

AMERICAN PUNITIVENESS AND MASS INCARCERATION: PSYCHOLOGICAL PERSPECTIVES ON RETRIBUTIVE AND CONSEQUENTIALIST RESPONSES TO CRIME*

Mark R. Fondacaro** and Megan J. O'Toole***

A recent National Academy of Sciences Report entitled, "The growth of incarceration in the United States: Exploring causes and consequences," examined the drivers of the fourfold increase in incarceration rates in the United States and provided a firm recommendation for significant reduction in incarceration rates (Travis, Western, & Redburn, 2014). Policy makers representing the entire political spectrum are now publicly airing their views on the need for reform. Although public sentiment is generally favorably disposed toward reform in the abstract, when confronted with specific examples of crime, they tend to favor more

*The authors would like to express their gratitude to Christopher Slobogin for feedback on an earlier draft of this article and to Joon Won Yoon for his diligent research assistance. This article is based on an invited discussion paper prepared for the Interdisciplinary Roundtable on Punitiveness in America, which was sponsored by the MacArthur Foundation and John Jay College of Criminal Justice, City University of New York, April 2–3, 2015.

**Mark R. Fondacaro is a Professor and Director of the Psychology & Law Doctoral Training Area at the CUNY Graduate Center and John Jay College of Criminal Justice. He received his Ph.D. in clinical psychology from Indiana University–Bloomington and his J.D. from Columbia University School of Law. He completed postdoctoral studies in social ecology in the Department of Psychiatry and Behavioral Sciences at Stanford Medical School. His major research interests center on procedural justice, behavioral science and criminal responsibility, and reform of the juvenile and criminal justice systems.

***Megan J. O'Toole is a doctoral candidate in the Psychology and Law Program at the CUNY Graduate Center and John Jay College of Criminal Justice, and received her B.A. from Wheaton College. She is currently a Graduate Research Fellow in John Jay's Research and Evaluation Center, where she advises New York City juvenile justice agencies in best practices and evidence generation. Her major research interests include public attitudes toward justice system reform, evidence-based justice system practices, and program evaluation.

New Criminal Law Review, Vol. 18, Number 4, pps 477–509. ISSN 1933-4192, electronic ISSN 1933-4206. © 2015 by the Regents of the University of California. All rights reserved. Please direct all requests for permission to photocopy or reproduce article content through the University of California Press's Reprints and Permissions web page, <http://www.ucpress.edu/journals.php?p=reprints>. DOI: 10.1525/nclr.2015.18.4.477.

punitive, retributive responses to crime. Retributive justifications for punishment that are deeply ingrained in our culture and our legal system, as well as our biological and psychological make-up, are a major impediment to constructive reform efforts. However, recent advances in research across neurobiological, psychological, and social levels of analysis suggest that following our retributive impulses to guide legal decision making and criminal justice policy is not only costly and ineffective in reducing crime, but unjust and increasingly difficult to justify morally. This article will draw on a body of research anchored in social ecological models of human behavior to argue for more forward-looking, consequentialist responses to crime that aim at the individual prevention of criminal behavior in the least restrictive and most cost-effective manner at both the front and back ends of our criminal justice system.

Keywords: *mass incarceration, retribution, consequentialism, punishment, psychology*

INTRODUCTION

The goal of this article is to provide a broad interdisciplinary audience, including legal scholars, not only a better understanding of how intuitive psychological processes that are largely automatic, unconscious, and biologically based fuel our desire to punish, but also how our predominant reliance on a commonsense folk psychology frame of reference for assigning legal responsibility in the criminal law can fuel harsh punishment and have the paradoxical effect of promoting criminal behavior.¹ Moreover,

1. This article is based on a discussion paper that was presented at an Interdisciplinary Roundtable on Punitiveness in American that was funded by the MacArthur Foundation and hosted by John Jay College of Criminal Justice, City University of New York. The Roundtable was an outgrowth of the recent National Academy of Sciences (NAS) report that explored the drivers of the fourfold increase in incarceration rates in the United States, including sentencing policy, race and incarceration, the war on drugs, individual and community effects of prison, and the effect of incarceration on crime rates (Travis, Western, & Redburn, 2014). The NAS report provided a firm recommendation for significant reduction in incarceration rates. The Roundtable was convened to explore interdisciplinary explanations of the harshness of American criminal justice policy and to consider what might be done to reduce undue reliance on punitive policies. The primary goals were to foster innovative interdisciplinary collaboration and to determine whether and how sustained exploration of these topics within and across disciplines might foster constructive changes in criminal justice policy. The Roundtable presentation on which this article is

