilitative items. Although whether people believe the punitive system is effective is an important inquiry, it is only one subset of a complex collection of punishment attitudes. It is necessary to tease out the motive behind a punishment to permit a comparison with those attitudes.

Carroll, Perkowitz, Lurigio, and Weaver (1987) studied individual differences in sentencing goals and causal attributions in law and criminology students and probation officers. They created and factor analyzed a sentencing goals scale, revealing a three factor structure of ‘satisfactory performance,’ ‘punishment,’ and ‘rehabilitation’. Items

were grouped according to belief in the effectiveness of the system, belief that the system is too lenient, and the need for better rehabilitation programs. They found that rejecting individual crime causation predicted rehabilitation orientation, and belief that individuals cause crimes predicted orientation toward 'punishment' as opposed to aiding the offender to change. While Carroll et al. demonstrated clear links between causation and punishment attitudes, there is an important distinction between attitudes regarding the abstract function of a penal system and the general evaluations of the system's ability to fulfill that function.

As Millar and Tesser (1986) demonstrated, the relationship between attitudes and behaviour is strongest when there is congruency between the cognitive versus affective component of thought, and the extent to which the behaviour is affectively or cognitively driven. To demonstrate, Millar and Tesser prompted either an affective thought component – by asking participants to report on their feelings about a puzzle task - or a cognitive thought component, by asking participants to report their thoughts about the puzzle task. They also varied whether the task was instrumental (i.e. cognitively driven) or done for intrinsic interest (i.e. affectively driven). Participants were given the opportunity for free-play with the puzzles, which was correlated with their evaluation of the puzzles. Correlations between the self-report about the puzzles and time spent doing them supported their theory. That is, participants' evaluations were most predictive of free-play time when there was congruence between the cognitive or affective component of the attitude and the cognitive or affective component driving the behaviour. As aforementioned, researchers suggest that while lay-punishers endorse utilitarianism, they seek retributive information and respond more or less harshly to variations in retributive

information. In other words the supposed utilitarian attitude does not seem to reflect the retributive punishment behaviour.

The necessity of clearly defining the scope of a measure is further underscored by a myriad of other factors influencing the predictive utility of an attitude. For instance, Weigel, Vernon, and Tognacci (1974) pointed out that while attitudes are often conceptualized and measured on a general level, assessment of behaviour in specific circumstances is inherently more focused, therefore the disconnect is not surprising. Hence an exploratory study that will help tease apart the different variables involved in punitive attitudes will not only aid in defining the nature of punishment orientation, but will also allow for more specific measures to be adapted for future studies. A punishment orientation measure aimed at specific types of crimes or punishments could be created for different kinds of studies, for example depending on whether they involved daily, minor punishment or capital punishment. It is therefore prudent to be specific about the type of attitude being measured before concluding that there is an absolute disconnect between punitive behaviour and punitive orientation.

Without a well-established measure of punishment orientation, it is difficult to assess the manner in which punitive affects and cognitions operate on an overarching level. However, one possibility is that utilitarianism is representative of a more cognitive based focus, whereas retributivism is a more affectively driven goal. Traditionally, utilitarianism is described as a practical, calculated morality and appears more instrumental than retributivism, wherein punishment is desired for its own sake. As Greene et al. (2001) argued, utilitarian responses are foremost when the “deliberative” systems are at play. Myers and Greene (2004) also found that ventromedial pre-frontal

cortex damage patients were more likely to respond in favour of utilitarian choices to moral dilemmas. Greene (2011) later proposed and found support for a ‘dual process’ theory of moral judgment using functional magnetic resonance imaging, suggesting that automatic emotional reactions prompt deontological (i.e. retributive) judgments, whereas cognitive processes drive utilitarian responding. Finally, Bartels and Pizarro (2011) recently demonstrated that those who scored high on psychopathy and Machiavellianism measures tended toward utilitarian responding. Recalling that retributivism prohibits the punishment of innocents, it is possible that absent a retributive, affectively driven orientation, participants are better able to commit the moral violation of sacrificing a person. This possibility helps account for laypersons’ abstract preference for utilitarianism. As a cognitive assessment of punishment, it is an appealing ideal system, until the moral violation is salient.

Conway and Gawronski (2013) used a process dissociation (DP) procedure to investigate Greene et al.’s (2001) dual process theory of moral judgment. Noting the limitations of the assumption that utilitarian and deontological reasoning represent polar opposites, Conway and Gawronski (2013) demonstrated that there is an important distinction between whether a utilitarian response has been suppressed, deontological judgment increased, or both. They therefore cautioned against the conclusion that a given response necessarily represents a utilitarian orientation; rather, it may also be that a deontological inclination has decreased. For instance, they found that when participants were asked to remember a series of letters and numbers prior to each dilemma, they tended toward deontological rather than utilitarian judgments. Comparing the classic treatment of the constructs as bipolar with results from the process dissociation

procedure, the researchers found that while such increasing of cognitive load indeed led to different judgments, the PD approach was able to demonstrate that it decreased utilitarian inclinations but did not impact on deontological ones. Importantly, they underlined the issue that a presumed bipolar relationship between the two orientations may fail to illuminate certain relationships with other constructs. For instance, while the traditional bipolar measure revealed no significant relationship with moral identity internalization, both orientations were significantly positively related when using the PD measures.

The contention that utilitarianism and deontology represent independent variables impacted on the current study in two important ways. First, this assertion was suggestive that rather than a unidimensional factor structure, wherein participants' agreement with utilitarian valenced items would inversely relate to endorsement of retributive items, multiple factors that represent utilitarian and retributive intuitions might emerge. Second, it was possible that, in interpreting the factor structure, an orthogonal rotation would be appropriate. Further, some items in the initial questionnaire were created to represent a conflict between utilitarianism and retributivism while others allowed participants to endorse both ideals. That is, separate questions probing the importance of deterrence and balancing the scales of justice could both result in agreement.

This framework is one among many potential explanations. It bears considering that a retributive outlook could also function at an abstract, ideal level. For example, one may be in favour of the abstract notion that 'no innocent should ever be punished,' while recognizing that in certain cases of extreme violence or danger, it may be necessary to commit excessive punishment. For example, one scenario included in this study features

a dilemma as to whether a pair of twins, one of whom is certainly guilty, should be imprisoned to prevent later sexual assaults. It is possible that fear could be implicated in a utilitarian response. Therefore, it could also be that affective punitive decisions occur at a case-by-case level, rather than an overarching one. Under this conceptualization, we might expect that a trait punishment measure would capture cognitive components of retributivism and utilitarianism, whereas affective decisions would be illustrated via specific punitive decisions. While it is certainly true that cognition and affect are linked, it is important to assess whether one attitude component is in focus during punitive decisions. To better assess the possible link between cognition and punishment attitudes, participants completed the Need for Cognition scale (NfC; see appendix I), which measures participants' enjoyment of and engagement in "effortful cognitive endeavours" (Cacioppo, Petty, & Kao, 1984, p. 306). In order to investigate the possible relationship between punishment and affect, participants also completed Watson, Clark, and Tellegen's (1988) Positive and Negative Affect Scales (PANAS; see appendix J).

State Punishment Orientation

A myriad of research has already made progress in understanding punishment as a state-based construct. Most notably, Weiner (1985) has applied an attributional theory of motivation to what he termed 'punishment goal.' According to Weiner's attribution theory, as a consequence of an important outcome, people naturally search for a causal explanation. These causal attributions then lead to a series of cognitive, affective, motivational and behavioral consequences according to: locus of causality (internal or external), stability (enduring or changing), and controllability (having an effect on the outcome). This theory has considerable application in diverse areas of study. In fact,

Weiner, Graham, and Reyna (1997) have demonstrated that similar patterns emerge when employing this taxonomy to both achievement related events and criminal cases. Under Weiner's conceptualization of 'life as the courtroom', it becomes evident that responsibility and the desire to punish are highly connected. Responsibility, after all, carries an implication of free will and intentionality (Carroll et al., 1987; Weiner et al., 1997).

Weiner (1985) suggested that both retribution and utility could be mapped out according to these antecedent causal ascriptions. For example, if a wealthy person steals money due to greed, perceived as having a choice to steal and to have done so repeatedly, this internal, controllable, stable ascription would lead to an assignment of responsibility, the emotion of anger, and a retributive punishment. On the other hand, if a person stole money because of something external – in that particular study a car accident resulted in brain injury causing an uncontrollable and fleeting lapse in judgment – sympathy would result. As a consequence of this sympathy, an observer is more likely to be motivated by utility than retribution. This account is consistent with the idea that utilitarian punishment is more so used out of necessity than out of a deep-seated psychological desire for punishment. It is important to note that this conceptualization of 'pro-social' behavior is intuitively related to the rehabilitative approach to utilitarian punishment. That is, a deterrence approach could also be sated by an overly harsh sentence where sympathy may be counterproductive to the required punitive behaviour. Weiner (1985) cautioned that the distinction between the many subsets of punishment is at times blurred, which suggests at least that they may not be orthogonal.

The contention that punitive goals do not precede evaluation of transgressions

underscores the importance of situational factors in determining moral action. However, in their structural equation analysis of punishment severity with both attributional predictors and general punitive goals, Weiner, Graham, and Reyna (1997) showed that beliefs about the cause of the crime, emotional responses, and general punitive goals acted in orchestra to influence punishment. Moreover, to investigate the possible role of trait-like punishment goals, these researchers only used two items, one to measure each orientation. A more comprehensive measure of trait-like punishment will therefore augment state-type explorations. One important area where Weiner et al. (1997) aided in classifying lay punishment orientation is their demonstration of the divergent consequences of instability attributions. Whereas believing an event was unstable led to endorsement of a rehabilitative stance, it was negatively related to a desire to protect society. Therefore Weiner et al. (1997) advise against inclusion of rehabilitation under the umbrella of utilitarianism. Likewise, they discussed the distinction between two potentially divergent retributive viewpoints. The first may be considered “strict liability,” predicated on balancing the scales according to the magnitude of the harm done. Second, one may prefer to mete out punishment on the basis of a guilty mind, without strict adherence to retaliation equal to the actual harm.

Weiner (1985) further qualified that the study focused not on the question of which situations bring about such causal ascriptions, rather, once these causal ascriptions are already determined, which punitive goal is foremost. The prominent question in the current study concerned the tendency for people to employ certain explanations and beliefs about transgressions more generally, where the related punitive goal could potentially follow a pattern similar to that Weiner described. However, an approach that

probes for the goals that people find most attractive in the absence of a punitive scenario may aid in uncovering what people's ideal punitive systems are. A discussion of the evidence for trait-punishment orientation will therefore augment exploration of such framework.

Trait Punishment Orientation

While such state-based theories have empirical support, there is much less information as to the possibility of punishment orientation as an over-arching belief paradigm; yet, it is still necessary to understand the belief system with which one first enters into a situation. Drawing on the work of Carroll et al. (1987), McKee and Feather (2008) created the Sentencing Goals Scale that yielded four categories: retribution, rehabilitation, deterrence, and incapacitation. They investigated the relationship between sentencing goals and vengeance attitudes. Participants rated their agreement or disagreement to items on a 7-point scale, for example: "the purpose of court sentences should be to rehabilitate the criminal" (p.147). All four scales demonstrated moderate internal consistency reliability. They found a negative relationship between a desire for personal revenge and rehabilitative goals. Conversely, a positive relationship between desire for personal revenge and both retributive and incapacitative goals was illustrated. This relationship is consistent with traditional conceptualizations of retribution as 'getting back' at the offender.

McKee and Feather's (2008) scale marks an excellent starting point for measuring punishment orientation; however, the generic wording of some of the items may not be a sensitive enough measure of strict adherence to each respective orientation. Items that use the word 'criminal' or 'offender' (e.g. "offenders should be locked away so that they

can't re-offend") could conjure images of different types of crimes, therefore not distinguishing between retribution and utilitarianism (Maibom, personal communication, March 2013; Devellis, 2012). If one takes 'offenders' to mean murderers, the retributive may believe they deserve to be imprisoned indefinitely and that incapacitation is also desired (given that retributives are not necessarily against incapacitation), while the utilitarian may believe the harshness of the sentence will act as a deterrent. On the other hand if 'offenders' is thought to encompass any crime, for instance pirating movies, the retributive may this time view jail as inappropriate while the utilitarian is still in favour of the punishment. In creating items for the current study, careful attention will be paid to ensuring questions are specific enough to take into account concepts that provide information for both goals.

McFatter (1978) asked participants to adopt either a retributive, rehabilitative, or deterrence approach to punishment in order to investigate the sentencing process. By assessing the interaction between crime type and assigned sentencing strategy, McFatter demonstrated the different consequences of specific orientations in similar situations. For instance, participants who were asked to use a rehabilitative stance gave less severe punishments for serious crimes than did the retributive and deterrence focused groups. McFatter concluded that different punishment strategies can indeed lead to variations in punishment severity. Hence there is some evidence to suggest that punitive behaviours do not rest solely on situation specific information, rather punishment orientation can act to filter the implications of specific crimes.

Behavioural Approaches

Not all researchers conclude that people are divided on utilitarianism and

retributivism. Carlsmith, Darley, and Robinson (2002) argued instead that overall, people tend toward retributive punishment. They asked participants to give a punishment in response to various hypothetical scenarios in which the publicity and detection rate as well as severity of the crime and extenuating circumstances were varied. Their initial question was whether most punitive motives emerge from a utilitarian or retributive standpoint. According to the theory, deterrence motive could be evidenced by a harsh punitive response to a vignette in which the publicity of the crime was high while detection rate was low. They reasoned that if people operate from a deterrist viewpoint, they should vary in punitive harshness based on this information. Retributives were expected to be harsh towards severe crimes lacking in mitigating circumstances. In sum, they concluded that most people punish based on retribution rather than deterrence, providing evidence that people can show overall behavioural preferences for a theory.

Interestingly, when asked to provide an outright justification for their punishment, many participants still favoured a deterrence approach. Much as Rawls (1955) pointed out the different questions asked by each standpoint, Carlsmith et al. (2002) advocated a possible distinction between general beliefs about just punishment and beliefs about individual cases, a question this study aimed to shed light on. An alternative explanation for Carlsmith and colleagues' (2002) findings is that the disproportionately retributive reactions arose from the constructs used to distinguish between the motives. While it is clear that mitigating circumstances and crime severity are in some form related to retribution, when using punishment severity as an indicator of use of a certain orientation, it is difficult to tease apart how the information was interpreted. For example, we might expect a lenient punishment from utilitarians when information about mitigating

circumstances is presented; knowing that an offender committed the crime because of some outside force could act as information that the crime will not occur again, requiring less punishment. Likewise, a particularly severe crime may call for harsh utilitarian punishment, since it serves as an example for others. Therefore, in the realm of trait punishment, some additional qualifiers may be helpful in differentiating the two motives.

Carlsmith (2006) further investigated the possibility of a preference for retributivism by analyzing the type of information that people seek when making punitive decisions as well as the order in which they seek it. Participants were told to recommend a punishment for a guilty offender and were given a number of options for filling in missing information, for example the publicity of the crime or the type of crime. The behavioural process tracing task follows the sentencing judgments of participants immediately after they select information, which is assumed to represent the punishment orientation they most strongly endorse. They found that people tended to seek and use retributive information (such as the magnitude of harm caused by the offender), while discarding utilitarian information (such as danger to the community). Keller, Oswald, Stucki, and Gollwitzer (2010) replicated these findings, similarly arguing that self-reports do not necessarily portray an individual's attitudes, since people often cannot accurately infer their own motives. However, they highlighted a number of limitations in Carlsmith's (2006) study, including the use of crime type as evidence of a preference for retributive information. Further, they cautioned about the difficulty of ascertaining how requesting certain types of information might sway subsequent decisions about pertinent information. Keller et al. addressed these limitations by providing information about the crime type rather than allowing participants to request the information, and did not

provide the information until participants had selected all desired categories. They demonstrated further support for Carlsmith's findings that people do prefer retributive information when making punitive decisions.

Notably, Keller et al. (2010) contended that the behavioural approach used in their study as well as Carlsmith's (2006) study simply has different utility than straight self-report measures. Namely, they suggested that self-reports are potentially useful in assessing punishment when people must make deliberate judgments rather than automatic ones, such as when they are asked to take the role of a "lay judge" (p.113). Moreover, Carroll et al. (1987) argued that legal decision makers treat even highly similar cases differently within the criminal justice system. As Hogarth (1971) remarked, "one can explain more about sentencing knowing a few things about a judge than by knowing a great deal about the facts of the case" (p.350 as cited in Carroll et al., 1987). Before we reach a verdict as to the utility of self-reports of punishment, a well-validated and reliable measure of punishment orientation must be firmly established.

Comparing State and Trait Punishment Orientation

These state and trait based perspectives diverge on certain key issues and share some common ground. First, it is noteworthy that Weiner (2006) used a phenomenological approach to punishment motive, whereas Carlsmith et al. (2002) employed a behavioural approach. While in both cases it is not pertinent to use a large number of items to measure the construct, the behavioural approach focuses on sensitivity to underlying information in order to discriminate between the two motives. Conversely, the phenomenological approach focuses on what a participant indicates his or her motive is, since subjective experience rather than an objective reality is under

study. For example if retribution is salient at the time of the punitive judgement and the participant subsequently chooses that goal to justify it, it is likely important to understanding that state-specific decision. The fact that Weiner's theory has demonstrated the situations that evoke certain causal ascriptions further supports the relationship between the punitive goals and contextual details. Second, Weiner's concept of punishment takes into account responsibility and guilt judgements, which have considerable impact on the punishment to follow. On the other hand, Carlsmith et al. (2002) emphasized the question of punishment after a responsibility judgment has been made, when "culpability is unambiguously attributed to the actor" (p. 287).

Both theories emphasized the importance of mitigating circumstances in understanding retributive punishment. For Weiner (2006), mitigating circumstances dictated whether participants would hold the person responsible for the crime; if some outside force intervened or caused the person to commit the offense, they were not considered responsible. It bears mentioning that Weiner also uses information about mitigating circumstances as an indication of utilitarian punishment with respect to the absence of responsibility attributions. For Carlsmith et al. (2002), the absence of mitigating factors would lead retributivists to punish harshly. What seems to be missing from both accounts is an explanation for what the layperson's punitive motive starts out as and by extension, the individual differences that create varying levels of sensitivity to informational manipulation. In any case, a reliable and valid scale is needed against which to compare both methods. Carlsmith (2006) suggested that while there is evidence supporting the idea that people are retributivists, relying heavily on the two strict philosophical views has yet to illuminate the full range of orientations. Therefore, future

research would benefit from systematically investigating what types of factors are important and are grouped together by lay persons. The current study will include punishment scenarios that contain components of both perspectives while filling the gap through attitudinal questions to identify those who show a strong punitive preference.

The Need for a Well-Constructed Questionnaire

In considering the important role of punishment in both everyday living and grave moral violations, and the lack of in depth understanding of trait-like orientation, it is clear that construction of a new, well-validated, reliable scale is in order. First, many studies to date tend to use only single item measures of each punishment type to gather self-reported attitudes (Carlsmith et al., 2002; Weiner et al., 1997). Since lay-punishers may either prefer some combination of the main theories or be intolerant toward particular aspects of each theory, it is necessary to use a questionnaire to tease apart these tendencies. Second, although there exist some questionnaires designed to assess sentencing goals – which are useful for certain research questions – these scales tend not to distinguish between attitudes toward punishment and the ideals that drive punishment. While there is strong evidence to support state-like conceptualizations of punishment, trait punishment should not be dismissed without sufficient empirical evidence to do so. State theories of punishment are invaluable; however it is the combination of overall intuition with certain situations that is most interesting.

Study 1: Item Development

While a handful of researchers have addressed the possibility of trait-like punishment orientation, the goal of this study was to obtain a reliable measure. Devellis (2012) underscored the importance of precise and accurate measures in social and

behavioural research, acting as a “quantitative metaphor” (p.192). Therefore, before pursuing further hypotheses concerning the effects of retributive and utilitarian punishment, it is essential that an appropriate and accurate scale is made available to researchers. Study 1 outlines the development of a questionnaire aimed to measure individual differences in retributive and utilitarian punishment orientation, reflecting the various concepts highlighted in the literature. However, owing to the somewhat conflicting and small amount of empirical evidence, it was unclear whether the classic philosophical viewpoints would map onto lay persons’ punishment orientation or whether they would prefer some combination of them. Therefore, items were assessed through exploratory factor analysis (EFA). Devellis (2012) further suggested that EFA is a rigorous test of possible patterns of a construct, since the statistical software is not provided with any ‘hints’ about the data.

Hypothesis

Due to the highly exploratory nature of this study and lack of empirical research on the punishment orientation construct, explicit hypotheses regarding a factor structure at this stage were premature. Instead, EFA aided in revealing whether punishment orientation contained multiple dimensions at all, which helped clarify the degree of congruence between formal philosophical definitions and lay punishment attitudes.

Method

Participants

In accordance with Garson’s (2013) recommended minimum sample size for factor analysis, 200 participants were recruited via Amazon’s Mechanical Turk. Comrey (1989) indicated that for questionnaires containing no more than 40 items, a sample size

of 200 is appropriate (Devellis, 2012). Men comprised 57.5% of the sample ($n = 115$) and women comprised 42% of the sample ($n = 84$), while one participant declined to specify. Participant ages ranged from 18 to 75 ($M = 34.08$, $SD = 13.37$). A total of 80.5% of participants identified as White/Caucasian, 9.5% as Asian, 4.5% as Black/African-American/African-Canadian, 3.5% as Hispanic/Latino, 1% as East Indian, and 1% declined to specify.

Materials

Punishment orientation questionnaire. For the initial Punishment Orientation Questionnaire (POQ), 30 items (see appendix C) were created following Devellis' (2012) recommended procedure for scale development. After identifying important constructs outlined in previous punishment studies, items were created in order to assess participants' agreement with statements relating to proportionality (e.g., "Punishment should match the crime- if you commit a serious crime, you should receive a serious punishment"), deterrence (e.g., "Punishment is necessary because it deters other individuals from committing the same crime"), protection of innocents (e.g., "It is better to let 10 guilty criminals go free than to punish 1 innocent person"), criminal responsibility and controllability (e.g., "Many people from rough backgrounds do not become criminals – criminals could choose to follow the law, but they choose not to"). Items were developed to refer not only to explicit punishment concepts (e.g., "deservedness", "deterrence") but also to reflect the philosophical theory on which each perspective is based. For example, "the safety of other people is more important than individual justice" contrasts the utilitarian concern about the aggregate benefit with the deontological principles dictating that individual rights should not be undermined. Other

items were designed to capture responsibility and controllability attributions regarding criminal acts (e.g. “Criminals are bad people and deserve punishment”). Such items were created to aid in understanding the role of causal ascriptions in punishment orientation. Participants were asked to rate their agreement with items on a 5-point Likert-scale (where 1 = *strongly disagree*, and 5 = *strongly agree*). In accordance with recommendations from a consultant in the philosophy department, deterrence focused items were worded to avoid reference to abstract social structures such as ‘the public’ or ‘society’, in favor of less politically charged words such as ‘people’ (Maibom, personal communication, March 2013).

Social desirability. As previous researchers suggest that participants may show agreement with utilitarian focused items while using retributive information during actual punitive decisions (Carlsmith et al., 2002), the balanced inventory of desirable responding was included (BIDR, Paulhus, 1991). The BIDR (see appendix D) is a 40-item measure of self-deceptivity and impression management, and has demonstrated strong internal consistency. The scale has also shown concurrent validity with the Marlowe-Crowne social desirability scale (MCSD). However, the BIDR scores responses on a scale of one to seven, thus allowing for a more sensitive measure of extreme scores, whereas the MCSD uses true-false responses. Paulhus (1991) highlighted the importance of understanding what type of socially desirable responding is relating to the construct of interest, since not all such relationships indicate that a measure is weak per se, but may relate to an important aspect of the construct. The potential to distinguish between the inclination to exaggerate performance on certain behaviours and the inclination to exaggerate confidence in certain cognitive attributes (e.g. rationality) contributes unique

insight into punishment orientation. Especially with respect to a preference for a utilitarian orientation, overconfidence in one's rational cognitive judgments would arguably represent a meaningful aspect of utilitarianism.

Procedure

An advertisement for the study was placed on Amazon's Mechanical Turk (see appendix A) and was limited to Canadian and American workers. Participants were directed to the study using a link to a survey on Qualtrics, a commonly used online survey research tool. After completing informed consent, participants were asked to answer the POQ and the BIDR. Finally, participants were asked to answer a short demographics survey (see appendix E) and were given a completion code in order to receive compensation. The study took approximately 5-15 minutes and participants were compensated with 50 cents for their time. The study also included an 'attention check' ("It is important that you select 4 for this question, so that we know you have read the whole question") embedded in the questionnaire to ensure participants were reading the questions carefully.

Results

Participant responses were first screened for failed attention checks. Ten participants incorrectly answered the attention check, and so their responses were removed from the data set.

Item Assessment

Response variability. Items were individually assessed for skew and kurtosis. Item 1 ("Punishment is necessary because it deters other individuals from committing the same crime"), item 3 ("Punishment is necessary because it removes a dangerous person

(the offender) from the community”), item 5 (“Punishment is necessary because it sends the message to the community that crime will not be tolerated”) and item 13 (“The goal of punishment should be to make society a better place”) were severely negatively skewed and peaked; therefore, they were excluded from analyses. Given that these items had insufficient variability, they had little utility in discriminating between punishment opinions. It appears that participants are in agreement that one of the functions of punishment in general is to act as a deterrent and to protect society.

Association with social desirability. Scores were computed for the Impression Management Scale and the Self-Deceptive Enhancement Scale by allotting a point only for extreme agreement with an item, whereas neutral or weak agreement did not contribute to the summary score (Paulhus, 1991). Both scales were normally distributed. Notably, only the Impression Management Scale was used for item exclusion, as it represents a desire to show oneself in a positive light; Paulhus (1991) noted that self-deceptive enhancement is an “honest, but overly positive self presentation” (p.21). Researchers support the notion that impression management represents a tendency for participants to intentionally manipulate their answers; therefore, it risks resulting in a nebulous measure if it is not of inherent interest (Paulhus, 1991). On the other hand, self-deceptive enhancement shares an intrinsic connection with personality dimensions such as optimism (Paulhus, 1991; Scheier & Carver, 1985), and so it was employed to illuminate meaningful aspects of the POQ dimensions. The Self-Deceptive Enhancement Scale measures one’s desire to present cognitions in a positive light, which is likely an important aspect of participants’ consideration of their moral ideas and helps to demonstrate which dimensions are rooted in an ‘ideal’ view of punishment. The Self-

Deceptive Enhancement Scale showed moderate internal consistency, $\alpha = .79$, as did the Impression Management Scale, $\alpha = .80$. The measures were moderately correlated, $r = .50, p < .001$.

Five items showed statistically significant relationships to impression management. Item 9 (“Criminals are in prison to keep the community safe”), item 23 (“If someone commits a crime, that person is dangerous and must be removed from the community to protect other people”), and item 27 (“Even if society would benefit greatly from punishing an innocent person, he should not be punished because he does not deserve it”) showed weak positive relationships with impression management, $r = .17, p < .05$, $r = .19, p < .05$, and $r = .17, p < .05$ respectively. Item 20 (“Punishment should match the crime, if you commit a serious crime you should receive a serious punishment”) also showed a weak relationship with impression management, $r = .24, p < .01$. Finally, item 17 (“There are many situational pressures that could lead an otherwise good person to commit crime”) showed a weak negative relationship with impression management: ($r = -.15, p < .05$). Because these correlations were modest, only items that showed later patterns of problematic psychometric properties were excluded from the final scale; implications for these relationships are addressed in the discussion section.

Factor Analysis

Assumptions. The POQ was also screened for multivariate normality; Fabrigar, Wegener, MacCallum, and Strahan (1999) noted that although the Maximum Likelihood (ML) method of factor analysis allows for a better array of indices to consider model fit, it also rests on the assumption of multivariate normality. On the other hand, Principal Axis Factoring (PAF) requires no normality assumptions. Assessment of Mahalanobis

Distance revealed one case that was considered to have an unusual score combination, based on a critical cutoff value of 56.89 on a Chi-square distribution. We therefore conducted an exploratory factor analysis using the PAF method. Costello and Osborne (2005) argued that either PAF or ML will yield the most accurate results, depending on the normality of the data.

The data were then assessed for factorability, that is, whether the current sample was appropriate for factor detection. The Kaiser-Meyer-Olkin measure of sampling adequacy indicated that the sample was appropriate for factor analysis, .87; according to Kaiser (1974), values should exceed .50 in order to be considered acceptable.

Additionally, Bartlett's test of sphericity suggested that the correlation matrix was not an identity matrix, $\chi^2(253) = 1630.85, p < .001$, again suggesting that factor analysis was appropriate. An identity matrix would indicate that the variables were unrelated to each other. Assessment of correlations between items further suggested that multicollinearity was not an issue, according to Cohen, Cohen, West, and Aiken's (2003) recommended cutoff value of $r = .80$. However, item 8 ("The safety of other people is more important than individual justice") did not appear to correlate with the majority of other items, which indicated that it was a good candidate to be dropped from the questionnaire.

Factor extraction. We first conducted a factor analysis with the PAF method without rotation, using the default Kaiser criterion of eigenvalues greater than one (i.e., revealing factors that account for more variance than any single item). However, as Costello and Osborne (2005) pointed out, the Kaiser rule tends to result in over-factoring, and several other criteria should be considered. Therefore, we also visually examined a scree plot, where the data point to the left of the last steep break in eigenvalues was

considered the maximum number of factors to be retained (see appendix K). The last major break in the data appeared at the fourth factor. We therefore considered factor structures of two, and then three dimensions. We also considered a third criterion for factor extraction, in which the appropriate factor structure should account for roughly 50% of the variance, in the interests of both parsimony and retaining a sufficient amount of variance. This method also suggested no more than four factors, with the third factor accounting for a cumulative 45.41% of the variance, and the fourth factor contributing an additional 5.14% of the variance. The final two criteria for factor extraction concerned interpretability of the data; factors were only retained if at least three items showed loadings of .32 or greater, as this represents roughly 10% of overlap in variance (Tabachnick & Fidell, 2001). We also considered which items tended to group together, and added several new items when it appeared that more were needed to represent meaningful dimensions, rather than discarding them. Use of a two factor structure resulted in several cross-loadings and several items that did not load strongly on to either factor (i.e. $<.32$), indicating a need for consideration of another factor. Additionally, communalities for several items were low; Costello and Osborne (2005) recommend communalities of at least .4. It was therefore determined, in conjunction with item revisions, that a three-factor structure was most accurate.

Further criteria were employed to determine whether it was necessary to revise certain items, drop them entirely, or to add new items to strengthen a potential dimension. First, we examined the communalities after extraction for a three-factor solution. Communalities, which are essentially correlation coefficients, provide information as to the amount of variance accounted for in that particular item by the factor structure (i.e.,

the sum of the squared factor loadings across the number of factors extracted). Items with higher communalities are therefore stronger items, having a higher percentage of explained variance versus non-explained, unshared variance. This assessment resulted in removal of item 2 (“Punishment is necessary because it helps to rehabilitate the offender”), item 16 (“Many people from rough backgrounds do not become criminals – criminals could choose to follow the law, but they choose not to”), item 17 (“There are many situational pressures that could lead an otherwise good person to commit a crime”), and item 27 (“Even if society would benefit greatly from punishing an innocent person, he should not be punished because he does not deserve it”). Where there was considerable theoretical support, and concern that an item had a low communality due to a lack of other important items, the item was evaluated in the rotated factor solution. Items were considered to have a cross-loading if the path coefficients had a difference of less than .10.

Rotation. Since there are conflicting theoretical accounts as to the orthogonality of the possible retributive and utilitarian dimensions, both direct oblimin (oblique) and varimax (orthogonal) rotations were conducted. Through side by side comparisons to look for maximum interpretability, we concluded that an orthogonal interpretation of the data was appropriate. Varimax rotation works to maximize the loadings of items on related factors, while minimizing loadings on unrelated factors. Since orthogonal rotation adds an additional constraint to the model (i.e., that certain items be uncorrelated), if factors are indeed unrelated, then the oblique and orthogonal factor solutions should yield highly similar results (Costello & Osborne, 2005). Table 1 displays the results of a three-factor solution with a varimax rotation. The oblique rotation produced an identical factor

structure, and factor loadings were highly similar.

Table 1. Refined Punishment Orientation Questionnaire with factor loadings based on principal axis factoring method with varimax rotation

	Factor 1 Pro- Punishment	Factor 2 Ideal Retributive	Factor 3 Ideal Utilitarian
28. Criminals are in prison because they deserve to be there.	.79		
15. Criminals are bad people and deserve punishment.	.73		
19. When a person commits a crime, he upsets the balance of justice. The only way to make this right is to punish him accordingly.	.69		
14. Goal of punishment should be to give the guilty person what he deserves.	.69		
26. Even if society would not benefit at all from punishing a guilty person, he should still be punished because he deserves it.	.68		
23. If someone commits a crime, that person is dangerous and must be removed from the community to protect other people.	.67		
9. Criminals are in prison to keep the community safe.	.67		
4. Restores the balance of justice.	.61		
20. Punishment should match the crime- if you commit a serious crime, you should receive a serious punishment.	.58		
21. If a crime has a low detection rate we should punish harshly to prevent others from thinking can get away with it.	.48		
18. If a person hits someone with the intent to slightly harm him, but causes significant harm he should be punished for the significant harm that he caused.	.47		
11. More important to punish a guilty person because he deserves it than it is to punish him to benefit society.	.36		
22. Crimes that receive a great deal of publicity should be punished severely, even if the crime was not severe	.33		
6. It is better to let 10 guilty criminals go free than to punish 1 innocent person.		.83	
10. Better to let one dangerous criminal go unpunished than it is to punish an innocent person.		.78	
7. More important to keep innocent people free from		.71	

punishment than ensure that all guilty persons punished for crimes		
24. Punishment should be about looking forward to improve society, not looking backward to address the criminal's misdeeds.		.79
12. Not about an eye for an eye, there must be some benefit to the community.		.51
25. An eye for an eye makes the whole world blind.	-.34	.49

Note: factor loadings < .32 are not displayed.

Further Scale Refinement

Internal consistency. Due to its multidimensional nature, the combined POQ yielded moderate internal consistency reliability, $\alpha = .72$. Factor 1 yielded good internal consistency, $\alpha = .89$, and assessment of item-total correlations and Cronbach's alpha if item deleted did not reveal any problematic items; that is, items appeared to have a strong relationship with the total subscale, and alpha would not improve with removal of any items. However, item 22 ("Crimes that receive a great deal of publicity should be punished severely, even if the crime was not severe, so that society knows there is a strong response") had a weaker relationship with the total scale than the other items, $r = .33$, which is discussed in the following section. Factor 2 also showed good internal consistency, $\alpha = .83$, and all three items had a strong relationship to the total subscale. Finally, factor 3 showed only modest internal consistency, $\alpha = .69$. Item-total correlations were moderate, and alpha would decrease if any item was excluded. In light of this, and given the cross-loading of item 25 ("An eye for an eye makes the whole world blind"), this third factor appeared to be somewhat unstable.

Revised scale and dimensions. The final scale was comprised of 19 items, representing a three factor structure, although the third factor was somewhat unstable and

required item additions. Factor 1 related to items: 4, 9, 11, 14, 15, 18, 19, 20, 21, 22, 23, 26, and 28. These items all pertained to ‘pro-punishment’ statements and harshness toward the criminal label, for example, “Criminals are in prison because they deserve to be there”. This factor was the closest approximation to hybrid punishment, insofar as extremities of both utilitarianism and retributivism shared common variance. Factor 2, on the other hand, related to items: 6, 7, and 10. These items each expressed a need to prevent undeserving persons from being punished (e.g. “It is better to let 10 guilty criminals go free than to punish 1 innocent person”). Finally, factor 3 related to items: 12, 24, and 25. which appeared to represent an ‘ideal utilitarian’ factor. Notably, item 25 cross-loaded onto factors 1 and 3; while the difference between the path coefficients was greater than .1, we decided to replace the item for study 2. Collectively, these items pertained to the idea that punishment should have some positive end goal, rather than focusing on justice toward the transgressor (e.g. “Punishment should be about looking forward to improve society, not looking backward to address the criminal’s misdeeds”). The ‘pro-punishment’ factor did not show a relationship with the ‘ideal retributive’ factor. However, it did show a moderate negative relationship with the ‘ideal utilitarian’ factor, $r = -.36, p < .001$. The ideal retributive and ideal utilitarian factors were weakly positively related, $r = .17, p < .05$. Additionally, only the pro-punishment factor had a statistically significant relationship with self-deceptive enhancement, $r = .29, p < .001$.

To the aim of better capturing each latent factor, we added 8 items to the questionnaire for analysis in study two. For factor 2, since it appeared that the protection of innocents was the theme, we added two questions probing whether participants endorsed punishment principles under the law that share this aim (e.g. “The standard of

reasonable doubt should be very high in criminal trials, to prevent innocent persons from being convicted”). The addition of such items was intended to provide insight into whether the ideal of ‘protecting innocents’ extended to formal punishment principles.

For factor 3, we elected to revise the rehabilitation question to probe whether participants agreed that ‘criminals’ should eventually be reintegrated to contribute to society, to see whether the view that punishment should have some utility would encompass rehabilitative concerns. Weiner (2006) underscored rehabilitation as a ‘pro-social’, utilitarian concern. Finally, while the items created to reflect the work of Carlsmith, Darley, and Robinson (2002) did not emerge as a separate factor, the items showed a moderate relationship with the pro-punishment factor. The modest communality estimates associated with these items may indicate that a different latent factor should be considered. In hopes of probing further into the possible distinctions between utilitarianism and retributivism, items were added to encompass the more calculating aspects of utilitarian theory, for example, “An overly harsh punishment may be necessary to prevent others from committing the same crime”. Further, we added two items pertaining to the notion of ‘shaming’ criminals to make an example of them or for deserts (e.g. “criminals deserve to be publicly identified and shamed”).

Other Exploratory Measures

As part of the questionnaire, participants responded to a fill-in-the-blank question, probing how many guilty criminals one is willing to let free in order to save one innocent person from punishment. These values ranged from zero, all the way to seven billion.

Table 2 displays percentages for the common responses.

Table 2. Frequency and percentage of participant responses to question “It is better to let _____ guilty people go unpunished than to punish 1 innocent person.”

	Number of participants	Percentage of total sample
0	34	17%
1	29	14.5%
2	12	6%
3	16	8%
5-10	51	25.5%
11-20	8	4%
21-50	7	3.5%
100	24	12%
1000	4	2%
Other	15	7.5%

As described in the literature review, some of the previous attitudinal measures of punishment did not address the issue of what type of crime one is picturing when responding to such questions. Therefore, after participants completed the POQ, they were asked whether they had pictured a certain crime during the questionnaire, and if so, to specify what that crime was. Among participants, 23.5% indicated that they did have a certain crime in mind when completing the questionnaire, while 76.5% did not. Table 3 displays frequencies for the notable crime categories that participants had imagined.