by broadening the focus from the psychology of why we punish others to encompass questions of what types of sanctions have what types of consequences for what types of offenders (and their victims, families, communities, and society at large), the chances of facilitating the goal of reducing incarceration rates are much improved. This will require a willingness to incorporate into criminal law empirical evidence that challenges moral intuitions and folk psychology concepts of culpability and just punishment when they are inaccurate or biased in ways that lead to unjust outcomes, much as psychologists have begun to do in the areas of eyewitness identification and false confessions. Criminal law also needs to broaden the focus for judging criminal responsibility and assigning punishment beyond retrospective social judgments about what a defendant was or was not thinking at the time of the alleged crime to a more contextual understanding of the interrelated biological, psychological, and social factors that drive both criminal behavior and compliance with the law. Finally, at the broadest level, the article suggests the promise of analogizing from the social-cognitive problem-solving models that psychologists have successfully developed to address retaliatory and instrumental aggression at the individual level, and bringing them up to scale at the policy level to help diminish the unreflective demand by the public and policymakers for unduly harsh punishment in the American criminal justice system.

The article begins with an overview of the emergence of distinct definitions and conceptualizations of punishment in the field of psychology, outlining the historical and contemporary role of consequentialist and retributive perspectives. Section I discusses the influence of early behavioral and cognitive subdisciplines of psychology on the conceptualization of punishment. These early influences, along with later influences, are then mapped onto retributive and consequentialist perspectives. This section lays the foundation for the argument that the early anchoring and later realignment of the criminal law with retributive justifications for punishment is increasingly out of step with recent scientific evidence on the causes, consequences, and prevention of criminal behavior, and has provided moral cover for unjust, unnecessary, ineffective, and unduly harsh

based focused on some of the ways in which the field of psychology has contributed to understanding the nature and consequences of various justifications for punishment in the American criminal justice system, and how the field might contribute to future collaborative research and policy reform.

criminal sanctions. Section II focuses on the question of why we punish offenders. This section highlights how the criminal law's narrow, myopic focus on *mens rea* and judgments of past mental state as a primary basis for judging criminal responsibility tracks common sense and folk psychology but reinforces and legitimizes short-sighted, biologically based responses to crime that are guided more by intuition than reason and can fuel harsh and ineffective punishment. Section III focuses on whom we punish, with an emphasis on whom we punish most often, most harshly, and why. The section examines the cumulative effects of implicit racial biases on decision making at various points in the legal system and how it contributes to disproportionate minority confinement and mass incarceration. Section IV focuses on the individual and social consequences of punishment for the punisher and the punished, and reviews empirical evidence indicating that the "nothing works" mantra that contributed to the rebirth of retribution in the criminal law has been put to rest by effective evidence-based interventions with both juveniles and adults. Section V provides suggestions for future research and policy reform, followed by some concluding remarks.

I. PSYCHOLOGICAL PERSPECTIVES ON PUNISHMENT

Since its inception, the discipline of psychology has reflected a dynamic tension between those who have emphasized the internal, subjective experience of the individual and those who have focused on the determinants and consequences of behavior. This distinction is reflected in the two primary orientations toward punishment in the field of psychology today: consequentialism and retribution.

Consequentialism has its early roots in the behavioral paradigm of Skinner (1965), which defines punishment as any consequence that follows an individual's behavior and decreases the likelihood that the behavior will recur in the future. According to this view, whether a response to a person's behavior serves as an effective punishment is determined solely on the basis of whether the person no longer behaves in the same manner in the future. More recent advocates of a consequentialist perspective suggest that behavior is determined by a confluence of social, psychological, and biological factors, and that punishment should be forward looking and focus on whether it has an effect on reducing the likelihood of criminal behavior in the future (Fondacaro, 2014; Greene & Cohen, 2004; Slobogin & Fondacaro, 2011).

The retributive justification for punishment has its roots in religion, philosophy, and common sense, and maps onto more intrapsychic and cognitive models of human behavior. A recent manifestation of a retributive model in the field of psychology is reflected in the empirical and conceptual work of Darley, Carlsmith, and Robinson (Carlsmith et al., 2002; Darley et al., 2000). The retributive model is squarely aligned with traditional legal models of criminal responsibility and just deserts. Punishment is viewed as a morally required response to wrongdoing, which serves to make the offender pay or suffer for the harm caused. Moreover, in its philosophically pure, deontological form, retributive punishment must be proportionate to the harm caused—no more and no less. Whether or not punishment has any effect on future behavior is incidental from a retributive standpoint. The justification for punishment is based on the presumption that people act for reasons and can consciously choose whether or not to obey the law (Morse, 2000). As a result, when they engage in criminal behavior it is presumed that they could have chosen to do otherwise and therefore deserve to be punished as payback for their moral failing.