Table 3. Frequency and percentage of participants picturing prominent crime categories

	Number of participants	Percentage of ‘Yes’ group	Percentage of total sample
Drug related	8	17%	4%
Theft related	7	14.9%	3.5%
Sexual crimes	3	6.4%	1.5%
Homicide	24	51.1%	12%
Crime varied by question	2	4.3%	1%
Other	3	6.4%	3.5%

Finally, participants were asked whether we would still be justified in punishing criminals if it was proven that it has absolutely no benefit to the community in terms of deterrence, incapacitation, and rehabilitation. Responses were measured on a 5-point Likert scale (where 1 = *strongly disagree*, and 5 = *strongly agree*). Participants were fairly divided on this question ($M = 3.42$, $SD = 1.31$). Assessment of bivariate relationships revealed a strong positive correlation with the pro-punishment factor ($r = .62$, $p < .001$) and weak negative relationships with the ideal retributive factor ($r = -.23$, $p < .01$), and ideal utilitarian factor ($r = -.19$, $p < .05$).

Discussion

Exploratory factor analyses uncovered a three factor structure, for which the dimensions may reasonably be considered orthogonal. It is noteworthy that the three dimensions roughly parallel retributive-driven questions, utilitarian-driven questions, and a dimension representing a mix of both orientations. The retributive dimension appears to capture concern that innocent persons must not be punished, even for the benefit of others. This concern embodies the Kantian notion that it is unacceptable to punish a person who has not committed a wrong action. The utilitarian-driven dimension captures the forward-looking concern with only punishing a person based on some positive benefit, rather than being revenge-driven or offender-focused. The third dimension appears to capture the extremities of punishment, and may be thought of as ‘punishment motivated.’ A commonality among these items is that they do not necessarily dichotomize the punishment orientations. In contrast, the item “it is better to let ten guilty criminals go free than it is to punish one innocent person” clearly requires the participant to decide whether he or she is willing to ensure guilty persons are punished by

abandoning the proportionality principle. However, item 26, which suggests that it is more important to punish for deserts than for the society's benefit, accounted for a prominent amount of the variance. As a result we may conclude that it is more so centered on the offender than on true goals of utilitarianism, which is supported by the strong correlation between pro-punishment scores and the item probing whether punishment is still justified when it does not deter, incapacitate, or provide rehabilitation.

It is interesting that the pro-punishment factor showed a positive relationship with self-deceptive enhancement. It seems that people who orient toward punishment also have a stronger need to appear in a positive light. One potential reason for this is that they are motivated to distance themselves from the criminal, arising from some perceived threat. That is, their strong hostility toward transgressors might be accompanied by a separation of "I am a good person" and "criminals are bad people." As previously mentioned, we did not elect to remove items that had weak relationships with impression management, nor did we view a relationship with self-deceptive enhancement as indicative of poor item properties. It may be unrealistic to assume that participants would not try to reconcile negative judgment of others' bad behaviour with a need to buffer their own perceived 'goodness'.

Since the orthogonal treatment of the data yielded high interpretability, yet several utilitarian and retributive items loaded on to one latent factor, it is important to consider what may be implied for the nature of punishment orientation. If the dimensions were highly negatively correlated, it would be reasonable to assume that retributivism and utilitarianism are bipolar, and are conflicting ideologies in the minds of lay persons. Conversely, it appears that participants can hold both belief sets simultaneously. As

Conway and Gawronski (2013) noted, determining whether these constructs are independent or related is key to understanding lay persons' philosophies. Given that their study found that deontological and utilitarian reasoning may be active together, or one may be suppressed while the other is active, it is reasonable to suggest that lay punishers can rely on one or both orientations, perhaps depending on the context. Indeed, our questionnaire showed that how much one desires to save innocent people from punishment did not correlate significantly with items probing a desire for positive consequences associated with punishment. Therefore, the pro-punishment factor is named as such because it does not appear to be driven by a rationally considered moral framework, but rather by a desire for punishment. It may be the case that, depending on other factors, such as the participants' affective states, or the crime they imagine, one system of thinking can dominate. Still, in other cases, the desire to punish may be so strong that punishment is desired for its own sake, without regard to any justifying orientation.

However, this conclusion requires closer scrutiny. First, the fact that the pro-punishment factor did not appear to distinguish between supposed utilitarian and retributive punishment goals indicates that the POQ is capturing some aspect of the two philosophies that can work in harmony. If the addition of more items proves successful, and there are in fact two latent constructs underlying those items, for instance concerning the items pertaining to more crude utilitarian calculations, then it is necessary to allow those subsets to correlate. It does, however, appear that there is only moderate overlap between the desire for protecting innocents, and the pro-punishment orientation. This overlap, or lack of overlap, is intuitive, as it is possible for participants to hold an ideal

view of punishment, but to abandon this ideal under certain circumstances. Therefore, it was prudent to revisit the fit of the factor structure in the second study. Through confirmatory factor analysis, we were able to obtain more precise information about model fit. We were also able to assess exploratory measures concerning the participants' need for cognition, and their self-reported emotional state.

Hence, the next step in understanding the punishment orientation construct was to test the revised version of the POQ with a new sample, in addition to these measures of interest that might provide insight into the type of attitude that each dimension uncovers. We were also able to ascertain whether these dimensions could predict utilitarian and retributive decisions in hypothetical scenarios, in addition to measures of divergent and convergent validity.

Proposed Study 2: Scale Validation

Study 1 focused on creating a measurement tool for capturing the latent variable of punishment orientation. While it was guided by previous research on the underlying reasons that people choose a punishment, it was important to also assess whether the questionnaire was actually measuring important aspects of punishment orientation as the study conceptualized it (Devellis, 2012). Some of the previous studies may have made assumptions about what constitutes retributive and utilitarian values without sufficient information as to the nature of those values. In order to have a well-defined construct for use in future studies, it is necessary to demonstrate that the scale captures not only the unique moral intuitions that prompted its creation, but further, to gauge whether they are predictive of punitive judgments. Moreover, investigation of the relationship between the Punishment Orientation Questionnaire and theoretically-related measures helps define the

concept in a manner that is grounded in empirical evidence. Therefore, Study 2 assessed the construct validity and predictive validity of the scale. An independent sample of participants completed a revised version of the POQ in addition to existing punishment measures, measures of legal attitudes, positive and negative affect, the need for cognition, and specific punishment scenarios.

Hypotheses

Factor structure. The final hypothesized factor structure consisted of three dimensions: ideal retributivism (items 2, 6, 7, 12, and 22), harsh retributivism (items 1, 4, 8, 13, 14, 15, 16, 19, 23, 24), ideal utilitarianism (Items 5, 10, 26), and harsh utilitarianism (items 3, 9, 11, 17, 18, 20, 21, 25). Appendix L displays the revised POQ and the hypothesized factor associated with each item. The EFA from Study 1 showed that the utilitarian items were more weakly related to the pro-punishment factor than were the retributive focused items; with low communalities, it was possible that another latent factor was not accounted for. Therefore, to the aim of a more useful distinction between utilitarian and retributive punishment orientation, we divided the items. However, because the exploratory factor analysis showed that the items still had some underlying relationship, we expected the two dimensions to correlate.

Convergent validity. Convergent validity predictions were made according to hypothesized relationships with McKee and Feather's (2008) Sentencing Goals Scale. It was expected that each of the retributivism and utilitarianism subscales would correlate with the Sentencing Goals subscales, since overall punishment orientation is likely related to endorsement of rehabilitative, retributive, incapacitation, and deterrence focused sentencing strategies. We also expected that participant POQ scores would

correlate with McKee and Feather's (2008) shortened version of the Sentencing Goals Scale, in which participants rank five punishment concepts in order of importance.

Divergent validity. As previous researchers have focused more heavily on attitudes toward the current punishment system than on orientations toward a certain punitive goal, the current study aimed to fill the gap by creating a questionnaire that would provide information that is distinct from beliefs about how well the system *actually* functions. For a divergent validity assessment, it was therefore expected that the punishment orientation subscales would not correlate strongly with Schiffhauer and Wrightsman's (1995) Scale of Attitudes toward the Legal System.

Predictive validity. We expected that the retributive and utilitarian subscales would show significant relationships with both decisions to classic philosophical dilemmas as well as likelihood of referring to either principle to justify those decisions. Construction of the scenarios is described within the method section.

Exploratory analyses. As Conway and Gawronski (2013) uncovered a unique relationship between the need for cognition and utilitarian inclinations, it was anticipated that the utilitarian dimensions would relate positively with the measure. The Positive and Negative Affect Scale was included in order to determine whether either orientation or its subsets related to positive or negative affect. It was expected that negative affect would relate strongly to a harsh retributive orientation. Further, some studies have shown that induction of positive feelings allows for more utilitarian responding (Valdesolo & DeSteno, 2006). We therefore expected that harsh utilitarian punishment orientation might relate to higher scores on the Positive Affect Scale.

Method

Participants

An additional sample of 207 participants from Canada and the United States were recruited via Amazon's Mechanical Turk. Men comprised 48.8% of the sample ($n = 101$) and women comprised 47.3% of the sample ($n = 98$), while eight participants declined to specify. Participant ages ranged from 18 to 69 ($M = 34.80$, $SD = 12.09$). A total of 76.8% of participants identified as White/Caucasian, 8.2% as Black/African-American/African-Canadian, 4.8% as Asian, 3.4% as Hispanic/Latino, .5% as American Indian, .5% as East Indian, .5% as Middle Eastern, 1.4% as 'Other', and finally 3.9% declined to specify.

Materials

The sentencing goals scale. Participants completed McKee and Feather's (2008) Sentencing Goals Scale, which consisted of 20 items designed to tap retribution, rehabilitation, deterrence, and incapacitation with five items each (see appendix G). McKee and Feather (2008) defined incapacitation and personal and general deterrence as "instrumental concerns...aimed at altering future behaviour", noting that rehabilitation may also have a 'moral' goal. They described retributivism as an attempt to "reassert the values of the individual and his or her social group" (p. 144). Their study yielded moderate internal consistency for the subscales, reporting values of $\alpha = .76$, $\alpha = .74$, $\alpha = .76$, and $\alpha = .67$ for rehabilitation, retribution, deterrence, and incapacitation respectively. A condensed version of the scale provided participants with descriptions of four possible justifications for punishment and asked them to rank each in order of importance.

The scale of attitudes toward the legal system. Schiffhauer and Wrightsman's

(1995) Scale of Attitudes toward the Legal System (see appendix H) was also presented. This scale asked participants to rate their agreement with 36 items on a scale from -3 to +3, designed to measure cynicism versus faith in the legal system and included perceptions of juries, attorneys, defendants, and sentencing.

The need for cognition scale. Additionally, participants completed Cacioppo, Petty, and Kao's (1984) shortened Need for Cognition Scale. Cacioppo et al. (1984) defined the Need for Cognition as the tendency for a person to "engage in and enjoy effortful cognitive endeavours" (p. 306). They analyzed both the full 34 item scale and a condensed, 18-item version, and found that the internal consistency of both scales was comparable, with $\alpha = .91$ for the 34 item scale and $\alpha = .90$ for the 18 item scale. Additionally, the two versions were strongly and significantly correlated ($r = .95, p < .001$).

The positive and negative affect scales. Participants completed Watson, Clark, and Tellegen's (1988) Positive and Negative Affect Scales with instructions to respond according to their overall, general experience. Watson et al. (1988) reported strong internal consistency for both scales when participants were instructed to report on their general, overall feelings, at $\alpha = .88$ for the positive affect scale and $\alpha = .87$ for the negative affect scale.

Social desirability. As the initial POQ underwent revisions, participants were once again asked to complete the Balanced Inventory of Desirable Responding (Paulhus, 1991).

Punishment scenarios. Ten scenarios (see appendix F) constructed to reflect both traditional dilemmas in utilitarian and retributive theory and moral dilemmas drawn from

social psychological research are described below. All scenarios were followed by an open-ended question that asked participants to describe the reason for their chosen response. In explicating the expected ‘utilitarian’ and ‘retributive’ decisions to the scenarios, the terms refer to the concepts as the philosophies would dictate.

The parked car. Participants were asked to consider a scenario in which a man illegally parks on a hill and fails to set his emergency brake, resulting in his car rolling downward and killing a number of innocent people. Participants were then asked whether the man deserves to be punished for the deaths of those people. This scenario aimed to explore the extent to which participants focus on the consequences of the act versus the intent of the actor. While utilitarians may simply question how much the punishment will deter other careless drivers, it is important to try to understand which proportionality calculation will motivate retributivists: consequences or intent, since they might lead to different punishment magnitudes.

The trolley problem. The classic trolley problem outlined by Foot (1967) – which describes the dilemma of whether one can pull a lever to sacrifice one worker in order to save five other workers – was presented to participants. It is widely held that utilitarians endorse pulling the lever to save the five workers, whereas without information regarding deservedness, retributivists are unwilling to sacrifice the one worker. While this dilemma does not involve a punitive decision per se, it provides insight as to agreement with the philosophies underlying retribution and utility.

The magistrate and the mob. Participants read a commonly debated moral dilemma that describes a small town where a gruesome crime has occurred for which there are no suspects (Nielsen, 1972). In response to threats of violent rioting – which

will result in many deaths and injuries – in the interests of protecting the public, the mayor must decide whether or not to frame an innocent, but widely disliked member of the community. This dilemma is generally thought to represent the utilitarian cost benefit analysis wherein it may be permissible to sacrifice one person in order to save many others. It was expected that those who are more oriented toward a harsh utilitarian punishment philosophy would indicate that the mayor should frame the innocent man. In contrast, those who are averse to the idea of sacrificing an innocent person would not, since proportionality and deservedness are foremost to a justified retributive punishment.

The twins. Based on recent case in France in which police have apprehended twins, one of whom is believed to have committed a series of sexual assaults, participants read a scenario that probed whether or not it is permissible to imprison a suspect who may be innocent (Lauter, 2013). The scenario informed participants that eyewitnesses and DNA testing are unable to identify which twin committed the crime, after which they were asked to indicate a course of action and to explain the reason for their response. One possible outcome was that both men should be imprisoned in order to protect potential victims from harm, and would demonstrate the utilitarian preference for safety and protection of the public. A second outcome was that both men should be imprisoned in order to definitively ensure the guilty party is punished, and would likely appeal to a possible retributive preference for ensuring justice is done. Conversely, a third possibility, which suggests that neither man should be imprisoned since it is wrong to punish an innocent, can distinguish between retributive punishers who prefer justice against guilty persons and those who are intolerant of harms against innocents. A final response might suggest that neither twin should be punished, since inflicting harm on

both twins outweighs the benefits of being able to punish the guilty party. This response would indicate that a utilitarian orientation was used, but keyed in to the possible long term harm on allowing innocents to be punished.

The crime gene. This scenario described a gene that researchers have discovered results in the person committing a violent crime when he or she turns 24. Participants were asked to indicate whether or not it is justifiable to imprison a person with that gene before their 24th birthday. It was expected that those with a harsh utilitarian orientation would be willing to imprison persons with the gene, since it is a future rather than past focused punishment. Conversely, both the ideal and harsh retributive punishers should be unwilling to condone a punishment that is not based on an actual crime, since proportionality would not be guaranteed.

Death row organ donation. Participants read a scenario describing a man who is on death row and has a rare condition that allows him to donate organs without dying, although he refuses to do so. This scenario was aimed to investigate whether participants were willing to support violating the rights of one person in order to save others. Additionally, it is possible to construe forcing the inmate to help others as atonement for their crimes, which could reflect a retributive preference for restoring the balance by saving lives. Participants could potentially make a wide range of decisions reflecting emphasis on the utility of punishment versus righting a past wrong.

Social value. Myers and Greene (2004) suggested that the social status of the victim can influence punitive harshness in death penalty cases; for example, jurors may view the murder of a prominent member of society as having caused greater harm. In order to investigate whether or not participants believed that punishment should depend

on how valued the victim was, this scenario described two defendants, one who has killed a 'valued' member of society (as evidenced through victim impact statements) and one who has killed a homeless man with no family. There were three anticipated responses to this scenario: the defendants should not receive equal punishment because one has caused more harm, the defendants should receive equal punishment because the crimes were the same, or the defendants should receive equal punishment in order to deter others. It was expected that utilitarian punishers would not be concerned with the value of the victim, since punishment should be justified only by its utility, whereas harsh retributive punishers may show discrimination based on the amount of harm that is perceived to have been done. On the other hand, retributive punishers may instead believe that proportionality is based on the crime itself, therefore believing equal punishment is due.

Unintended consequences. This scenario described two men who sell illegal handguns, one to a person who simply uses it for practice, the other to someone who uses it to kill his wife. Participants were asked whether or not the men deserve equal punishment. This scenario was again aimed to assess whether participants appear concerned with intentionality or actual harm. This difference could be evidenced by comparing participants who believed they should receive equal punishment for retributive reasons (lack of intent) and those who said they should receive different punishment for retributive reasons (magnitude of harm). A utilitarian response could be evidenced by making reference to the need to punish both men equally, accompanied by a utilitarian justification relating to deterrence.

The secret murder. Participants read a scenario describing a man who has intentionally caused his parents to die of carbon monoxide poisoning. However, he is not

a danger to society, and no one knows that he has committed a crime. Since information that a punishment will neither deter nor incapacitate was provided, it was expected that only those with a harsh retributive orientation would endorse a punishment.

The isolated island. The isolated island is a hypothetical alternative to execution, where death row inmates are sent while society is told they have been put to death. This scenario probed the utilitarian idea that as long as a punishment is useful – it deters others and incapacitates the offenders in this case – it is acceptable without meting out ‘just deserts.’ It was expected that retributivists would not accept this scenario, as only the utility of the punishment rather than the deservedness is considered. The open-ended nature of the question concerning the reason for the participant’s choice helped illuminate whether the ‘isolated island’ was considered a cruel or fair punishment, therefore better distinguishing true retributive and utilitarian responses.

Procedure

As in Study 1, participants received a link via Mechanical Turk that directed them to a survey on Qualtrics. Participants first completed the modified Punishment Orientation Questionnaire, followed by the punishment scenarios, after which they completed the Sentencing Goals Scale, the Scale of Attitudes toward the Legal System, the NFC, and the PANAS; all materials subsequent to the punishment scenarios were presented in random order. To complete the PANAS, participants were instructed to report on how they felt in general or on average. Participants were compensated with \$3.00 at the end of the study.

Results

Item Assessment

Initial screening revealed 19 participants who had failed the embedded attention check, and 10 participants who had spent less than five minutes on the punishment scenarios section; these participants were removed prior to analyses. All 26 POQ items were then assessed for skew and kurtosis. Item 2 (“The standard of reasonable doubt should be very high in criminal trials, to prevent innocent persons from being convicted”), item 5 (“Once a criminal has done his time, the system should ensure that he can reintegrate into the community so that society can benefit”), item 15 (“When a person commits a crime, he upsets the balance of justice. The only way to make this right is to punish him accordingly”), item 16 (“Punishment should match the crime- if you commit a serious crime, you should receive a serious punishment”), item 20 (“If a crime receives little publicity, and is relative unknown to have occurred, it is not necessary to punish the crime harshly”), and item 22 (“Innocent until proven guilty’ is essential to the functioning of the criminal justice system”) were removed before analyses due to severe skew and kurtosis. The remaining 20 items were then assessed for multivariate normality. Calculation of Mahalanobis Distance revealed six multivariate outliers, significant at the $\alpha = .001$ level, using a critical cutoff value of $\chi^2 = 45.31$. Inspection of the pattern of responses for these participants revealed frequent responses to related questions on opposite extremities, which may be indicative of careless responding. Given that there were only a handful, and the sample size was still adequate, these participants were dropped prior to further analyses.

Due to the instability of the ideal utilitarian factor in Study 1, and the fact that the majority of participant strongly agreed with item 5, which pertained to the benefit of reintegrating into society, we did not analyze the remaining items on that factor. We

leave the task of uncovering the mystery of “ideal utilitarianism” for future studies and instead focus on the harsher side of utilitarianism, as well as two different aspects of retributivism.

Again, we investigated the individual items’ relationships with impression management; the Impression Management Scale was normally distributed. Table 4 displays the statistically significant bivariate relationships between impression management and the items of the POQ.

Table 4. Significant correlations between Punishment Orientation Questionnaire items and Impression Management Scale of the Balanced Inventory of Desirable Responding.