In an early and influential article on the social psychology of punishment reactions, Miller and Vidmar (1981) noted that when the fields of psychology and other social sciences have attempted to define punishment in the context of the legal system, they typically have done so in confusing ways. However, they noted that regardless of whether the primary justification for punishment is crime prevention or redress of a past injustice, the methods used to deliver punishment often involve elements of both consequentialist and retributive motivations. More recently, Keller, Oswald, Stucki, and Gollwitzer (2010) have presented a more comprehensive and refined framework that conceptualizes justifications for punishment as follows:

Punishment justifications can be differentiated beyond the dichotomy of retribution and utilitarianism. Dimensions of classification refer to (a) which social entity the punishment is directed at (i.e., the perpetrator, the victim, or the general public), (b) whether the sanction focuses rather on negative or rather on positive aspects, (c) whether the sanction is backward-oriented (i.e., retributive in nature) or forward-oriented (i.e., utilitarian in nature). Retribution, for example, mainly aims at the perpetrator, focuses on negative aspects, and is largely backward-oriented. Special prevention aims at the perpetrator, but is forward-oriented, and can imply both negative aspects (deterrence) and positive aspects (i.e., rehabilitation). General prevention is

forward-oriented and includes negative and positive aspects, but aims at the general public. Incapacitation aims at the perpetrator, is forward-oriented, and implies a negative sanctioning form (such as incarceration). (p. 100; internal citations omitted)

II. PSYCHOLOGICAL RESEARCH ON WHY WE PUNISH OTHERS

The answer to the question of why we punish others depends to a large extent on where and at what level of analysis we look. Since the law and commonsense notions of folk psychology start with psychological functioning and presumed motivations of the individual, the analysis will start there. However, many different and more complex explanations for human behavior become increasingly evident as the psychological functioning of the individual begins to be placed at the center of the dynamic factors that influence individual behavior within and across social, psychological, and biological levels of analysis. Community and developmental psychologists have referred to this broader frame of reference as a social ecological model of human behavior (Bronfenbrenner, 1979; Moos, 1973). More recently, leading researchers from the subdiscipline of psychology and law have urged their colleagues to argue for more awareness of contextual influences on human behavior when addressing issues of law and policy (Darley, Fulero, Haney, & Tyler, 2002).

The American legal system encompasses both utilitarian and retributive justifications for punishment but is most deeply grounded in principles of retribution, just desert, and an evil-doer theory of crime. Generally, individuals are held criminally liable when they commit a voluntary act that was actuated by a guilty mind. Researchers who have studied punishment from a retributive standpoint have characterized the desire to punish as the result of a dual process mechanism, which is triggered initially by intuitive, largely unconscious, emotion-based processes and then loosely regulated by top-down rational, cognitive processes (Carlsmith & Darley, 2008). According to this view, people have a general desire to punish others who have committed a harmful and immoral act, and feel that justice is served when the person receives just desert or proportionate punishment for the harm they caused. The consequences of the punishment are largely incidental to the function of restoring the moral imbalance by having the person pay

back their debt for the intentional harm caused with the currency of personal suffering (Kaplan, Weisberg, & Binder, 2000). In the American system of criminal justice, the prototypic symbol for and means of delivering personal suffering is incarceration.

The retributive paradigm of criminal responsibility has its heuristic roots in what Judge Hoffman (2014) refers to in *The Punisher's Brain* as second-party punishment—that is, when one person perceives they have been harmed by another person, they respond in a retaliatory way. When individuals anticipate responding to personal harm with punishment, they experience activation of the dorsal striatum, a subcortical region of the brain associated with anticipation of reward and satisfaction (de Quervain et al., 2004; Knutson, 2004). Even when they are not the direct victims of harm but are punishing as a third-party for harm that was committed against someone else, they experience stimulation of reward centers of the brain (Strobel et al., 2011). Although investigations of the neural correlates of such third-party punishment indicate that the magnitude of punishment is linked to activation of brain regions associated with affective processing, including the amygdala and medial prefrontal and posterior cingulate cortex, determination of criminal responsibility is associated with activation of the prefrontal cortex, suggesting the potential influence of rational, deliberative processes (Buckholz et al., 2008). Overall, second-party and third-party punishment appear to be supported by common neural mechanisms.