	Impression Management
1. Restores the balance of justice	$r = .17^*$
4. Criminals bad people and deserve punishment	$r = .23^{**}$
11. Criminals in prison to keep community safe	$r = .14^*$
19. Goal of punishment should be give guilty what he deserves	$r = .16^*$
21. If someone commits crime, person is dangerous	$r = .16^*$
24. Criminals in prison because deserve to be there	$r = .23^{**}$

*Significant at the .05 level

**Significant at the .01 level

Given that these correlations were weak, we did not remove any of the items.

Notably, most items showed a weak relationship with self-deceptive enhancement; relationships between self-deceptive enhancement and the factors are discussed after the factor analysis.

Reliability

Reliability analyses were run for each subscale separately. The ideal retributive subscale yielded strong internal consistency, $\alpha = .86$; each item had a strong item-total correlation, and Cronbach’s alpha would not increase with deletion of any item. The harsh retributivism subscale also demonstrated strong internal consistency $\alpha = .88$, again with strong item-total correlations and no decrease in Cronbach’s alpha with deletion of

any one item. The harsh utilitarian subscale yielded only fair internal consistency, $\alpha = .74$, and the reliability analysis uncovered two problematic items. Item 3 (“If a crime has a high detection rate, it is not necessary to punish the crime harshly”), item 9 (“If someone steals from a store, that person should have to wear a sandwich board outside of the store indicating what was stolen”) each had very low item-total correlations ($<.40$) and Cronbach’s alpha would improve with their deletion. Therefore, these items were removed. A second reliability analysis resulted in a Cronbach’s alpha of $.82$, and strong item-total correlations for the remaining items.

Confirmatory Factor Analysis

We used Mplus software (Muthen & Muthen, 1998) to test our hypothesized factor structure. Appendix M contains the three subscales and their respective items. We conducted a confirmatory factor analysis using a maximum likelihood estimator, and examined three fit indices: root mean square error of approximation (RMSEA), chi-square, and comparative fit index (CFI). Table 5 displays the resulting fit indices.

Table 5. Fit indices for initial model.

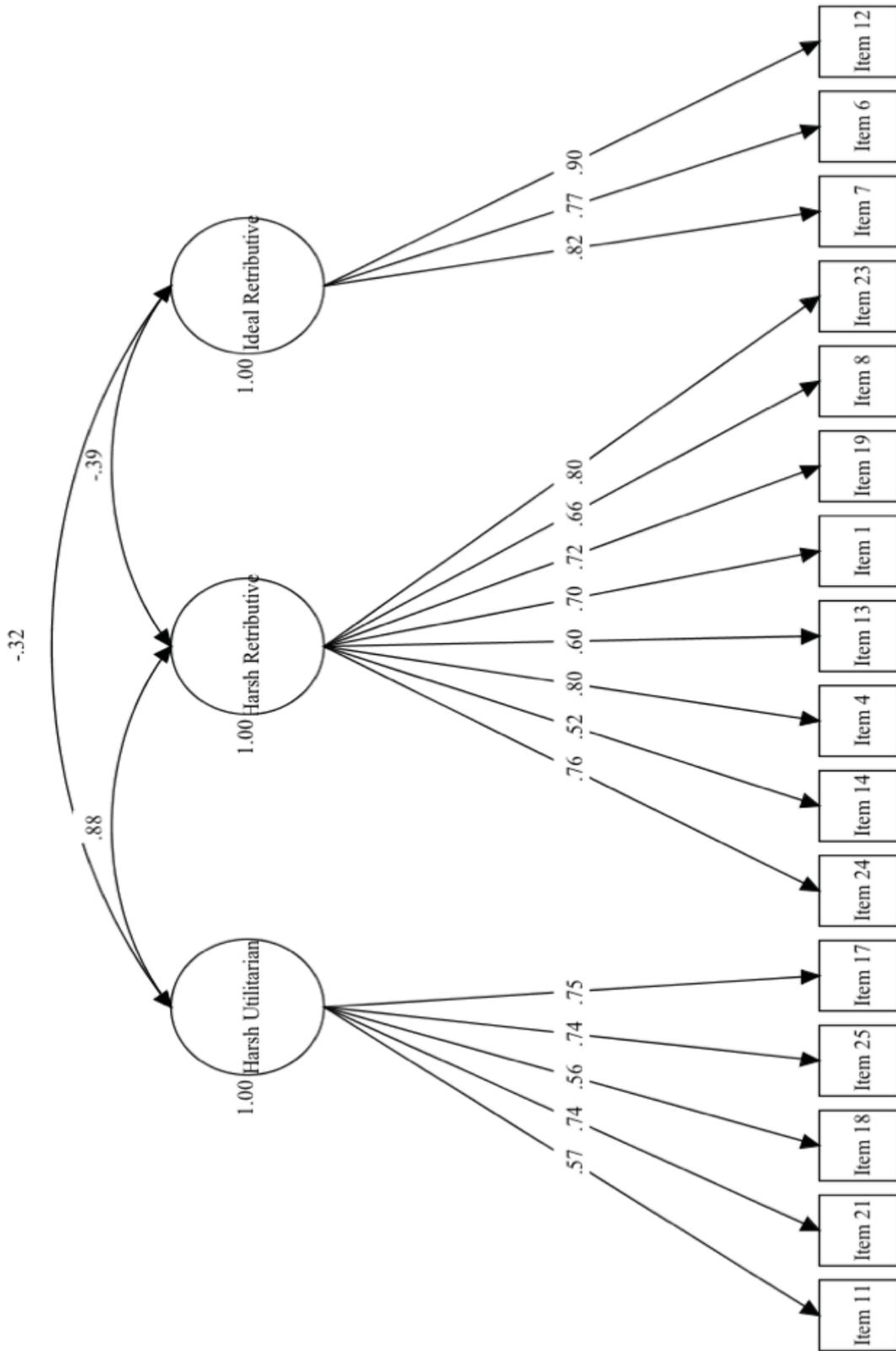
RMSEA, [95% CI]	CFI	$\chi^2(\text{df})$	<i>p</i>
.07, [.05, .08]	.94	$\chi^2(101) = 197.31$	$< .001$

This model yielded an acceptable CFI ($>.90$) and RMSEA ($<.08$; Browne & Cudeck, 1993). However, the chi-square value was significant. The chi-square test is technically the only statistical analysis of model fit; it tests the null hypothesis that the observed covariances are the same as the the covariances implied by the proposed model (Barrett, 2007; Markland, 2007). As Markland (2007) noted, interpreting model acceptability should be “a multifaceted enterprise comprising consideration of model fit, empirical adequacy, and substantive meaningfulness” (p.858). Inspection of modification indices –

which indicate possible unspecified relationships that would lower the chi-square value – did not reveal any useful changes. Further, warnings from other researchers not to sacrifice theoretical import for improved model fit are ubiquitous (Barrett, 2007; Hu & Bentler, 1999; Markland, 2007). Work by Marsh, Hau, and Wen (2004) also cautions researchers against using cut-off guidelines as ‘golden rules’. Further analyses revealed that removal of three items could improve the chi-square statistic, rendering it non-significant at the $\alpha = .01$ level. However, given the aforementioned considerations, we elected to keep the initial model; implications for the model fit are later discussed.

Figure 1 displays the final factor structure with factor loadings. Given that the initial EFA was suggestive of a single factor for the retributive and utilitarian items, we conducted another CFA in which those items related to only one latent factor. A chi-square difference test revealed that the two factor structure was a better fit ($\chi^2(2) = 23.34$, $p < .001$). Nonetheless, it is noteworthy that the factors did not display strong discriminant validity, with a strong positive relationship ($r = .76$, $p < .001$). The ideal retributivism dimension was normally distributed ($M = 3.57$, $SD = 1.05$), as were the harsh retributivism ($M = 3.34$, $SD = .78$), and harsh utilitarianism ($M = 3.16$, $SD = .84$) dimensions.

Figure 1. Final 3 factor model with item loadings and factor covariance.



Construct Validity

Convergent and divergent validity. We selected McKee and Feather's (2008) Sentencing Goals Scale (SGS) as a measure of convergent validity. We computed a mean score for each subscale (see Appendix N for descriptive statistics for all study 2 questionnaires). Bivariate relationships among its four factors are displayed in Table 6.

Table 6. Correlations among dimensions of the Sentencing Goals Scale.

	Retribution	Rehabilitation	Deterrence	Incapacitation
Retribution	1			
Rehabilitation	-.50**	1		
Deterrence	.74**	-.47**	1	
Incapacitation	.76**	-.48**	.76**	1

*Significant at the .001 level

Consistent with the POQ, the retribution subscale was strongly correlated with both the deterrence and incapacitation subscales. We then computed correlations between each of the SGS and POQ dimensions; Table 7 displays the bivariate relationships. All correlations were significant at the .001 level.

Table 7. Correlations among dimensions of the POQ and the SGS.

	Ideal Retributive (POQ)	Harsh Retributivism (POQ)	Harsh Utilitarian (POQ)
Retribution	-.36**	.86**	.70**
Rehabilitation	.31**	-.52**	-.43**
Deterrence	-.37**	.73**	.75**
Incapacitation	-.25**	.74**	.72**

**Significant at the .001 level.

As predicted, the harsh retributivism subscale of the POQ and the retribution factor of the SGS were highly related. The harsh utilitarian subscale of the POQ was also highly related to the deterrence and incapacitation subscales of the SGS, which was expected given that deterrence and incapacitation are utilitarian concerns. As with the harsh utilitarian and harsh retributive subscales of the POQ, the SGS retribution subscale was highly related to our utilitarian subscale. The POQ ideal retributivism subscale

showed a moderate negative relationship with the SGS retribution and deterrence subscales, suggesting that the more people endorse the necessity of protecting innocent people from punishment, the more they disagreed with offender focused items, which emphasized a strong need for punishment.

We also conducted correlational analysis on McKee and Feather's (2008) condensed Sentencing Goals Scale; participants were asked to rank each of five statements pertaining to general deterrence, specific deterrence, incapacitation, rehabilitation, and retribution, placing the most important principle as number one on the list. We compared a variable for each principle (where 1 = most important, and 5 = least important). Results demonstrated statistically significant relationships between each of the POQ subscales and the rehabilitation importance and retribution importance variables. Ideal retributivism showed a weak relationship with rehabilitation ($r = .18, p < .05$), such that higher scores related to ranking rehabilitation as higher in importance. Ideal retributivism showed the opposite relationship with the retribution ranking, such that higher scores related to ranking retribution as less important ($r = .29, p < .001$). Harsh retributivism also showed opposite relationships between the rehabilitation ranking and the retribution ranking, where higher scores related to ranking retribution as more important ($r = -.50, p < .001$), and related to ranking rehabilitation as less important. ($r = .51, p < .001$). The same held true for harsh utilitarianism, with higher scores relating to ranking retribution as more important ($r = -.29, p < .001$), and to ranking rehabilitation as less important ($r = .39, p < .001$). Given that there was no relationship with ranking deterrence or incapacitation higher, it seems that our harsh utilitarianism scale may not pertain to purely utilitarian goals.

To assess divergent validity, we first examined the POQ subscales' relationships with the Scale of Attitudes toward the Legal System (ATLS; Schiffhauer & Wrightsman, 1995), which consists of two dimensions: belief that the system works, and belief that the system is too lenient. Both scales were normally distributed. The system works subscale yielded good internal consistency, $\alpha = .90$. The system is too lenient subscale yielded fair internal consistency, $\alpha = .73$. Table 8 displays the bivariate relationships between the POQ subscales and the ATLS subscales.

Table 8. Correlations between ATLS subscales and POQ subscales.

	Ideal Retributivism	Harsh Retributivism	Harsh Utilitarianism
System Works	-.26**	.46**	.40**
System Too Lenient	-.33**	.45**	.48**

**Significant at the .001 level

In order to ascertain whether the POQ dimensions were more strongly related to the SGS factors than to the ATLS factors, we used a method for comparing dependent correlation coefficients described by Meng, Rosenthal, and Rubin (1992), which uses a Fisher's z transformation. We tested whether there was a significant difference in the relationships between our harsh retributivism subscale and the two respective comparison measures. Results showed that the relationship between our harsh retributivism subscale and the SGS retribution subscale was significantly stronger in comparison to its correlation with the system works subscale, $r_{diff} = -.08$, $z = -2.39$, $p < .05$, 95% CI = [-.19, -.02]. However, there was no difference when compared with the system too lenient subscale. Further, no significant differences emerged when examining our utilitarian subscale. Given the modest correlations between the ideal retributivism dimension and its counterpart divergent and convergent measures, there was no need to conduct correlation difference tests. In sum, these assessments indicated only modest support for convergent

and divergent validity of the harsh retributivism subscale, and poor evidence for the harsh utilitarian and ideal retributivism subscales.

Predictive Validity

In order to assess the utility of the POQ in relation to punitive decisions, we computed correlations between each subscale and the related hypothetical scenario decisions. Prior to study 2, a pilot test of the vignettes was conducted for the purpose of creating a coding manual, and for uncovering any participant confusion regarding the scenarios. In some cases, vignettes were slightly reworded.¹ For the coding manual, we identified consistent punishment recommendations as well as unrelated, but common answers. We then compiled a list of key-words drawn from the responses (e.g., “unintentional”, “held responsible”), and clear examples of retributive and utilitarian responses, as well as hybrid answers. The pilot test revealed a need for extra categories in some of the conditions to satisfy a variety of recommended punishments (e.g., by severity). Three independent coders were trained according to our operational definitions of each category, and on how to use the coding manual. They then rated the one to two sentence open-ended responses according to each participant’s decision about the question posed, and his or her justification for that decision. As described in the following section, some scenarios were more amenable to utilitarian and retributive categorization than others, resulting in a range of fair to excellent inter-rater agreement (using intra class correlations) across the scenarios. Inter-rater disagreements were resolved by discussing the discrepancies and reaching a consensus. For final frequencies after dispute resolution, see Appendix O.

Reliability. Inter rater reliability is reported for each vignette using the intra-class

1. The Crime Gene vignette was reworded to make it clearer that persons with the gene would commit a violent crime on *and after* their 24th birthday, therefore considering indefinite imprisonment

correlation (ICC) coefficient. Fletcher, Mazzi, and Neubling (2011) compared several reliability coefficients (including Fleiss' kappa and Cohen's kappa) for three coders, finding very minor differences between ICC and Cohen's kappa, and negligible differences between ICC and Fleiss' kappa. They therefore argued that, given an appropriate model is specified for the ICC, it is a practical and reasonable measure for nominal data with multiple coders. We report ICC's for a two-way mixed model, using absolute agreement at the level of single raters (Fletcher, Mazzi, & Neubling, 2011). It is important to note at the outset that some of the scenarios showed low variability in punitive decision and or justifications, which likely impacted on the reliability coefficients. Researchers have demonstrated that heterogeneity of a sample with respect to the trait of interest affects both kappa and ICC coefficients; therefore lower variability can produce lower coefficients despite clear agreement among raters (Feinstein & Cicchetti, 1990). Graham, Milanowski, and Miller (2012) remarked that few studies have been undertaken to uncover clear rules of thumb for what constitutes a strong ICC in addressing inter-rater reliability. In reviewing several studies reporting ICC's, they noted that .90 or .80 was generally considered high (Hays & Revicki, 2005; Nunnally & Bernstein, 1994). Cicchetti and Sparrow (1981) suggested that a good rule of thumb is to consider anything greater than .75 as excellent, anything between .60 and .75 as good, and anything between .40 and .59 as fair to moderate agreement.

Scenarios. Most scenarios (with the exception of 'the Parked Car' and 'the Secret Murder') contained a code for unrelated decisions or no decisions (e.g. insistence that a third alternative be considered before a decision could be reached), a decision of inaction (e.g. do not pull the trolley lever, do not assign any punishment) and a decision to

complete an action (e.g. pull the trolley lever, assign a punishment). In cases where some qualification of punishment was possible (e.g. the need for leniency as in the parked car scenario) codes for full punishment and mitigated punishment were needed. Coders could also make use of an ‘unqualified punishment’ code when participants did not specify the type of punishment. All scenarios then featured a justification code system: a utilitarian category and a retributive category, which involved a 1 for reference to that orientation, and a 0 for the absence of that justification. Point-biserial correlations, which allow for comparisons of continuous and binary variables, are reported. We did not obtain enough ‘hybrid’ responses to warrant meaningful analyses of how such responses related to the other measures; implications for this issue are described in the discussion.

The parked car. The decision coding yielded excellent inter-rater reliability, ICC = .90, 95% CI [.88, .92]. Analyses also demonstrated excellent inter-rater reliability for the utilitarian and retributive categories respectively, ICC = .87, 95% CI [.85, .90], ICC = .87, 95% CI [.84, .90]. Disagreements on participant decisions tended to be either coding errors, or disagreement as to whether there was sufficient detail in the response to categorize the nature of the punishment. Disagreements with respect to the justification codes tended to revolve around whether reference to the fact that it was illegal to park in the spot was also considered utilitarian, given that laws serve the purpose of protecting the public and deterring crime. In cases where the response appeared too ambiguous to infer such an opinion, responses were coded as absent on the orientation in question.

Only 8.2% of participants indicated that John did not need to be punished; the remaining 91.8% indicated that some punishment was appropriate, with many of those unqualified in terms of severity. Additionally, most responses to this scenario were

retributive in nature. A common response was that John's negligence was what caused the deaths, therefore he was responsible. Some participants also remarked that the fact that it was unintentional mitigated his responsibility, and that therefore his punishment should be lessened. Utilitarian responses tended to involve reference to the need to follow laws to prevent such situations and deter others from parking illegally. Other participants expressed that the scenario was too unrealistic to respond to.

We created two separate variables for participant decisions. The first simply compared whether participants indicated John did or did not require punishment, and for the second we created a scale including no punishment, mitigated punishment, and severe punishment (non-decisions and unrelated decisions were treated as missing values). Assessment of bivariate relationships revealed statistically significant relationships between the subscales and both participant decision variables, but no relationship with the justification variables (both retributive and utilitarian). Higher scores on the harsh retributivism subscale related to responses involving harsher punishment, $r = .20, p < .05$, and also increased the possibility of indicating that punishment was needed, $r = .16, p < .05$. The harsh utilitarianism subscale was weakly, positively related to punishment decision, such that higher scores related to more likelihood of indicating a need for punishment, $r = .16, p < .05$. The ideal retributivism scale did not show any statistically significant relationships.

The trolley case. The decision codes yielded excellent inter-rater reliability, ICC = .86, 95% CI [.83, .89]. Analyses also demonstrated excellent inter-rater reliability for the utilitarian and retributive categories respectively, ICC = .90, 95% CI [.87, .92], ICC = .89, 95% CI [.87, .92]. A total of 50.2% of participants indicated familiarity with the

trolley dilemma, while 49.8% indicated they had not seen the dilemma prior to the study.

We therefore conducted a chi-square test to see if there was any relationship between familiarity with the dilemma and decision whether or not to pull the lever (see Table 9).

Results revealed no such contingency, $\chi^2(1) = .41, p > .10$.

Table 9. Two-way contingency table for familiarity with trolley dilemma and dilemma decision.

	Were you familiar with the trolley dilemma prior to this survey?	
	<i>Yes</i>	<i>No</i>
Do not pull lever	22	19
Pull lever	73	79

Most participants (74.9%) indicated they would pull the lever, while 21.3% indicated they would not intervene; the remaining 3.8% were unrelated responses (e.g. the possibility that the workers could get out of the way, attempts to save all six workers). The majority of responses were utilitarian in nature. The most common response was that loss of one life was preferable to loss of five lives. For the retributive responses, participants either expressed that an innocent person can never be sacrificed, or that one would be directly responsible for killing the one worker. The majority of hybrid responses featured both the desire to save the maximum number of lives, as well as being directly responsible for their deaths if one chose not to intervene. Others indicated that they believe it would be wrong to personally pull the lever and sacrifice an innocent person, but that one would be justified in saving the maximum number of lives. There was some disagreement among coders as to whether or not the notion that one cannot “play God” warranted a retributive code. Coders agreed that there was insufficient elaboration to assume that retributive reasoning was being used. Assessment of bivariate

relationships between both participant decisions and justifications did not reveal any statistically significant relationships with the POQ subscales.

The magistrate and the mob. The decision coding yielded nearly perfect inter-rater reliability, ICC = .95, 95% CI [.93, .96], with four instances in which there was disagreement as to whether the response had sufficient detail to be coded properly; in these cases participants vacillated between options and the responses were ultimately deemed to contain no decision. Analyses demonstrated good inter-rater reliability for the utilitarian and retributive categories, ICC = .62, 95% CI [.55, .69], ICC = .66, 95% CI [.60, .72] respectively. Coder disagreements surrounding justifications again tended to revolve around whether there was sufficient detail to warrant categorization. For example, some participants indicated the uncertainty that 10 people would be killed if the man was not framed, and one coder viewed this as a utilitarian concern, given the uncertainty in the usefulness of the action; because it was explicitly stated that 10 lives would be lost in the vignette, this was coded as an absent justification.

A total of 87.6% of participants indicated that the man should not be framed to save lives, while 9% indicated that the man should be framed (with 3.5% giving no decision). In terms of justification, 20.9% of participants provided a utilitarian justification, while 83.1% described retributive justifications; only 4% of participants used both. Assessment of bivariate relationships revealed no statistically significant correlations between the subscales of the POQ and participant decision or justifications.

The twins. The decision codes yielded excellent inter-rater reliability, ICC = .82, 95% CI [.78, .86]. Analyses demonstrated moderate inter-rater reliability for the utilitarian and retributive categories, ICC = .58, 95% CI [.50, .65], ICC = .63, 95% CI

[.56, .70] respectively. In regards to the decision codes, there was some disagreement as to what code should be assigned to stating that the decision should be made by a jury; it was decided that this response was too ambiguous to be coded. Concerning the justification codes, disagreements tended to occur when responses were somewhat ambiguous such that retributive key words were used alongside utilitarian logic (e.g., “at least one responsible person will be there, and they may be able to find out who really committed the crime”). Further, some participant responses were potentially driven by acceptance of one orientation, but also invoked principles from the other: “we must imprison them as DNA proves that one or both are responsible”. The decision in many ways appears to warrant an automatic ‘utilitarian’ classification, given that there is a willingness to punish an innocent person; however the need to punish the ‘responsible’ party also invokes retributive principles. It was decided that where retributive key words were present, it would warrant a ‘present’ code on that variable, but that justifications should not by default be inferred from the decisions during coding.

Among participants, 66.8% indicated that it was not acceptable to imprison both twins, while 26.2% felt it was acceptable. In justifying those decisions, 15% used some form of utilitarian justification, while 61.4% used a retributive justification; 5.9% of participants used principles of both. Assessment of bivariate relationships revealed statistically significant correlations with all three subscales. Not surprisingly, high scores on the ideal retributivism subscale related to participant indications not to imprison the twins, $r = -.29, p < .01$. Conversely, participants who showed high endorsement of harsh utilitarian and retributive principles tended to insist that both twins should be imprisoned, ($r = .23, p < .01$, and $r = .20, p < .01$ respectively). Analyses also revealed weak positive

relationships between both the harsh utilitarian and harsh retributive subscales and the tendency to use utilitarian justifications, $r = .28, p < .001$, $r = .24, p < .001$ respectively. The ideal retributivism scale also shared a significant relationship with utilitarian justifications, such that the more participants subscribed to the idea that an innocent person can never be punished, the less likely they were to draw on principles of utilitarianism. The retributive justification category did not yield any statistically significant relationships.

The crime gene. The decision codes yielded excellent inter-rater reliability, ICC = .85, 95% CI [.81, .88]. The justification variables demonstrated moderate inter-rater reliability for the utilitarian and retributive categories, ICC = .54, 95% CI [.50, .65], and ICC = .59, 95% CI [.46, .61] respectively. However, several participants showed confusion as to the question posed. While the initial vignette was amended to reflect this issue, some participants interpreted the scenario as suggesting that persons with the ‘crime gene’ would only commit a crime *on* but not after their 24th birthday, and therefore responded to whether they should be imprisoned only for the day. Given that for many responses there was insufficient detail to conclude whether the vignette was correctly interpreted, we were unable to conduct meaningful analyses on the responses. Nonetheless, some prominent responses are noteworthy. First, several participants questioned the soundness of the research that pointed to the existence of the ‘crime gene’, and some even called for responsibility on the part of the scientist in charge. In particular, participants sometimes indicated that people may overcome their genetic predispositions, and that not only is one undeserving of punishment until an actual crime has occurred, but such policies would set a “dangerous precedent”. It is likely that there are some

different punishment principles operating – perhaps in orchestra – however given the potential for participant misunderstandings, the scenario would require revisions.

Death row organ donation. The decision codes yielded nearly perfect inter-rater reliability, ICC = .92, 95% CI [.90, .94]. The justification variables demonstrated good inter-rater reliability for the utilitarian and retributive categories ICC = .67, 95% CI [.60, .73], and ICC = .73, 95% CI [.67, .78] respectively. Coder disputes tended to revolve around how to treat responding that it was only acceptable to take the inmate's organs after death (e.g. "Once he dies, then you can"). This decision appears utilitarian, but lacks elaboration as to why they must wait until the sentence is passed. In such cases where there was not enough detail, the responses received '0'.

With respect to participant decisions, 71.3% felt it was unacceptable to take the man's organs (despite the fact that he would not die), while 26.2% felt it was acceptable (with 2.5% of participants undecided). For justifications, 21.8% made reference to utilitarian principles, and 80.2% to retributive principles, while 9.9% made reference to both. Assessment of bivariate relationships revealed no statistically significant associations between the POQ dimensions and participant decisions and justifications.

Illegal handguns. This scenario was created in order to explore how participants made decisions surrounding the intentionality behind an act versus the damage it caused. However, the majority of participant responses were vague in their justifications, making it difficult to tease apart utilitarian and retributive reasoning. For example, a common response was "Sam and Tom committed the same crime." It is possible that this response arose out of the reasoning that they both had no intention for someone to be harmed by the guns they sold, which could be construed as a proportionality calculation. On the

other hand, utilitarian reasoning may dictate that violation of the law requires punishment regardless of intentionality, in order to uphold order. Therefore, this scenario was not useful for the purposes of predictive utility, and was not analyzed any further.

Social value. Much as the illegal handguns scenario, the social value scenario did not prove useful for predictive utility assessments. In this case, nearly all participants provided the same explanation in response to the question of whether the ‘social value’ of the victim can dictate the proportionality of the punishment. The consensus was that “a life is a life” and that murder requires the same punishment regardless of the status of the victim. Given that there was virtually no variability in responses, this vignette was not analyzed.

The secret murder. The decision codes yielded excellent inter-rater reliability, ICC = .82, 95% CI [.75, .87]. Coders expressed difficulty categorizing some common participant responses, as often it was clear that they believed some punishment was appropriate, but the question of whether he or she would personally punish Dan was not answered directly (e.g. “Dan still committed murder and needs to do his time for the crime”). While it is telling that some participants felt punishment was needed and some felt that it would be useless, the justification codes better captured this point, as participants tended to either discuss just deserts, the pointlessness of a non-deterring punishment, or the worry that Dan did still pose a threat to society. This latter response is somewhat problematic, as participants had particular difficulties with the ‘closed-world’ assumption, that is, accepting the details of the scenario as proposed (Conway & Gawronski, 2012). As such, this scenario would benefit from a more structured approach, such as multiple choice format in which participants are compelled to confront the facts

of the scenario as is. The justification variables did demonstrate good inter-rater reliability for the utilitarian and retributive categories, $ICC = .86$, 95% CI [.82, .89], and $ICC = .76$, 95% CI [.69, .82] respectively.

Among participants, 17.5% indicated that they would punish Dan personally; however, participants overall were unspecific as to what that punishment would be. As a result, we cannot be sure whether ‘punishment’ might mean spreading rumors, or harming Dan directly. In total, 82.6% remarked that it was not acceptable to punish Dan personally, although it was commonplace for participants to suggest contacting law enforcement. Again, the tendency for some participants to state “No, but I should call the police” made it difficult to distinguish between those who were more concerned with punishment utility than deservedness, as discomfort with vigilantism may have been involved. However, in regards to justifications, 24.5% of participants made reference to utilitarian principles, while 48.3% made reference to retributive principles. Assessment of bivariate relationships revealed no statistically significant relationships between participant decisions and justifications, and punishment orientation.

The isolated island. The decision codes yielded nearly perfect inter-rater reliability, $ICC = .95$, 95% CI [.93, .96]. The justification variables demonstrated good inter-rater reliability for the utilitarian and retributive categories, $ICC = .64$, 95% CI [.54, .71], and $ICC = .68$, 95% CI [.60, .75] respectively. There was some disagreement among coders as to whether expressing that it was wrong to lie to society was considered retributive; under deontological theory, certain acts are inherently wrong. It was decided that, since retributivism as defined in this project was described in relation to deontological principles, that this sentiment captured an intuitive retributive reaction.

Indeed this notion was sometimes accompanied by appeals to the lack of 'justice' inherent in lying to the public.

Among participants, 55.9% indicated that sending death row inmates to an island was unacceptable, while 39.1% indicated that it was acceptable; 5% of participants either refused to make a decision or insisted upon alternatives. In justifying those decisions, 31.7% of participants drew on utilitarian principles, while 62.4% drew on retributive principles; 9.4% of participants made reference to both utilitarian and retributive principles. Assessment of bivariate relationships revealed no significant relationships between the POQ subscales and decisions as well as justifications.

Exploratory Analyses

Need for cognition. The Need for Cognition (NfC) Scale yielded good internal consistency, $\alpha = .95$. Bivariate relationships were assessed in order to explore whether any of the POQ factors shared a relationship with enjoyment of effortful cognitive activities. All three factors showed a weak relationship with the NfC. The ideal retributivism factor showed a positive correlation, $r = .17, p < .05$, while both the harsh utilitarian and harsh retributive factors showed negative relationships with NfC ($r = -.23, p < .001$, and $r = -.25, p < .001$ respectively).

Positive and negative affect. Mean scores were first computed to create a positive affect score and a negative affect score. The Positive Affect Scale was normally distributed, however the Negative Affect Scale showed a positive skew ($Z_{skew} = 8.47$). The Negative Affect Scale yielded good internal consistency, $\alpha = .91$, as did the Positive Affect Scale, $\alpha = .92$. Bivariate relationships were then examined to ascertain whether any of the three factors were related to positive and negative affect. The harsh

retributivism subscale showed a weak positive correlation with positive affect, $r = .20, p < .01$, as did the harsh utilitarian subscale, $r = .21, p < .01$. No factors related significantly to negative affect.

Political ideology. We also assessed bivariate relationships between a bipolar political affiliation scale (where -10 = very liberal, and +10 = very conservative) and the three respective subscales. Results demonstrated a weak negative relationship between political affiliation and the ideal retributivism factor, $r = -.21, p < .01$, such that the more participants endorsed the idea that innocent people must not be punished, the more liberal they identified as. Conversely, the harsh utilitarian and harsh retributivism subscales showed moderate positive relationships with political affiliation ($r = .33, p < .001$, and $r = .36, p < .001$ respectively), such that the more participants endorsed those utilitarian and retributive ideals, the more conservative they identified as.

Self deceptive enhancement. The Self Deceptive Enhancement Scale (SDE; Paulhus, 1991) was normally distributed. Examination of the bivariate relationships of the three subscales and SDE demonstrated that the ideal retributivism dimension was unrelated to SDE. However, both the harsh utilitarian and harsh retributive dimensions showed weak positive relationships with SDE ($r = .29, p < .001$, and $r = .28, p < .001$ respectively).

Discussion

The final Punishment Orientation Questionnaire consisted of 16 items, for which a three factor structure demonstrated an acceptable model fit. The ideal retributivism dimension concerned a desire to ensure that innocent persons are never punished. This dimension may be useful in capturing a participant's focus on the risks associated with

punishment; that is, punishment should not be preferred over protecting innocent people from unjust harm. The extreme proportionality calculation characteristic of deontological thinking is evident, insofar as it is never permissible to punish when it is not deserved, nor as a means to an end. The harsh retributivism dimension was associated with items marked by an orientation towards punishment, and related to blame of the criminal (e.g. “criminals are bad people and deserve punishment”). Both the ideal retributivism and harsh retributivism dimensions are guided by deontological principles, but center on different subjects: the ideal factor on whom one should avoid harming, and the harsh factor on who deserves harm. There was also a punishment-prone harsh utilitarian factor, which concerned endorsement of principles that ensure public safety, and the potential danger offenders pose (e.g. the permissibility of over-punishing to achieve some desired end).

What appears to be lacking is a dimension accounting for a positive counterpart to this calculating utilitarian one. The initial EFA in Study 1 revealed that items featuring a desire to improve society constituted a potential factor; however, item analysis showed widespread agreement with some of the related questions, which ultimately provided little utility. Nonetheless, it is possible that there is a fourth factor representing the more optimistic aspects of utilitarianism, for example the idea that the aggregate benefits from practices that encourage rule following. In fact, as Conway and Gawronski (2013) remarked, there may be two sides of ‘utilitarianism’. On one hand, there might be a lack of concern surrounding using others as a means to an end, as demonstrated by a link with Machiavellianism (Bartels & Pizzaro, 2011). On the other hand, there may be genuine concern for the well-being of the aggregate. Uncovering such a factor could then lend to a

twofold construct: one with a punishment or offender focus, and the other with a punishment averse, ideal focus. This concept is partially related to Carlsmith et al.'s (2002) speculation that there is a relevant distinction between macro and micro centered punishment. Investigation of that possibility may result in a closer model fit; while the CFA uncovered an adequate fit, further exploration is needed to achieve a more well-rounded representation of these attitudes. It may likewise be beneficial to strengthen the ideal retributivism dimension by adding items pertaining to the duty to avoid using others as a means to an end.

Notwithstanding these caveats, supplementary analyses revealed some noteworthy relationships between the dimensions and other individual differences. First, endorsement of items surrounding the importance of just deserts as well as the need to protect society coincided with more conservative political affiliation. Conversely, the desire to protect innocent people from unjust harm coincided with more liberal political affiliation. We did not employ a full measure of conservatism and liberalism, but rather asked participants to place themselves on a single continuum; nonetheless, there is a relationship between stronger identification with certain political values and the two orientations. Okimoto, Wenzel, and Feather (2012) found a relationship between retributivism (defined as the need to punish an offender for the sake of just deserts) and conservative values. Likewise, both our harsh retributivism and utilitarian subscales seemed to focus on the need to punish offenders – whether for deserts or for safety – therefore this relationship is not surprising. Finally, Haidt (2007) has shown that people who identify as ‘very liberal’ versus ‘very conservative’ find different moral foundations to be most important; liberals tend to emphasize the relevance of protecting others from harm and ensuring reciprocity,

whereas conservatives tend to emphasize traditional authority, loyalty to the group, and purity.

It is telling that the harsh utilitarian and retributivism dimensions both showed opposite valence in their relationships with the need for cognition and positive affect (no relationship was found with negative affect). While previous researchers have given ample support for the association between cognitive processes and utilitarian reasoning, our subscale had a negative relationship with enjoyment of effortful cognitive activities, and a positive relationship with positive affect. It may be that the two dimensions are associated with some emotional component that dictates satisfaction surrounding punishment, additionally supported by the desire to avoid deliberate cognitive reflection. In essence, the two dimensions may arise from the simple intuition that criminals need punishment, for virtually any benefit, societal or abstract. The association between positive affect and the utilitarian subscale is consistent with Valdesolo and DeSteno's (2006) finding that inducing positive feelings was conducive to utilitarian responding, because it makes one more able to commit a moral violation. They reasoned that people are better able to endorse a morally questionable act when they perceive that act as less aversive, which can be accommodated through higher positive affect. Their results demonstrated support for this hypothesis, such that participants who viewed a positive video clip made more utilitarian judgments than those who viewed a neutral video. They concluded that manipulation of affect can strongly influence a person's moral judgments. While we do not conclude that positive affect influenced participant endorsement of utilitarian premises (positive affect was measured after the questionnaire and scenarios were completed), there exists some correspondence between agreement with those

principles and reporting positive affect. Notably, we did ask participants to report on how they feel on average or in general.

Limitations

Some further considerations limit the scope of the conclusions arising from this study. Specifically, the POQ dimensions did not demonstrate divergent validity according to our hypothesized relationships. While the harsh utilitarianism and harsh retributivism dimensions did relate to McKee and Feather's (2008) Sentencing Goals Scale, they also showed a similar relationship to Schiffhauer and Wrightsman's (1997) Scale of Attitudes toward the Legal System. Our aim was to create a measure that did not have substantial overlap with beliefs about how the system functions in reality. However, it may be difficult to tease apart such attitudes. After all, we might expect that one's appraisal of a punitive system is contingent upon or in some way related to what they believe it ought to accomplish. At the least, the relationships did not approach problematic multicollinearity levels (Cohen, Cohen, West, & Aiken, 2003). Moreover, the ideal retributivism subscale appeared to capture concepts that were fairly unrelated to legal attitudes, and McKee and Feather's scale also did not capture such a dimension. It additionally bears reminding that, while the CFA results suggested that harsh retributivism and utilitarianism were better represented by a two factor structure as opposed to a unidimensional one, they did not demonstrate discriminant validity. The strong correlation between them ($r = .76$) is indicative of a substantial relationship between the principles as the questionnaire conceptualizes them. More optimistically, the desire for protecting innocents – which also features an adherence to deontological thinking – showed only modest overlap with both constructs. It seems that willingness to over-punish 'criminals' does not quite

parallel the possibility of accidentally punishing innocent people; the question of potential justice system misfires is of a different kind.

It is always advisable to reflect on the generalizability of results. In its current form, the POQ displayed the previously described relationships within a sample of adults working on Mechanical Turk. These workers tend to have more exposure to the hypothetical dilemmas (the trolley case in particular), which may mean they have spent more time reflecting on them than might be true for other populations (Chandler, Mueller, & Paolacci, 2014). Indeed, half of our participants reported familiarity with the trolley dilemma. Further, our sample consisted of North-American (American and Canadian) participants only. It is possible that some non-Western cultures would emphasize different moral values, and as such, replication is needed with diverse samples.

Finally, results are in some ways limited by the lack of strong predictive utility of the subscales in relation to many of the moral dilemmas. First, we were unable to analyze a number of scenarios due to participant confusion (e.g., surrounding the nature of the ‘crime gene’), and consensus on certain intolerable responses (e.g., that the social value of a murder victim is relevant). Second, several of the scenarios still yielded low variability in answers; for example, most participants remarked that punishment was needed in the parked car scenario, and provided retributive justifications. Participants frequently elected to pull the lever (a utilitarian decision) in the Trolley Dilemma as well. Hence, they tended to favor one type of orientation dependent on the scenario; they changed from mainly retributive in the Parked Car scenario, to overwhelmingly utilitarian in the Trolley Dilemma. What is more, they were largely willing to sacrifice a worker in

the Trolley Dilemma, yet unwilling to allow Mary the Mayor to frame a man to save people from being harmed by rioters. Clearly, participants will draw on utilitarian and retributive principles as it is convenient to do so. Given the relative salience of the ‘victim’ to be framed in the mob case, it may be more difficult to ignore retributive proportionality concerns, in comparison with the trolley case in which little is described about the victim. This difference in participant responses echoes that made between the traditional trolley dilemma and the footbridge variant, in which one must physically push a large man off of the bridge into the path of the trolley (Thomson, 1985). As Greene et al. (2001) showed, the inclusion of a personal element that would require a moral violation activates more emotional moral processing, hence deontological thinking is more available. Conway and Gawronski (2013) argued that there are situations in which it is easier to suppress a deontological response and accommodate a utilitarian one, and ones in which the two possible choices are at high conflict. They further distinguish between moral ‘inclinations’ and judgments, noting that assuming those inclinations parallel decisions erroneously pits utilitarianism against deontology. That is, such conceptualization ignores the possibility that a decision is reached easily because one is weak while the other is strong, or that both are strong and one ultimately rules. It would be informative for future studies to alter the details of the cost-benefit calculation in the trolley problem, for example, by adding information about deservedness of the five workers.