In dealing with the question of why we punish, many behavioral scientists who study this issue refer back to the legal and philosophical definitions of retribution, which reflect the deontological view that consequences do not matter and that punishment must be proportionate to what is deserved, no more and no less. Psychologists who have embraced these foundational definitions and attempted to address the issue empirically have pitted punisher preferences for retributive versus utilitarian justifications for punishment such as deterrence and incapacitation, and generally found a preference for retributive justifications (Carlsmith, 2006; Carlsmith & Darley, 2008). These retributive preferences, as operationalized by the investigators, typically have been interpreted as validation of the fundamental, retributive nature of the motive to punish (Darley, 2009). This view is then extrapolated to specific public policy preferences for retributive legal sanctions (Carlsmith & Darley, 2008). Although this body of research has made valuable and seminal contributions to the understanding of the human motivation to punish, there are several limitations of this

line of research, analysis, and application that deserve additional empirical investigation. These issues are outlined briefly below.

First, the authors anchor their work in deontological definitions of retribution. However, in their operationalization and efforts to validate the construct, they do not adequately assess or probe the degree to which the construct they are dealing with goes beyond the bounds of proportionate punishment. For example, they fail to assess whether in the minds of the punishers just desert goes beyond proportionate punishment to issues of revenge, retaliation, and infliction of suffering and abuse, particularly toward those who are perceived as outsiders (e.g., criminal evil doers, minorities). The need to address this issue empirically is particularly salient as we move from description to policy recommendations, since whether and to what extent criminal justice policy should track specific public sentiments and moral intuitions may be influenced by whether those sentiments map only onto proportionate punishment or encompass revenge and gratuitous harm and physical and psychological abuse as well (Haney, 2006; Kolber, 2012).

Second, efforts to compare and contrast preferences for utilitarian versus retributive justifications for punishment do not sample the full range of consequential responses to crime, focusing narrowly on incapacitation and deterrence, failing to separate out general and specific deterrence, and using incomplete operationalizations of the utilitarian justifications for punishment under investigation. Typically missing from the analysis is any consideration of treatment, especially more recent evidence-based, multi-systemic risk management models of intervention that have shown considerable promise in preventing recidivism (Slobogin & Fondacaro, 2011).

For example, in an early empirical investigation of the relative importance of just deserts versus utilitarian justifications for punishment, Darley, Carlsmith, and Robinson (2000) limited their examination of utilitarian motivations to incapacitation, which in turn was operationalized narrowly based on information about whether the defendant had a past history of a similar or related offense. The results indicated that participants generally increased their punitiveness in response to moral seriousness but not likelihood of future harm, and the researchers concluded that just deserts was the primary motivation for punishment for intentionally committed crimes. However, they did not assess the full range of potential consequentialist justifications for punishment, nor did they comprehensively assess the construct of incapacitation. Clearly, even within the narrow limits of

incapacitation as the representative utilitarian justification for punishment, a wider range of psychological, social, and biological factors might be of relevance to lay decision makers regarding judgments of risk assessment and the prevention of future harm.

In a related study, Carlsmith, Darley, and Robinson (2002) again pitted retributive and consequentialist justifications for punishment against each other, this time focusing on deterrence as the utilitarian justification. In their initial validation study, they presented Princeton undergraduates with definitions of just deserts and utilitarian justifications taken from standard philosophical texts, and had them classify various pieces of information as indictors of either just desert philosophy (magnitude of harm, extenuating circumstances) or deterrence philosophy (detection rate, publicity). Then, in a series of experimental studies, they had participants rate the severity of punishment they thought was warranted in response to various situations involving *intentionally* committed crimes that varied along dimensions of just desert and deterrence. They found that the severity of recommended sentences was more strongly associated with the just desert–related factors than the deterrence-oriented ones. They also found that, despite the fact that participants expressly endorsed deterrence theory, their underlying motivations for punishment revealed sensitivity to factors associated with just desert and insensitivity to factors related to deterrence. They attempted to explain their findings by drawing on the distinction made by Brickman, Folger, Goode, and Schul (1981) between macro-level and micro-level justice. This distinction suggests that when people are asked to make judgments about what should happen in response to crime in general, they resort to more utilitarian concerns. However, when they are asked to resolve a specific case, they rely on just desert-based considerations to avoid imposing undeserved punishment.