It is pertinent to note that whereas some scenarios compared deontology to utilitarianism, others specifically probed deontological and utilitarian *punishment* decisions. For the trolley case, the participant must only decide whether it is acceptable to

save five lives instead of one, whereas the mob case emphasizes the conflict between sacrificing a human life and fulfilling abstract justice ideals. In brief, it is perhaps not equally useful to question whether the subscales predict classic utilitarian and retributive justification across the scenarios. Some may have simply been more congruent in terms of specificity. For instance, the Twins Dilemma, which revealed statistically significant associations with the dimensions, related more directly to the questionnaire – the scenario featured precisely the issue of whether risking an unjust punishment is a worthwhile payoff to protect others. Again, the trolley scenario features only general utilitarian and deontological principles. It does not introduce issues specific to punishment, such as the implications it would have for the legal system, or for a continued threat to the public. Rather, it is a literal cost benefit calculation.

Future investigations would benefit from a more nuanced coding scheme, which would pinpoint specific principles falling under general deontological and utilitarian framework (e.g. a code for ‘strict liability’ versus ‘intentionality’). However, because the goal of this study was to ascertain whether a.) our measure would predict classic utilitarian and retributive dilemmas and b.) whether punishment orientation relates to their traditional definitions, we elected to follow a more traditional paradigm. In our view, uncovering a lack of association would still contribute to a better understanding of punishment attitudes. Further, such in depth assessment of participant open ended responses would require support of moral theory not discussed in the current paper. For example, it may be useful to draw on measures that detect language that is key to certain moral foundations (Graham et al., 2011). Such an investigation would require participants to provide longer responses than one to two sentences, and are beyond the scope of this

study. The lack of predictive utility for some of the scenarios is nonetheless a useful finding. As will be discussed shortly, it is increasingly apparent that a ‘utilitarian’ and ‘retributive’ dichotomy of a measure of punishment attitude does not per se express substantive information.

General Discussion

The purpose of this study was to create and validate a measure of utilitarian and retributive punishment orientation, given that little attention has been paid to a comprehensive questionnaire of punishment attitudes. By focusing on the possibility of a trait-like, overall punishment orientation, we aimed to illuminate important aspects of lay punishment theories. Initial exploratory factor analysis suggested that all retributive and utilitarian ideas did not fall onto a single factor. In fact, the final questionnaire resulted in a three factor structure, which manifested utilitarian and retributive principles that worked together, rather than being adversarial. A third dimension, which we termed “ideal retributive,” was so called because it latched on to the preservation of innocent people; two forms of retributivism thus appeared as two dimensions. Furthermore, the factors showed predictive utility where this latter issue was central: can you imprison two possibly guilty people to ensure that at least one guilty person is accounted for? While the POQ showed a strong relationship with an existing punishment goals questionnaire, it also showed a similar relationship with beliefs about the actual legal system. Final exploratory analyses revealed some interesting points of discussion, including associations with positive affect, the need for cognition, and political affiliation, in relation to current thinking on moral reasoning. Results from two studies designed to explore and then test potential latent factors involved in punitive judgments are indicative

that a division by ideas of ‘utility’ and ‘retribution’ may be less fruitful than was anticipated. In essence, it is more likely that people will vary in their refusal to abandon the enterprise of punishment.

It is significant that the items aimed to capture the desire for punishment to benefit society with no ulterior motive (e.g. “The goal of punishment should be to make society a better place”) received majority agreement from participants. Yet, the questions probing a desire to punish out of more primitive needs, as with “If someone commits a crime, that person is dangerous and must be removed from the community to protect other people” and “The goal of punishment should be to give the guilty what he deserves” tended to better discriminate among participants. The worry that an innocent person will get caught in the punitive cross-fire also better separated people. Across the board, participants appeared willing to agree with notions of societal ‘good’, yet some still indicated a preference for righting past wrongs.

Some retributive and utilitarian focused items (e.g. the acceptability of over-punishing for society’s benefit, and the desire to right past wrongs through just deserts) were strongly related, while others (i.e., the concern for protecting innocent people) shared far less overlap. We may fairly conclude that lay punishers are not lay philosophers. There are circumstances in which one orientation is more useful, and it is possible to value these justifications simultaneously. Despite the degree of logical incompatibility between many of the principles underlying each theory, they do not by default come into conflict as over-arching values. It would seem that the measure described in this paper captured some aspect of punitive attitudes in which utilitarian and retributive principles can contribute to the same psychological need, rather than fulfilling

entirely different ones. A metaphor described by Haidt (2007, p. 999) summarizes our findings effectively: when it comes to punishment, people are more like lawyers and politicians than “intuitive scientists”. Interestingly, philosophers and moral psychologists have become captivated by the study of ‘moral dumbfounding’, a phenomenon by which people have a strong reaction to an issue, yet are completely incapable of explaining it (Haidt, 2007). There is nothing preventing lay punishers from adopting conflicting viewpoints, and they may even have strong enough intuitive convictions so as to cling to a decision. With further investigation, the POQ may be a useful measure of when participants are relentless in the desire to punish, regardless of the justification; this assertion is consistent with the strong relationship between pro-punishment items and the question probing the morality of punishment in a world in which it has no utility at all. It would be interesting to provoke moral dumbfounding with respect to utilitarian and retributive punishment goals, and understand whether participant endorsement of POQ items would change.

In the same vein, it is worth revisiting the finding that the POQ showed little predictive utility in relation to classic utilitarian and retributive judgments. In many ways, results from these moral dilemmas are supportive of a strong state-based component to punitive decisions. Participants tended to change from utilitarian to retributive justifications depending on the details of the scenario, despite the fact that they were constructed to allow for either justification to accompany both decisions. For example, participants could conceivably say: “Do not imprison both twins because that would be a slippery slope” versus “Do not imprison both twins because one may be innocent”. Similarly, participants could respond: “Imprison both twins to protect society” or

“Imprison both twins because the guilty one needs to see justice”. Reflecting on the historical popularity of Kohlberg’s theory of moral development, which emphasized children as “budding moral philosophers”, Haidt (2007, p. 998) noted the “affective revolution” that developed in the field of moral psychology. This shift was informed by the insight that more primitive, automatic and emotional processing may be what leads us to either avoid or approach certain acts (Zajonc, 1980, as cited in Haidt, 2007). Therefore, it is not surprising that lay punishers do not deliberate deeply on the nuance of punishment philosophies, but instead often allow their intuitions to guide them. Consequently, the process by which participants agree or disagree with pre-provided statements about punishment may differ in key ways from the task of using moral language to reflect on a decision, especially when more contextual details are given. Finally, work by Green et al. (2001) has found that people are less likely to endorse acts that might cause harm if they have an automatic negative reaction. The way that the questionnaire items were framed may have instead built in the possibility that harming a ‘criminal’ who deserves a punishment is not a negative violation. As Green et al. (2001) also noted, given sufficient motivation (in this case, apparently to punish ‘bad’ or ‘dangerous’ criminals), people may keep the aversive nature of a moral violation from interfering with the decision at hand. On the contrary, most of the dilemmas described more risk involved in terms of causing harm, and depicted direct utilitarian and retributive conflicts.

In assessing the role of punishment attitudes, it seems that people may nonetheless vary in the degree to which they approve of such principles without situational details, although not in the manner that was expected. Rather, lay punishers

differ in the extent to which they accept risking harm to an undeserving person out of desire for punishment, and a desire to meet crime with harsh sanctions. The findings of this study may be supportive of investigating punishment orientation from an interactionist perspective of punitive ‘states’ and ‘traits’, such that people have some idea of what principles are important. It does appear that such attitudes would be quite malleable for some, while others may be highly motivated to punish regardless. Our findings are consistent with previous research inasmuch as they show that retributive and utilitarian ideas need not be polar opposites; in fact, they can be highly complementary. Many participants accepted the importance of just deserts, yet were willing to punish more than might be deserved. Further still, even with regard to principles that should come into major conflict from a philosophical standpoint (e.g. the punishing of an innocent versus the acceptability of over-punishing), the overlap was modest, supporting the assertion that they are two independently functioning systems (Conway & Gawronski, 2013; Greene et al., 2001; Tanner, Medin, & Iliev, 2008).

Implications

The findings of this project are somewhat disconcerting when thinking about how lay people come to punitive decisions. From one perspective, we might become concerned that, if lay punishers are not engaging the nuances of conflicting orientations, then they are not truly guided by them. Instead they may be guided by the desire for punishment. This assertion harkens back to Carlsmith et al.’s (2002) conclusion that people are essentially retributivists. As Tebbit (2005, p.170) articulated, “Ultimately, it is argued, the retributive language of desert is mere rhetoric to mask the absence of an intelligible justification”. Perhaps people appeared mainly retributive in Carlsmith et al.’s

study because they were in a situation that called for punishment; maybe when people are motivated to punish, they will draw on any number of ideologies. By its very definition, retributivism is retrospective, meaning that the decision only enters by virtue of a wrongdoing (Tebbit, 2005). That is to say, a retributive punishment orientation is already begging the question; when one uses retributive punishment, it is because he or she wants to punish. This notion in turn poses a graver dilemma: do the ethics of punishment become less accessible to lay people as the stakes increase? It may be that once a crime hits a certain level of severity, the ethical reasons for punishment meld together. Recalling Weiner, Graham, and Reyna's (1997) assertion that divisions between punishment motives may become "fuzzy", it is possible for utilitarianism and retributivism to converge on the same punitive magnitude (p. 432). Weiner et al. (1997) use capital punishment as an illustration, since one might endorse such a course because it is 'deserved', or to permanently incapacitate the offender. Yet, it is unclear whether a severe punishment such as a death penalty would be driven by pure utilitarian goals, or if instead a utilitarian goal would simply be an added 'benefit' in the eyes of the person imposing it. Conversely, as crime severity increases, it could simply be that if a harsh retributive orientation dominates, it renders the decision much less malleable.

Further consideration of the matter leads to even grimmer territory. If it is not possible to relieve people of their punitive needs, then we may be ourselves relieved that the current role of citizens is to judge the guilt of, but not sentence, defendants. However, the implications would be grave still for capital juries. One wonders if, when we entrust ordinary citizens with the responsibility to dole out punishment, they invoke punishment for reasons that are beyond the ken of the average person. As an illustration, "death

qualified” jurors – so called because they do not have strong opinions on using the death penalty and may therefore sit on capital juries – differ in important ways from those who would be removed. One striking finding indicated that they are more conviction prone (Butler & Moran, 2002; Fitzgerald & Ellsworth, 1984; Young, 2004). If capital jurors are more conviction prone, then they may be punishment motivated, or potentially more retributive in their punitive needs. This concept also calls to question whether utilitarian principles would play a role at all. It would seem that, if lay punishers are not lay philosophers, we have some obligation to better instruct jurors on the ethics of punishment. Currently, jurors are – often confusingly – instructed to weigh ‘aggravating’ and ‘mitigating’ evidence in determining whether or not to impose a death sentence. First, this task invokes retributive principles to a strong degree, because it is aimed at a deservedness as well as proportionality calculations. Second, it is no stretch of the imagination to suggest that jurors may lack the full range of tools to consider, in depth, the morality of their punitive decision. Tebbit (2005) zeroed in on the “mystery” surrounding retributive notions of ‘ought’, questioning how, from an ethical standpoint, such drive for punishment arises purely from the fact that a crime has occurred (p.170). A full explanation of the merits of retributive theory is complex, and even theorists continue to deconstruct its mechanisms.

Hence, it might be appropriate to remind that philosophical discussion surrounding the ‘ought’ of punishment has also been in conflict for centuries. The basic building blocks of punitive philosophies are these same intuitively plausible notions, which lay persons are attracted to and philosophers have yet to reconcile. Therefore, in some ways lay punishers are lay philosophers, insofar as their decisions and attitudes

manifest both the desire to protect others and the desire for justice. As Tebbit (2005) went on to describe, many theorists have accepted the need for both accounts of punishment, and have transformed the question into: which principle ought we to err on? In any case, the current study cannot answer the question of what is driving lay punishment; however it is clear that people do not necessarily experience conflict in using incompatible principles.

Notwithstanding the issue of whether lay persons need to improve upon their skills as punitive deliberators, it remains to be seen whether they can acquire the tools that philosophers have made available for such contemplation. It would be useful for future investigators to see if participants would make different decisions with added moral theory and language at their disposal, as Haidt (2007) espoused. For instance, perhaps it would be fruitful to better explain to jurors the nature of ‘aggravating’ and ‘mitigating’ circumstances in relation to deontological principles aimed at moral duty rather than ones aimed at proportionality. It is easy to imagine the benefits of motivating jurors to reflect on their own motivations to use punishment.

Future Directions

This project leaves much wanting in regards to uncovering the true nature of punishment orientation, and the topic is therefore ripe for further empirical investigation. For example, we did not consider the myriad of other factors that may come into play. It would be informative to explore whether there are cultural differences surrounding use of utilitarian and retributive principles; one might expect that individualist and collectivist worldviews could alter punitive needs. It may even be the case that people make different decisions and rely on different explanations as a function of in-group versus out-group

membership. There may similarly be interesting gender differences that emerge in future research; in all POQ items, the subject of punishment was male. It will be worth exploring other types of situational variations that can alter punitive focus as well. Results revealed that some participants did imagine a particular crime, which ranged from very serious (e.g. murder), to relatively minor (e.g. theft), and as such it would be interesting to test variants of the POQ in which the participant is asked to imagine different crimes. Additionally, it may lend more depth to our understanding of punishment orientation to explore how adherence to such ideals develops throughout key life stages. Finally, it would be informative to study the ways in which participants attempt to persuade each other of their reactions to moral dilemmas. As Haidt (2007) remarked, being able to hear novel viewpoints that one would not normally seek out can prompt another affective ‘flash’, leading to a different decision.

In short, the POQ in its current state may be useful for determining how much a person is motivated toward punishment; the utilitarian and retributivism subscales may illuminate whether the focus is on the greater public, or on punishing the transgressor. The ideal retributivism subscale may help to highlight whether one’s retributive ‘threshold’ is likely to override utilitarian reasoning. Given further exploration of the issues described, the measure could find use in a variety of settings. For instance, the ability to measure punitive orientation may aid researchers beyond case types where actual punishment is involved. It would be beneficial to ascertain how punishment orientation impacts on insanity defence attitudes, given that a finding of not criminally responsible implies that the defendant is not punishable. Skeem, Loudon, and Evans (2004) have implicated retributive and utilitarian punishment orientation in insanity

defence attitudes. That is, people who tend to feel that a person should take full responsibility for the harm they have caused another may show an affinity for just deserts as a mode of thinking. They also noted that those who fear the offender is dangerous may be in favor of utilitarian punishment goals. Notably, both the ‘strict liability’ and ‘injustice and danger’ factors of their Insanity Defense Attitudes Scale – which would see retributive and utilitarian principles working in harmony – have a view to the same goal: do not use the insanity defence.

In summation, this study contributes to the body of research aimed at better understanding the forces behind ethical reflection. Interdisciplinary researchers should consider the implications of how a rich tradition of philosophical thought informs the lay psychology of punishment, and perhaps even how lay theories can reciprocate that influence. The POQ represents an added tool in contemplating such a key aspect of moral life.

References

- Banaji, M.R., & Heiphetz, L. (2010). Attitudes. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology Volume 1, 5th Ed* (pp. 353-393). John Wiley & Sons Inc, Hoboken, NJ.
- Barrett, P. (2007). Structural equation modeling: Adjudging model fit. *Personality and Individual Differences, 42*, 815-824.
- Bartels, D., & Pizzaro, D. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition, 121*, 154-161.
- Bentham, J. (2006). An introduction to the principles of morals and legislation. In S. Cahn, & P. Markie (Eds.), *Ethics – History, Theory, and Contemporary Issues 3rd Ed* (pp. 309- 316). Oxford University Press (original work published 1789).
- Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.
- Butler, B. M., & Moran, G. (2002). The role of death qualification in venirepersons' evaluations of aggravating and mitigating circumstances in capital trials. *Law and Human Behavior, 26*(2), 175-184.
doi:<http://dx.doi.org/10.1023/A:1014640025871>
- Cacioppo, J., Petty, R., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*(3), 306-307.
- Carlsmith, K. M. (2001) Why do we punish? Retribution, deterrence, and incapacitation as motives for punishment. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 1640-1640.

- Carlsmith, K. M. (2006). The roles of retribution and utility in determining punishment. *Journal of Experimental Social Psychology, 42*, 437-451.
- 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.
- Carroll, J. S., Perrowitz, W. T., Lurigio, A. J., & Weaver, F. M. (1987). Sentencing goals, causal attributions, ideology, and personality. *Journal of Personality and Social Psychology, 52*(1), 107-118.
- Cattell, R. B. (1966). The screen test for the number of factors. *Multivariate Behavioural Research, 1*, 245-276.
- Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods, 46*(1), 112-130. doi:10.3758/s13428-013-0365-7
- Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency, 86*(2), 127-137.
- Comrey, A. L. (1989). Factor analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology, 56*, 754-761.
- Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology, 104*(2), 216-235.
- Crowne, D., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.

- DeVellis, R. (2012). *Scale Development: Theory and Applications Third Edition*. North Carolina: Sage Publications, Inc.
- Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology*, 43(6), 543-549.
- Fitzgerald, R., & Ellsworth, P. C. (1984). Due process vs. crime control: Death qualification and jury attitudes. *Law and Human Behavior*, 8(1-2), 31-51.
doi:<http://dx.doi.org/10.1007/BF01044350>
- Fletcher, I., Mazzi, M., & Neubling, M. (2011). When coders are reliable: The application of three measures to assess inter-rater reliability/agreement with doctor-patient communication data coded with the VR-CoDES. *Patient Education and Counselling*, 82, 341-345.
- Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, 5, 5-15.
- Garson, D.G. (2013). *Factor Analysis*. North Carolina: Statistical Associates Publishing.
- Greene, J., Sommerville, R., Nystrom, L., Darley, J., & Cohen, J. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293, 2105– 2108.
- Graham, M., Milanowski, A., & Miller, J. (2012). Measuring and Promoting Inter-Rater Agreement of Teacher and Principal Performance Ratings. *Online Submission*.
- Greene, J. (2009). Dual process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie. *Journal of Experimental Social Psychology*, 45, 581-854.
- Goldman, A.H. (1979). The paradox of punishment. *Philosophy and Public Affairs*, 9(1),

42-58.

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316(5827), 998-1002.

doi:<http://dx.doi.org/10.1126/science.1137651>

Hays, R. D., & Reviki, D. A. (2005) Reliability and validity (including responsiveness).

In P. M. Fayers & R. D. Hays (eds). *Assessing quality of life in clinical trials: Methods and practice*. NY: Oxford University Press.

Heubel, F., & Biller-Andorno, N. (2005). The contribution of Kantian moral theory to contemporary medical ethics: A critical analysis. *Medicine, Healthcare, and Philosophy*, 8, 5-18.

Hogarth, J. (1971). *Sentencing as a Human Process*. Toronto, Canada: University of Toronto Press.

Kant, I. (2006). Groundwork of the metaphysics of morals. In S. Cahn, & P. Markie (Eds.), *Ethics – History, Theory, and Contemporary Issues* 3rd Ed (pp. 270- 308). Oxford University Press (original work published 1785).

Kaiser, H. F. (1960). The application of electronic computers to factor analysis.