The authors failed to consider the extent to which their results were influenced by the way they defined and operationalized the various justifications for punishment. Note that the utilitarian justification was limited to general deterrence operationalized in terms of likelihood of detection and how publicly visible the punishment would be. This operationalization obviously omitted specific deterrence, rehabilitation, and restorative justice as consequential responses to crime that may be likely to elicit more micro-analytic and context sensitive judgments about punishment. Moreover, Carlsmith and colleagues (2002) defined “just deserts” as a sentence “that society assigns as the just punishment that the criminal *deserves* for the wrong

that he did. ‘Wrong’ is indexed by both severity and context of the crime” (p. 291). It should not be surprising then that severity of punishment should be highly correlated with just deserts defined as severity of the crime. Moreover, all of the scenarios in this study and the prior study by Darley and colleagues (2000) involved crimes that were described as unequivocally intentional. In actual criminal trials, the issue of intentionality and mental state is typically the key issue in dispute and is tied to degree of criminal responsibility. Any evaluation of the inherent desirability or justice of lay notions of retribution would seemingly turn at least in part on the accuracy of lay judgments that a criminal act was committed in a manner that was voluntary and intentional.

In contrast to deontological justifications for punishment, with retribution anchored in just desert as the prototypic example, consequentialist responses to crime justify punishment based on their ability to prevent repeat and future crimes and to promote specific public policy objectives. Although typical ideological and philosophical debates have contrasted retributivism with utilitarianism, a consequentialist position broadens the discussion and adds breadth and depth to the typical characterization of the scope and limits of utilitarianism. So, for example, deontologists often characterize the utilitarian position as one where anything goes, as long as in an aggregate cost-benefit analysis, there is a net social benefit (Carlsmith & Darley, 2008). Although minor consideration is sometimes paid to utilitarian methods and goals of specific deterrence and rehabilitation, the primary discussion typically centers on general deterrence and incapacitation. Discussion of the utilitarian view sometimes begins with a cramped, caricatured straw man characterization of the position, followed by its application in extreme and unrealistic examples—such as punishing random or innocent individuals for the purpose of general deterrence or extreme over-punishment, such as lengthy imprisonment or the death penalty for offenses such as petty theft. Although a thorough and complete critical analysis of this position is beyond the scope of this article, even the most fervent utilitarians, especially those armed with any statistical and empirical sophistication, would presumably reject these extreme practices on the grounds that the net social demoralization and loss of political legitimacy would far outweigh any net social benefits derived.

In any case, notwithstanding any real or misconstrued limitations of utilitarianism, consequentialists can draw on a wide range of methods to promote recidivism reduction and crime prevention. For example, with

respect to general deterrence, the consequentialist approach goes beyond the criminal justice system as a means of promoting general compliance with legal rules and the maintenance of social stability. The promotion of positive commitment to the rule of law and responsible citizenship is seen as the outcome of a wide range of interconnected socializing institutions such as family, education, work, neighborhood, community, state, nation, and beyond (Carlsmith et al., 2002; Fondacaro & Weinberg, 2002; Fondacaro, 2011). However, within the realm and realistic reach of the criminal justice system, the primary goal is the individual prevention of future criminal behavior among those who have already violated the law by drawing on principles and practices of specific deterrence, rehabilitation, risk assessment and management, restorative justice, and incapacitation as a last resort (Slobogin & Fondacaro, 2011). Although public awareness of punitive sanctions for criminal behavior may serve as a proactive, general deterrent to crime in the traditional sense, such consequences are incidental or secondary to the goal of having the criminal justice system respond to law violators in a way that will decrease the likelihood that specific offenders will continue to engage in criminal behavior or victimize others in the future. As noted above, the proactive prevention of crime and promotion of abiding by the law among the general public is seen as being driven primarily by social institutions and social conditions whose influences far outweigh those of any potential legal threat of severe punishment.

One of the often overlooked strengths of the consequentialist position is that it provides a metric and lever for promoting specific public policy objectives. So, for example, if broadly speaking, your primary policy goal in the criminal justice system is to reduce recidivism, then you can explicitly measure the extent to which the particular sanction or intervention strategy is advancing that goal. Moreover, there is no necessity of any pre-commitment to a particular policy or intervention strategy—it becomes an empirical question, one that can be addressed by comparing any policy or sanction to an alternative or baseline condition—the prototypic baseline being the status quo. If the question becomes, what is the most effective way to reduce the likelihood that someone who has committed a particular crime will do so in the future when they are no longer under the jurisdiction of the criminal justice system, we can begin with the usual prison sanction as a baseline starting point and then compare that, in an incremental way, with alternative responses such as longer or shorter prison sentences, rehabilitation, supervised probation, restitution, restorative justice, and so on. There is

also no required pre-commitment to a single policy goal; it is possible to compare interventions and sanctions on multiple outcome criteria including costs, public approval, and collateral benefits and damage.