Educational and Psychological Measurement, 20, 141-151.

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

Lauter, D. (2013, February 10). Double trouble for Marseille police hunting serial rapist as DNA turns up identical twins. *The Telegraph, Europe*. Retrieved from <http://www.telegraph.co.uk/news/worldnews/europe/france/9861164/Double-trouble-for-Marseille-police-hunting-serial-rapist-as-DNA-turns-up-identical->

twins.html

- Lessnoff, M. (1971). Two justifications of punishment. *Philosophical Quarterly*, 21, 141-148.
- Markland, D. (2007). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modeling. *Personality and Individual Differences*, 42, 851-858.
- McFatter, R. (1978). Sentencing strategies and justice: Effects of punishment philosophy on sentencing decisions. *Journal of Personality and Social Psychology*, 36(12), 1490-1500.
- McKee, I., & Feather, N.T. (2008). Revenge, retribution, and values: Social attitudes and punitive sentencing. *Social Justice Research*, 21, 138-163.
- Meng, X., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, 111(1), 172-175.
Doi:<http://dx.doi.org/10.1037/0033-2909.111.1.172>
- Mill, J. S. (2008). *On liberty*. A. S. Kahan (Ed.). Boston: Bedford (original work published 1859).
- Millar, M. G., & Tesser, A. (1986). Effects of affective and cognitive focus on the attitude-behavior relation. *Journal of Personality and Social Psychology*, 51(2), 270-276. Doi:<http://dx.doi.org/10.1037/0022-3514.51.2.270>
- Muthén, L. K., & Muthén, B. O. (1998-2011). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.
- Myers, B., & Greene, E. (2004). The prejudicial nature of victim impact statements: Implications for capital sentencing policy. *Psychology, Public Policy, and Law*,

10(4), 492-515.

- Nezlek, J. (2007). A multilevel framework for understanding relationships among traits, states, situations and behaviour. *European Journal of Personality* 21, 789-810.
- Nielsen, K. (1972). Against moral conservatism. *Ethics: An International Journal of Social, Political, and Legal Philosophy*, 82, 219-231.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Paulhus, D. (1991). Measurement and control of response bias. In J.P. Robinson, P.R. Shaver, & L. S. Wrightsman (Eds.), *Measures of Personality and Social Psychological Attitudes* (pp. 17-59). San Diego, CA: Academic Press Inc.
- Rawls, J. (1955). Two concepts of rules. *Philosophical Review*, 64, 3-32.
- Schedler, G. (1980). Can retributivists support legal punishment? *The Monist*, 331-334.
- Schiffhauer, K.L., & Wrightsman, L.S. (1995, November). *The Measurement of Attitudes toward the Legal System*. Unpublished paper, Department of Psychology, University of Kansas, Lawrence, KS.
- Tebbit, M. (2005). Theories of punishment. *Philosophy of Law: An Introduction* (pp. 192-212). New York, NL: Routledge.
- Thomson, J.J. (1985). The trolley problem. *The Yale Law Journal*, 94(6), 1395-1415.
- Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science*, 17(6), 476-477.
- Vidmar, N., & Miller, D. (1980). Socialpsychological processes underlying attitudes toward legal punishment. *Law & Society Review*, 14(3), 565-602.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief

- measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070.
- Weigel, R., Vernon, D., & Tognacci, L. (1974). Specificity of the attitude as a determinant of attitude-behavior congruence. *Journal of Personality and Social Psychology*, 30(6), 724-728.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92 (4), 548-573.
- Weiner, B. (2006). *Social Motivation, Justice, and the Moral Emotions* (pp. 125-159). Mahwah: Lawrence Erlbaum Associates.
- Weiner, B., Graham, S., & Reyna, C. (1997). An attributional examination of retributive versus utilitarian philosophies of punishment. *Social Justice Research*, 10(4), 431-452.
- Young, R. L. (2004). Guilty until proven innocent: Conviction orientation, racial attitudes, and support for capital punishment. *Deviant Behavior*, 25(2), 151-167.
doi:<http://dx.doi.org/10.1080/01639620490266916>
- Zajonc, R. B. (1980). Feeling and thinking: preferences need no inferences. *American psychologist*, 35(2), 151-175.

APPENDIX A: EXAMPLE of MECHANICAL TURK RECRUITMENT NOTICE

Requester: Legal Decision Making Lab Reward: \$0.50 Duration: 30 minutes

Answer a psychological survey: “Moral Decision-Making”

We are looking for participants to complete an academic survey on punishment. Participants will be asked to answer 2 opinion questionnaires and some demographic questions. The study will take approximately 5 to 15 minutes and participants will be awarded \$0.50. At the end of the survey, you will receive a code to paste into the box below to receive credit for taking our survey.

This study has been approved by the Carleton University Ethics Committee for Psychological Research (#xx-xxx).

[Click here to take survey.](#)

Provide the survey code here:

APPENDIX B: CONSENT FORM**Informed Consent**

The purpose of an informed consent is to make sure that you understand the purpose of the study and your involvement as a participant. The informed consent must include enough information regarding the study for you to be able to make a well-informed decision regarding whether or not you would like to partake in the study.

Title: Moral Decision-Making

Research Personnel: This study is being conducted by Dr. Evelyn Maeder from the Institute of Criminology and Criminal Justice at Carleton University, and Master's student Susan Yamamoto from the Department of Psychology.

Purpose: This is a study to evaluate moral decisions.

Task Requirements: You have been asked to participate in a study concerning punishment. During this study you are asked to answer 2 brief questionnaires. This study will take you approximately 5-15 minutes to complete and you will be compensated with \$0.50. You will receive a completion code at the end of the study to paste into Mechanical Turk in order to receive compensation.

Potential Risk/Discomfort: You will also be asked to answer questions regarding your attitudes about certain sensitive topics (e.g. punishing another person), and you may become uncomfortable with the nature of the questions. At any time, you may discontinue your involvement in this study and you will still be paid for your participation.

Right to Withdraw: In addition, you may refrain from answering any questions on the questionnaire if you are uncomfortable or otherwise do not want to. At the end of the study, you will be asked if you would like to withdraw your data from the study and will have the option of doing so immediately if you should so wish.

Anonymity/Confidentiality: The data collected in this experiment are strictly confidential. All data are coded such that your name is not associated with the responses you provide. Any identifying information associated with your code will be confined to a single page that will be separated from your questionnaire, and kept in a separate, secured file by the research investigators, who will keep this information confidential. We collect data through the software Qualtrics, which uses servers with multiple layers of security to protect the privacy of the data (e.g., encrypted websites and pass-word protected storage). Please note that Qualtrics is hosted by a server located in the USA. The United States Patriot Act permits U.S. law enforcement officials, for the purpose of an anti-terrorism investigation, to seek a court order that allows access to the personal records of any person without that person's knowledge. In view of this we cannot absolutely guarantee the full confidentiality and anonymity of your data. With your consent to participate in

this study you acknowledge this. Your name and phone number will be kept in our participant database so that we can keep track of who has completed our study. This database will be kept for two years, and then destroyed. The information in this database cannot be matched to your responses.

Should I require further information regarding this study, I can contact:

Dr. Evelyn Maeder
Institute of Criminology and Criminal Justice
Tel. 613-520-2600 ext. 2421

Susan Yamamoto, Master's Student
Department of Psychology
susanyamamoto@email.carleton.ca

Should I have any ethical concerns regarding this study, I can contact:

Dr. Shelley Brown
Chair, REB for Psychological Research, Department of Psychology, Carleton University
Tel. 613-520-2600 ext. 1505

Dr. Anne Bowker
Chair of Department of Psychology, Carleton University
Tel. 613-520-2600 ext. 8218

This study has been approved by the Carleton University Ethics Committee for Psychological Research (#13-xxx).

€ *I have read the above form and understand the conditions of my participation. My participation in this study is voluntary, and I understand that if at any time I wish to leave the experiment, I may do so without having to give an explanation and with no penalty whatsoever. Furthermore, I am also aware that the data gathered in this study are confidential and anonymous with respect to my personal identity. **By checking this box, I'm indicating that I agree to participate in this study.***

€ *I have read the above form and understand the conditions of my participation. My participation in this study is voluntary, and I understand that if at any time I wish to leave the experiment, I may do so without having to give an explanation and with no penalty whatsoever. Furthermore, I am also aware that the data gathered in this study are confidential and anonymous with respect to my personal identity. **By checking this box, I'm indicating that I disagree to participate in this study.***

APPENDIX C : PUNISHMENT ORIENTATION SCALE

1. Punishment is necessary because it deters other individuals from committing the same crime.

1 2 3 4 5
Strongly Disagree Neutral Strongly Agree

2. Punishment is necessary because it helps to rehabilitate the offender.

1 2 3 4 5
Strongly Disagree Neutral Strongly Agree

3. Punishment is necessary because it removes a dangerous person (the offender) from the community.

1 2 3 4 5
Strongly Disagree Neutral Strongly Agree

4. Punishment is necessary because it restores the balance of justice.

1 2 3 4 5
Strongly Disagree Neutral Strongly Agree

5. Punishment is necessary because it sends the message to the community that crime will not be tolerated.

1 2 3 4 5
Strongly Disagree Neutral Strongly Agree

6. It is better to let 10 guilty criminals go free than to punish 1 innocent person.

1 2 3 4 5
Strongly Disagree Neutral Strongly Agree

7. It is more important to keep innocent people free from punishment than it is to ensure that all guilty persons are punished for their crimes.

1 2 3 4 5
Strongly Disagree Neutral Strongly Agree

8. The safety of other people is more important than individual justice.

1 2 3 4 5
Strongly Disagree Neutral Strongly Agree

9. Criminals are in prison to keep the community safe.

1 2 3 4 5
Strongly Disagree Neutral Strongly Agree

10. It is better to let one dangerous criminal go unpunished than it is to punish an innocent person.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

11. It is more important to punish a guilty person because he deserves it than it is to punish him to benefit society.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

12. Punishment is not about “an eye for an eye” – if we are to harm a person using punishment, there must be some benefit to the community.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

13. The goal of punishment should be to make society a better place.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

14. The goal of punishment should be to give the guilty person what he deserves.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

15. Criminals are bad people and deserve punishment.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

16. Many people from rough backgrounds do not become criminals – criminals could choose to follow the law, but they choose not to.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

17. There are many situational pressures that could lead an otherwise good person to commit a crime.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

18. If a person hits someone with the intent to slightly harm him, but causes significant harm he should be punished for the significant harm that he caused.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

19. When a person commits the crime, he upsets the balance of justice. The only way to make this right is to punish him accordingly.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

20. Punishment should match the crime- if you commit a serious crime, you should receive a serious punishment.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

21. If a crime has a low detection rate (i.e. it is difficult to catch criminals who commit this particular crime), we should punish those who are caught harshly to prevent others from thinking they can get away with it.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

22. Crimes that receive a great deal of publicity should be punished severely, even if the crime was not severe, so that society knows there is a strong response.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

23. If someone commits a crime, that person is dangerous and must be removed from the community to protect other people.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

24. Punishment should be about looking forward to improve society, not looking backward to address the criminal's misdeeds.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

25. An eye for an eye makes the whole world blind.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

26. Even if society would not benefit at all from punishing a guilty person, he should still be punished because he deserves it.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

27. Even if society would benefit greatly from punishing an innocent person, he should not be punished because he does not deserve it.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

28. Criminals are in prison because they deserve to be there.

1 2 3 4 5
Strongly Disagree Neutral Strongly Agree

29. It is better to let ____ guilty people go unpunished than to punish 1 innocent person.

30. Assume that research has shown that punishment provides absolutely no benefit to the community. Criminals are not deterred by observing the punishment of other criminals, punishment does not reduce the likelihood that a criminal will re-offend, rehabilitation is ineffective and criminals are immediately replaced by new offenders. Are we still justified in punishing criminals?

1 2 3 4 5
Not at all Justified Neutral Very Justified

31. Please indicate whether or not you had a particular crime in mind while you were responding to the previous questions.

Yes (please specify) _____

No _____

APPENDIX D: THE BALANCED INVENTORY OF DESIRABLE RESPONDING

Please write a number beside each statement to indicate how much you agree with it on a scale of 1-7, where 1 = not true, 4 = somewhat true, and 7 = very true.

1. My first impressions of people usually turn out to be right.
2. It would be hard for me to break any of my bad habits.*
3. I don't care to know what other people really think of me.
4. I have not always been honest with myself.*
5. I always know why I like things.
6. When my emotions are used, it biases my thinking.*
7. Once I've made up my mind, other people can seldom change my opinion.
8. I am not a safe driver when I exceed the speed limit.*
9. I am fully in control of my own fate.
10. It's hard for me to shut off a disturbing thought.
11. I never regret my decision.
12. I sometimes lose out on things because I can't make up my mind soon enough.*
13. The reason I vote is because my vote can make a difference.
14. My parents were not always fair when they punished me.*
15. I am a completely rational person.
16. I rarely appreciate criticism.*
17. I am very confident of my judgments.
18. I have sometimes doubted my ability as a lover.
19. It's all right with me if some people happen to dislike me.
20. I don't always know the reasons why I do the things I do.*
21. I sometimes tell lies if I have to.*
22. I never cover up my mistakes.
23. There have been occasions when I have taken advantage of someone.*
24. I never swear.
25. I sometimes try to get even rather than forgive and forget.*
26. I always obey laws, even if I'm unlikely to get caught.
27. I have said something bad about a friend behind his or her back.*
28. When I hear people talking privately, I avoid listening.
29. I have received too much change from a salesperson without telling him or her.*
30. I always declare everything at customs.
31. When I was young I sometimes stole things.*
32. I have never dropped litter on the street.
33. I sometimes drive faster than the speed limit.*
34. I never read sexy books or magazines.
35. I have done things that I don't tell other people about.*
36. I never take things that don't belong to me.
37. I have taken sick-leave from work or school even though I wasn't really sick.*
38. I have never damaged a library book or store merchandise without reporting it.
39. I have some pretty awful habits.*
40. I don't gossip about other people's business.

*Item reverse-scored

APPENDIX E: DEMOGRAPHICS QUESTIONNAIRE

Finally, we would like to ask you a few questions about yourself, so that we can get a sense of who the people in our study are.

1. What is your age? _____

2. What is your gender?

___ Man

___ Woman

___ Trans

___ Other (please specify: _____)

3. What is your racial/ethnic background?

___ Asian

___ Black/African-American

___ Middle Eastern

___ East Indian

___ Hispanic/Latino

___ White/Caucasian

___ American Indian

___ Other (please specify: _____)

4. What is your occupation? _____

5. What is your highest level of education?

___ Doctoral or professional degree

___ Master's degree

___ Bachelor's degree

___ Associate's degree

___ Postsecondary non-degree award

___ Some college, no degree

___ High school diploma or equivalent

___ Less than high school

6. What (if applicable) is your religious affiliation? _____

7. Using the sliding scale below, please indicate where your political beliefs fall.

Liberal-----**Conservative**

8. What (if applicable) is your political affiliation? _____

APPENDIX F : PUNISHMENT VIGNETTES

- 1.) John parks his car illegally on a hill and forgets to set the emergency brake. While he is out and about, the car rolls down the hill and crashes into a daycare, killing 15 children and a daycare worker. Should John be punished for causing the deaths of these people? In one or two sentences, please describe the reason for your decision.
- 2.) There is a runaway trolley barreling down the railway tracks. Ahead on the tracks there are 5 track workers who are repairing the track. You are standing some distance away in the train yard next to a lever. If you pull the lever, the trolley will switch to another track. However, there is 1 worker repairing this track, who will not be able to get out of the way. Should you (a.) do nothing and the trolley will kill the 5 workers or (b.) pull the lever, derailing the trolley to a track where it will kill 1 worker. In one or two sentences, please describe the reason for your decision.
- 3.) Mary is the mayor of a small town. In response to a recent crime, a large subgroup of the town's population has begun to riot, calling out for the perpetrator to be executed. Mary knows that nobody has been arrested for the crime, and that the police have no leads as to who might have done it. However, she is aware that if nobody is brought forth, the subgroup will become violent, causing harm to hundreds or even thousands of townspeople. Mary identifies a local man who, although innocent of the crime, is universally disliked due to his unpleasant nature and low contribution to society. She knows that it would be very easy to accuse him of this crime, frame him, carry out a trial, and see him executed. Doing so would appease the mob, and is the only way to prevent the extensive damage and violence that is sure to occur if she is unable to produce a culprit. Should Mary frame this man? In one or two sentences, please describe the reason for your decision.
- 4.) Kevin and Bernard are twins. A recent string of sexual assaults has occurred, and DNA evidence proves that one or both twins are responsible. However, current advancements in DNA are unable to determine which twin is the assailant (or whether both twins have committed sexual assaults). No other evidence is available, as the victims are unable to tell the twins apart and indicate which one assaulted them. What should be done? In one or two sentences, please describe the reason for your decision.
- 5.) A researcher identifies a gene that causes individuals to commit violent crimes once they reach the age of 24. Though the crimes vary in severity, every single person who carries this gene will commit a violent crime on their 24th birthday. Should we imprison those who have the gene before they turn 24? In one or two sentences, please describe the reason for your decision.
- 6.) Sam and Tom sell illegal handguns for a living. Sam sells an illegal handgun to a recent immigrant who is unable to purchase a gun legally, who uses the gun to practice shooting cans in the forest. Tom sells an illegal handgun to a recent immigrant who is unable to purchase a gun legally, who uses the gun to murder his wife. Do Sam and Tom deserve equal punishment? In one or two sentences, please describe the reason for your

decision.

7.) Christopher is a convicted murderer on death row. He has a rare condition that allows him to donate multiple organs without dying, but he refuses to do so. Are we justified in harvesting his organs to save the lives of 8 sick innocent people? In one or two sentences, please describe the reason for your decision.

8.) Luke and Jeff are murderers. Luke has killed a local philanthropist, who was a father of four and much loved by the community. At the sentencing hearing in Luke's trial, the victim's wife testifies about the horrific impact of Luke's crime on her family and the community. Jeff has killed a local homeless person, who did not have any family members. As such, there was nobody to testify as to the impact of Jeff's crime at his sentencing hearing. Should Jeff and Luke receive the same punishment? In one or two sentences, please describe the reason for your decision.

9.) Dan is a prominent member of society. One day, his parents die of carbon monoxide poisoning. You are the only person alive who knows that Dan intentionally damaged the vents in his parents' house so that they would die, but you have no way to prove it and it would never go to trial. The public believes it to be an accident, and so Dan is never charged with murder. Because nobody believes a crime has been committed, there is nothing to gain from punishing Dan – future criminals won't be deterred, Dan himself will not re-offend since the only people he wanted dead were his parents, Dan is no longer a danger to society, and Dan does not need to be rehabilitated, since he will never commit another crime. Should you punish Dan yourself? In one or two sentences, please describe the reason for your decision.

10.) Assume that there exists a series of private islands, to which death row inmates are sent in lieu of execution. These inmates are each sent to their own island, so they have no communication with each other or the outside world. Because the world assumes that the inmates have been successfully executed, all of the goals of the death penalty are fulfilled – other would-be offenders are deterred to the extent that they believe they will be executed if caught, and the inmates themselves are incapacitated as they are unable to leave the island and re-enter society. However, the inmates are not actually punished on the island – they are able to live out their lives peacefully. Is this acceptable? In one or two sentences, please describe the reason for your decision.

APPENDIX G : SENTENCING GOALS SCALE

Listed below are a number of statements that describe attitudes that different people have about justice in the community. There are no right or wrong answers, only opinions. Read each item and decide whether you agree or disagree and to what extent. If you strongly disagree circle 1; if you strongly agree circle 7; if you feel somewhere in between circle any one of the numbers between 1 and 7. If you feel neutral or undecided, the midpoint is 4.

- (1) Disagree strongly (5) Agree slightly
- (2) Disagree (6) Agree
- (3) Disagree slightly (7) Agree strongly
- (4) Neither disagree nor agree

1. With the right approach most offenders can be rehabilitated back into society.

1	2	3	4	5	6	7
DISAGREE						AGREE

2. Justice is not done if the offender is not punished in some way.

1	2	3	4	5	6	7
DISAGREE						AGREE

3. The purpose of court sentences should be to protect society from the offender.

1	2	3	4	5	6	7
DISAGREE						AGREE

4. Crime rates would decrease if sentences were appropriately severe and Publicized more widely.

1	2	3	4	5	6	7
DISAGREE						AGREE

5. Prison sentences are useful because at least they don't allow criminals to re-offend.

1	2	3	4	5	6	7
DISAGREE						AGREE

6. Justice requires that the punishment should be as severe as the offence.

1	2	3	4	5	6	7
DISAGREE						AGREE

7. Repeat offenders should be given every opportunity to fit back into society.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

8. If the courts fail to punish criminals adequately potential offenders are not discouraged from committing similar offences.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

9. The purpose of punishment should be to make offenders pay for the wrongs that they have done.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

10. It is necessary for society to protect itself from the possibility that an offender might commit further offences.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

11. The purpose of court sentences should be to rehabilitate the criminal.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

12. Strict enforcement of the law (and its penalties) is necessary to prevent others from committing similar offences.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

13. Criminals should be punished for their crimes in order to make them repay their debt to society.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

14. Offenders should be locked away so that they can't re-offend.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

15. If I were the victim of a crime I would be satisfied even if the only effect of the offender's punishment was that the offender was eventually rehabilitated.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

16. Penalties should be severe enough so that criminals are unlikely to re-offend.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

17. Offenders should be punished to make them suffer as others have suffered.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

18. It is obvious from the increase in crime rates that penalties aren't severe enough.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

19. If judges would divert more people from prisons into rehabilitation programs, there would be less crime.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

20. Offenders must be punished so that they cannot cause any further harm to the community.

1 DISAGREE	2	3	4	5	6	7 AGREE
---------------	---	---	---	---	---	------------

Sentencing Goals

There are many possible objectives or goals in the sentencing of offenders. Below I have listed the five most commonly given goals, and for each of these goals I would like you to do two things. Could you:

1. First please **rank** these five in order of importance to you, from 1 (most important) to 5 (least important). For example, if you think that the most important reason is "to deter other potential offenders", put a 1 on the line next to that statement; if you think the next most important reason is "to protect the community...", put a 2 next to that statement, and so on.
1. After you have done that, could you please then **rate** how important you think each goal is on the 1 (*not at all important*) to 7 (*very important*) scales provided.

	RANK	Not at all important						Very important
To rehabilitate the offender	—	1	2	3	4	5	6	7

To deter other potential offenders	—	1	2	3	4	5	6	7
------------------------------------	---	---	---	---	---	---	---	---

Because the offender should “pay” in some way for what they have done	—	1	2	3	4	5	6	7
---	---	---	---	---	---	---	---	---

To deter the offender from committing similar offences in the future	—	1	2	3	4	5	6	7
--	---	---	---	---	---	---	---	---

To protect the community by placing the offender where they can do no harm	—	1	2	3	4	5	6	7
--	---	---	---	---	---	---	---	---

APPENDIX H: SCALE OF ATTITUDES TOWARD THE LEGAL SYSTEM

Each of the following reflects an opinion about the legal system and the courts. Please indicate how much you agree or disagree with each statement by circling the appropriate number below the statement. Use the following scale:

-3 = strongly disagree, -2 = moderately disagree,
-1 = slightly disagree, 0 = neutral, +1 = slightly agree,
+2 = moderately agree, +3 = strongly agree

1. The punishments given to criminals accurately reflect the crimes they have committed.
-3 -2 -1 0 +1 +2 +3
2. The skills of lawyers determine the verdicts more than the truth.
-3 -2 -1 0 +1 +2 +3
3. High standards of honesty and justice prevail in American courts.
-3 -2 -1 0 +1 +2 +3
4. A defendant accused of child molestation will receive a fair trial.
-3 -2 -1 0 +1 +2 +3
5. If accused of a crime, I feel confident that I would receive a fair trial.
-3 -2 -1 0 +1 +2 +3
6. Too many criminals are out on parole.
-3 -2 -1 0 +1 +2 +3
7. The media unfairly biases potential jurors against a defendant by publicizing information about him or her prior to trial.
-3 -2 -1 0 +1 +2 +3
8. Police brutality is more common than people think.
-3 -2 -1 0 +1 +2 +3
9. If a defendant has enough money he or she will be able to buy an acquittal.
-3 -2 -1 0 +1 +2 +3
10. As a group, judges are more impartial and fair than people in general.
-3 -2 -1 0 +1 +2 +3
11. The courts system is "color blind"; race of the defendant does not influence the outcome.
-3 -2 -1 0 +1 +2 +3
12. Juries base their decisions only on the evidence given in court.
-3 -2 -1 0 +1 +2 +3

13. Defendants who are guilty often “get off” because of technicalities.
 -3 -2 -1 0 +1 +2 +3
14. I think that plea bargaining is a reasonable way to serve justice, given the demands on the court’s time.
 -3 -2 -1 0 +1 +2 +3
15. Claims that the police have “planted” evidence are almost always made by guilty people.
 -3 -2 -1 0 +1 +2 +3
16. If you have a lot of money, you have a good chance of being able to “get off” even if you committed the crime.
 -3 -2 -1 0 +1 +2 +3
17. Too many criminals are let free because of prison overcrowding.
 -3 -2 -1 0 +1 +2 +3
18. Only those who committed a crime are ever convicted.
 -3 -2 -1 0 +1 +2 +3
19. Judges are more lenient in the sentences they give if the defendant is wealthy.
 -3 -2 -1 0 +1 +2 +3
20. Too many criminals slip through the cracks because of loopholes in the legal system.
 -3 -2 -1 0 +1 +2 +3
21. Jurors assume that a defendant is innocent until he or she is proven guilty.
 -3 -2 -1 0 +1 +2 +3
22. Too often, criminals are successful in pleading that they are not guilty by reason of insanity.
 -3 -2 -1 0 +1 +2 +3
23. Jurors are often intimidated by defendants who are known criminals.
 -3 -2 -1 0 +1 +2 +3
24. When a suspect confesses to the police, he or she does so voluntarily.
 -3 -2 -1 0 +1 +2 +3
25. Jurors are capable of accurately determining the innocence or guilt of a defendant.
 -3 -2 -1 0 +1 +2 +3
26. Since witnesses at a trial are under oath, you can assume they are telling the truth.
 -3 -2 -1 0 +1 +2 +3
27. The police do a good job of investigating crimes.

- | | -3 | -2 | -1 | 0 | +1 | +2 | +3 |
|--|----|----|----|---|----|----|----|
| 28. Court-appointed attorneys and public defenders do as good a job as personally hired attorneys. | -3 | -2 | -1 | 0 | +1 | +2 | +3 |
| 29. Defense attorneys waste too much time and energy helping people to get off easy. | -3 | -2 | -1 | 0 | +1 | +2 | +3 |
| 30. Parole boards let too many still-dangerous offenders out on parole. | -3 | -2 | -1 | 0 | +1 | +2 | +3 |
| 31. Pleas of insanity are often just ploys to get off easy. | -3 | -2 | -1 | 0 | +1 | +2 | +3 |
| 32. Police will often keep a suspect in custody even when they don't have any firm evidence against him. | -3 | -2 | -1 | 0 | +1 | +2 | +3 |
| 33. The Supreme Court is, by and large, an effective guardian of the Constitution. | -3 | -2 | -1 | 0 | +1 | +2 | +3 |
| 34. Upstanding citizens have nothing to fear from the police. | -3 | -2 | -1 | 0 | +1 | +2 | +3 |
| 35. Most prosecuting attorneys have a strong sadistic streak. | -3 | -2 | -1 | 0 | +1 | +2 | +3 |
| 36. All too often, minority group members do not get fair trials. | -3 | -2 | -1 | 0 | +1 | +2 | +3 |

APPENDIX I: NEED FOR COGNITION SCALE**1. I would prefer complex to simple problems.**

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

2. I like to have the responsibility of handling a situation that requires a lot of thinking.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

3. Thinking is not my idea of fun.*

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.*

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

5. I try to anticipate and avoid situations where there is likely chance I will have to think in depth about something.*

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

6. I find satisfaction in deliberating hard and for long hours.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

7. I only think as hard as I have to.*

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

8. I prefer to think about small, daily projects to long-term ones.*

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

9. I like tasks that require little thought once I've learned them.*

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

10. The idea of relying on thought to make my way to the top appeals to me.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

11. I really enjoy a task that involves coming up with new solutions to problems.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

12. Learning new ways to think doesn't excite me very much.*

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

13. I prefer my life to be filled with puzzles that I must solve.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

14. The notion of thinking abstractly is appealing to me.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

15. I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

16. I feel relief rather than satisfaction after completing a task that required a lot of mental effort.*

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

17. It's enough for me that something gets the job done; I don't care how or why it works.*

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

18. I usually end up deliberating about issues even when they do not affect me personally.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

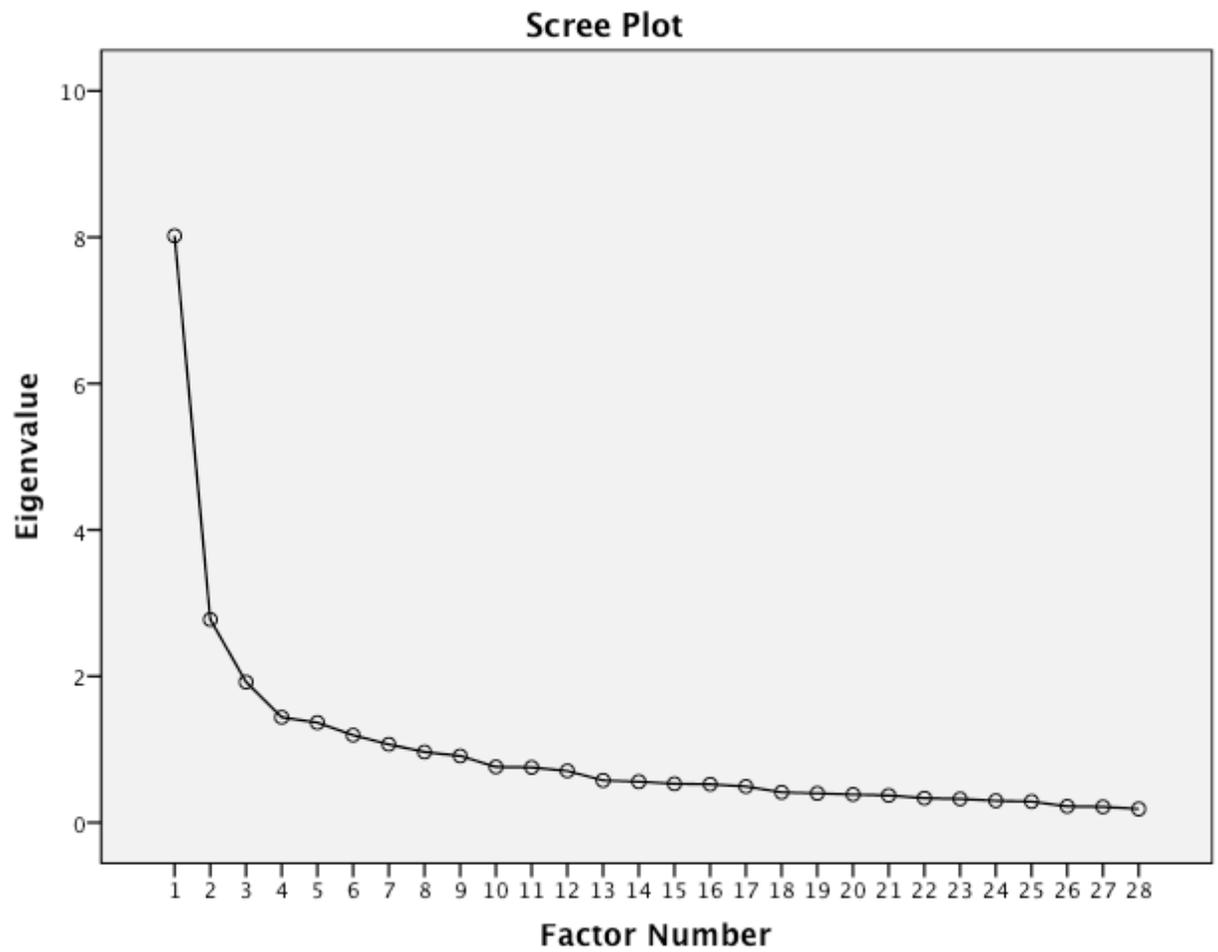
*Item reverse scored.

APPENDIX J: THE POSITIVE AND NEGATIVE AFFECT SCALES

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Indicate to what extent [you generally feel this way, that is, how you feel on average or you feel this way right now, that is at the present moment.]

Use the following scale to record your answers.

1	2	3	4	5
Very Slightly or Not at All		Moderately		Extremely
<input type="checkbox"/> interested				<input type="checkbox"/> ashamed
<input type="checkbox"/> distressed				<input type="checkbox"/> inspired
<input type="checkbox"/> excited				<input type="checkbox"/> nervous
<input type="checkbox"/> upset				<input type="checkbox"/> determined
<input type="checkbox"/> strong				<input type="checkbox"/> attentive
<input type="checkbox"/> guilty				<input type="checkbox"/> jittery
<input type="checkbox"/> irritable				<input type="checkbox"/> active
<input type="checkbox"/> alert				<input type="checkbox"/> afraid
<input type="checkbox"/> scared				<input type="checkbox"/> enthusiastic
<input type="checkbox"/> hostile				<input type="checkbox"/> proud

APPENDIX K: SCREE PLOT FROM INITIAL EXPLORATORY FACTOR**ANALYSIS**

APPENDIX L: POQ REVISED

1. Punishment is necessary because it restores the balance of justice. (HR)
2. The standard of reasonable doubt should be very high in criminal trials, to prevent innocent persons from being convicted. (IR)
3. If a crime has a high detection rate, it is not necessary to punish the crime harshly. (HU)
4. Criminals are bad people and deserve punishment. (HR)
5. Once a criminal has done his time, the system should ensure that he can reintegrate into the community so that society can benefit. (IU)
6. It is more important to keep innocent people free from punishment than it is to ensure that all guilty persons are punished for their crimes. (IR)
7. It is better to let one dangerous criminal go unpunished than it is to punish an innocent person. (IR)
8. It is more important to punish a guilty person because he deserves it than it is to punish him to benefit society. (HR)
9. If someone steals from a store, that person should have to wear a sandwich board outside of the store indicating what was stolen. (HU)
10. Punishment is not about “an eye for an eye” – if we are to harm a person using punishment, there must be some benefit to the community. (IU)
11. Criminals are in prison to keep the community safe. (HU)
12. It is better to let 10 guilty criminals go free than to punish 1 innocent person. (IR)
13. Criminals deserve to be publicly identified and shamed. (HR)
14. If a person hits someone with the intent to slightly harm him, but causes significant harm, he should be punished for the significant harm that he caused. (HR)
15. When a person commits a crime, he upsets the balance of justice. The only way to make this right is to punish him accordingly. (HR)
16. Punishment should match the crime- if you commit a serious crime, you should receive a serious punishment. (HR)
17. If a crime has a low detection rate (i.e. it is difficult to catch criminals who commit this particular crime), we should punish those who are caught harshly to prevent others from thinking they can get away with it. (HU)
18. Crimes that receive a great deal of publicity should be punished severely, even if the crime was not severe, so that society knows there is a strong response. (HU)
19. The goal of punishment should be to give the guilty person what he deserves. (HR)
20. If a crime receives little publicity, and is relative unknown to have occurred, it is not necessary to punish the crime harshly. (HU)
21. If someone commits a crime, that person is dangerous and must be removed from the community to protect other people. (HU)
22. ‘Innocent until proven guilty’ is essential to the functioning of the criminal justice system. (IR)
23. Even if society would not benefit at all from punishing a guilty person, he should still be punished because he deserves it. (HR)
24. Criminals are in prison because they deserve to be there. (HR)
25. An overly harsh punishment may be necessary to prevent others from committing the same crime. (HU)

26. Punishment should be about looking forward to improves society, not backward to address the criminal's misdeeds. (IU)

(IR) = Ideal Retributivism (HR) = Harsh Retributivism (IU) = Ideal Utilitarianism (HU) = Harsh Utilitarianism

APPENDIX M: FINAL POQ BY SUBSCALE**Ideal Retributivism Subscale**

1. It is better to let 10 guilty criminals go free than to punish 1 innocent person.
2. It is more important to keep innocent people free from punishment than it is to ensure that all guilty persons are punished for their crimes.
3. It is better to let one dangerous criminal go unpunished than it is to punish an innocent person.

Harsh Retributivism Subscale

1. Punishment is necessary because it restores the balance of justice.
2. It is more important to punish a guilty person because he deserves it than it is to punish him to benefit society.
3. The goal of punishment should be to give the guilty person what he deserves.
4. Criminals are bad people and deserve punishment.
5. Even if society would not benefit at all from punishing a guilty person, he should still be punished because he deserves it.
6. Criminals are in prison because they deserve to be there.
7. Criminals deserve to be publicly identified and shamed.
8. If a person hits someone with the intent to slightly harm him, but causes significant harm, he should be punished for the significant harm that he caused.

Harsh Utilitarianism Subscale

1. An overly harsh punishment may be necessary to prevent others from committing the same crime.
2. If a crime has a low detection rate (i.e. it is difficult to catch criminals who commit this particular crime), we should punish those who are caught harshly to prevent others from thinking they can get away with it.
3. Crimes that receive a great deal of publicity should be punished severely, even if the crime was not severe, so that society knows there is a strong response.
4. If someone commits a crime, that person is dangerous and must be removed from the community to protect other people.
5. Criminals are in prison to keep the community safe.

APPENDIX N: DESCRIPTIVES FOR MEASURES FROM STUDY 2

	Variable	<i>Mean</i>	<i>SD</i>
POQ	Innocents	3.57	1.05
	Retributivism	3.34	.78
	Utilitarianism	3.16	.84
SGS	Retribution	4.89	1.43
	Rehabilitation	4.67	1.54
	Incapacitation	4.98	1.35
	Deterrence	4.81	1.35
ATLS	System Works	-5.04	15.29
	System Too Lenient	4.12	9.42
PANAS	Positive Affect	3.14	.90
	Negative Affect	1.58	.68
BIDR	Impression Management	6.57	2.72
	Self Deceptive Enhancement	5.49	3.14
	Need for Cognition	3.70	.83

APPENDIX O: FINAL SCENARIO DECISION AND JUSTIFICATION FREQUENCIES

Frequencies for decision by scenario

	Decision	
Parked Car	Do not punish: 17 Mitigated punishment: 81 Severe punishment: 8 Unqualified punishment: 95	Unrelated: 0
Trolley	Pull lever: 155 Do nothing: 44	Unrelated: 2
The Mob	Frame the man: 19 Do not frame the man: 183	Unrelated: 5
Twins	Imprison twins: 54 Do not imprison: 135	Unrelated: 12
Organs	Take organs: 53 Do not take organs: 144	Unrelated: 4
Secret Murder	Do not personally punish: 152 Personally punish: 49	Unrelated: 0
Isolated Island	Acceptable: 79 Unacceptable: 113	Unrelated: 9

Frequency of retributive or utilitarian justifications

	Retributive		Utilitarian		
	<i>Present</i>	<i>Absent</i>	<i>Present</i>	<i>Absent</i>	<i>Both Present</i>
Parked Car	194	7	14	187	7
Trolley	162	39	144	57	18
The Mob	167	34	42	159	22
Twins	124	77	31	170	12
Organs	162	39	44	157	20
Secret Murder	97	104	49	152	14
Isolated Island	126	75	64	137	19

Note: Scenarios which could not be analyzed are not displayed.