In its deontological form, retribution immunizes itself from such comparative, empirical, and critical analysis. Some empirically oriented researchers sympathetic to retributive justifications for punishment have seemingly recognized this limitation and have jettisoned deontology for what they have termed “empirical desert.” According to Robinson and Darley (2007), empirical desert is based on the notion that human beings have an intuitive sense of justice that requires retributive punishment for criminal behavior, and that the amount of punishment required can be determined empirically through assessment of public preferences. As has been pointed out by Slobogin and Brinkley-Rubinstein (2013), this is a giant step forward from the retributivist tradition because it places retribution squarely on the consequentialist playing field. The empirical desert position now provides a basis and metric for comparison to alternative justifications for punishment and responses to criminal behavior. Even among some of the staunchest retributivists, there has been an apparent implicit assumption that adherence to principles of retribution serves to maintain social stability and the moral order and legitimacy of the legal system. It now becomes an empirical question whether retributive punishment is actually the most effective means of promoting these policy objectives. It also places squarely on the table questions about whether there is a universal moral consensus on what constitutes just desert and just punishment, and whether different subgroups within society, particularly those from minority groups disproportionately represented in American prisons, share that consensus.

In an initial attempt to subject empirical desert theory to empirical scrutiny, Slobogin and Brinkley-Rubinstein (2013) examined three general hypotheses underlying the theory: (1) whether there is consensus about the ordinal ranking of crimes and the appropriate punishments; (2) whether dissatisfaction with the law promotes noncompliance, and what type of dissatisfaction promotes the most noncompliance; and (3) whether assessments of desert are affected by preventive considerations, and to what extent laypeople prefer desert-based dispositions over prevention. Based on the results of a series of seven studies, they concluded the following:

- (1) while consensus on the ordinal ranking of traditional crimes is relatively strong, agreement about appropriate punishments—which arguably is the

type of agreement empirical desert requires in order to work—is weak; (2) the relationship between people’s willingness to abide by the law and the law’s congruence with their beliefs about appropriate punishment is complex and not necessarily positive; further, any noncompliance that results from the law’s failure to reflect lay views about desert is probably no greater than the noncompliance triggered by a failure to follow lay views about the role utilitarian goals should play in fashioning criminal dispositions; (3) while the relative crime control benefits of a desert-based system and a prevention-based system are hard to evaluate (and are not directly examined here), people are willing to depart from desert in cases that do not involve the most serious crimes if they believe that preventive goals can be achieved in some other way. (p. 77)

In an earlier attempt to identify the complexity of punishment goals, Weiner, Graham, and Reyna (1997) used causal attribution theory to examine retributive and consequential goals of punishment and the various methods of attempting to achieve those ends. The aim of their study was to examine the cognitive and affective processes that guide punishment decisions in response to criminal behavior. They placed subjects in the position of assuming the role of a judge in determining the punishment for a defendant accused of murder. They found that retributive desires to punish the specific defendant were related to causal attributions of personal control and responsibility, which were tied to emotional reactions of anger and low sympathy, as well as to general philosophical beliefs about retributive punishment. Alternatively, the goal of specifically deterring the defendant from committing future crimes was related to attributions of causal stability, which related to expectancies about future crime. Specific deterrence was also predicted by a general philosophical orientation toward utilitarianism. Emotional reactions were not significant mediators in the utilitarian context. Both general philosophical orientations to punishment were predictive of severity of punishment, but only through the context-specific judgments of the defendant. In terms of punishment severity, the descending order of influence of the various factors examined was “expectations about future criminal acts ($B = .248$), inferences of responsibility ($B = .193$), feelings of sympathy toward the accused ($B = -.187$), and endorsement of specific retributive goals ($B = .170$)” (Weiner et al., 1997, p. 445).

Overall, Weiner and colleagues (1997) concluded that punishment goals are state-like rather than enduring intrapersonal traits and are influenced by the situational context of criminal behavior. They also found that the

utilitarian methods of both rehabilitation and protection of society were related to attributions of causal stability but in opposite ways: instability was negatively related to societal protection-oriented means but positively related to rehabilitation, suggesting the inadequacy of a unidimensional view of utilitarian objectives. After noting shortcomings of both retribution and general deterrence, they concluded: “Rehabilitation has less in common with other utilitarian objectives than does retribution. This again suggests that the grouping of utilitarian objectives may not be a useful way of thinking, and that a different taxonomy of philosophies of punishment is needed. Rehabilitation, rather than the more embracing utilitarianism, may emerge as the ‘highest’ moral position” (Weiner et al., 1997, p. 450).

III. WHOM DO WE PUNISH?

To begin to understand the depth and scope of punishment in America from a psychological standpoint, it helps again to begin the analysis at the individual, psychological level on the issue of whom we punish. In fact, the American penchant for individualism and myopic focus on attributions of personal responsibility may, paradoxically to some, help to fuel the intensity and scope of punishment in the American criminal justice system (Fondacaro, 2011; Weiner, 1995; Weiner et al., 1997). To briefly reiterate what contributors to the Roundtable have said about the legal basis for criminal responsibility in America, there are two primary requirements: a *mens rea* that occurs concurrently with and actuates an *actus reus*. In more common parlance, criminal responsibility is the result of a guilty mind that sets into motion a bad act. Psychological theory and research conducted over the past century provides some insight into how each of the separate but interrelated links in this retrospective social judgment process amplifies the likelihood that the individual will be held criminally responsible and punished relatively severely. Moreover, defendants with certain legally irrelevant personal characteristics may receive particularly harsh judgments and sanctions.

The links of this decision-making chain will be taken up individually, and then the cumulative effects and systemic outcomes of this process will be examined. The first link in this chain is the focus on *mens rea* or the mental state of the individual to judge culpability. Here, the focus is on

determining what the person was or was not thinking at the time of the crime. Did the person rationally and subjectively consider engaging in the illegal behavior? If they did so, they would be judged criminally culpable for a related criminal act. Those judged to have thought about purposely causing the prohibited outcome would be judged more culpable than those who were judged to be consciously aware of a substantial and unjustifiable risk of the harmful outcome. The crux of the analysis focuses on conscious awareness and matters of degree. A person who lacked any conscious awareness of the harmful outcome and was not careless or negligent would not be judged criminally culpable for accidental behavior. The task here is to determine the person's mental state at the time of the crime.

The focus here is almost entirely cognitive. What was the person thinking at the time of the crime? Why did they do what they did, and were they morally responsible for their behavior? Perceptions of intentionality carry great weight in such judgments, and this all conforms nicely to common sense. However, trying to understand human behavior by focusing primarily on what a person was or was not thinking at the time of acting in a criminal manner is like trying to understand the nature of water by focusing only on hydrogen atoms (Fondacaro, 2000). Now, it is certainly understandable why humans have built-in tendencies to make their best guesses about human behavior on this basis, particularly when they are not aided by modern advances in behavioral science theory and research. However, the past hundred years or so of systematic research on the causes and consequences of human behavior suggests the following:

1. A major share of human behavior is not guided by rational deliberation but is influenced and largely controlled by internal and external processes that are outside the person's conscious awareness (Bargh & Chartrand, 1999; Gazzaniga, 2011; Libet, 1985).
2. The sense of conscious will is largely illusory, and cognitions (*mens rea*) do not directly cause human behavior but have an offline, indirect influence on later behavior (Baumeister, Masicampo, & Vohs, 2011; Wegner & Wheatley, 1999).
3. The focus needs to go beyond cognition and factor the very substantial influence of emotion and affective arousal on behavior into appraisals of culpability (Greene & Haidt, 2002).
4. People have a built-in tendency to anthropomorphize and to read intentionality into behavior even when it is not present, which

presents problems when culpability turns so heavily on judgments of intentionality (Heider & Simmel, 1944).

5. When judging the behavior of other people, individuals in American culture tend to overemphasize the influence of intrapersonal factors and dispositions, and tend to underestimate the influence of situational influences (Nisbett & Ross, 1980).
6. Although the focus is on the visual characteristics and physical make-up of the individual being judged, each person is embedded in multiple, interrelated micro-social contexts that are mostly invisible to the eye but have ongoing influences on the person's behavior (Brickman, Folger, Goode, & Schul, 1981; Bronfenbrenner, 1979; Moos, 1973).
7. Even accepting the legal and folk psychology presumption of a direct causal link between cognitions and behavior, there are situational factors and personal characteristics of defendants, including race, that can contribute to biases in judgments of mens rea and culpability, which are often implicit and outside the conscious awareness of decision makers (Kang et al., 2012).
8. Recent advances in the behavioral and neurosciences are making it increasingly possible to comprehend and in many cases visualize previously invisible social and biological factors that dynamically influence human behavior at both conscious and unconscious levels (Freeman, 2011; Greene & Cohen, 2004).
9. The increasing ability to quantify the combined influences of biological, psychological, and social factors on human behavior is gradually leaving less room for the kind of legal fictional homunculus necessary to justify retributive punishment on moral grounds (Fondacaro, 2014).
10. This broad telescopic lens now available to help understand the interrelated biopsychosocial influences on behavior also is beginning to provide new insights about levers at each of these levels that can and have been used to help change human behavior (i.e., prevent crime, reduce recidivism, and promote socially responsible behavior) (Fondacaro, 2014).

Although it is understandable that common sense and practical experience throughout evolutionary heritage would lead individuals to focus on intentionality and conscious thought processes as the basis for making judgments about culpability, new tools are now available to help make these judgments more accurately, if we choose to use them. As discussed

above, psychological research has demonstrated that individuals have strong retributive urges based largely on judgments about what another person was thinking at the time of engaging in criminal conduct. Many people are especially motivated to punish with negative and psychologically painful sanctions behaviors that they judge to be motivated or accompanied by conscious intention. However, the mere documentation of these sentiments and preferences does not mean that the criminal law has to continue to fly blind to factors other than cognition or *mens rea* in making judgments about legal responsibility and punishment. Moreover, determinations of criminal responsibility and guilt, whether grounded in principles of free will or determinism, need not be inextricably linked to the infliction of human suffering as a consequence.

Given the vast body of empirical studies regarding psychological issues in the legal context, it is somewhat surprising that few studies have looked at the issue of the reliability and validity of judgments of *mens rea*. Perhaps this reflects the fact that folk psychology presumptions are so deeply imbedded in the collective psyche and legal culture that it seems self-evident that jurors and judges make reliable and valid judgments of *mens rea* and mental state. Or, perhaps, the thought that such judgments are fraught with inevitable and systematic bias is so threatening to the sense of justice, even among dogged empirical scientists, that this question is both consciously and unconsciously avoided. This is perhaps an empirical question worthy of investigation in its own right. Maybe the lack of research on this topic reflects a pragmatic if not cynical calculation that addressing these fundamental issues at the core of the legal system is mere tilting at windmills. Perhaps there is even a touch of fear-based concern that such research would be perceived as subversive to traditional Western liberal values of autonomous individualism. In any case, one need not answer these questions in order to examine the few studies that have been conducted and to critically analyze their findings.

Some of the earliest research in this area was conducted by Severance and Loftus (1982), who demonstrated that jurors have difficulty making reliable distinctions between the distinct *mens rea* categories. A team of researchers with interests in the relationship between neuroscience and the law has revisited this area of research more recently. In their initial study using the Model Penal Code (MPC) *mens rea* categories, Shen, Hoffman, Jones, Greene, & Maoris (2011) found that subjects had particular difficulty distinguishing between mental states of knowing and reckless, but concluded

that overall, their findings supported the accuracy or validity of the behavioral assumptions underlying the MPC classification scheme. However, close examination of their findings suggested that when juror ratings are compared to the gold standard ratings of criminal law experts, the reliabilities of their judgments of negligent, reckless, and knowing mental states were at or below 50 percent, and the reliability for judging a mental state as purposeful was only 78 percent. Particularly alarming was the fact that 47 percent of the scenarios where the “correct” mental state was reckless, were judged to have been either knowing or purposeful, which would have resulted in over-punishment from a retributive standpoint. Such an error rate would be particularly alarming in cases involving homicide, where reckless homicide is considered manslaughter but either knowing or purposeful homicide is considered murder; under such conditions, a defendant wrongly convicted of murder could potentially be subjected to a death sentence rather than the five or so years he would have served for manslaughter.

Furthermore, although the authors refer to the “accuracy” of the judgments made, the findings did not directly bear on issues of accuracy or validity of the assessment. Under the MPC, purposeful, reckless, and knowing mental states are subjective standards. That is, they are presumed to reflect the actual state of subjective, conscious awareness of the defendant. Demonstrating a 78 percent agreement between jurors and legal experts in their judgments of a purposeful mental state does not answer the question of whether anyone was accurate or correct in judging the subjective experience of the defendant—a critical requirement of the MPC and a foundational assumption of moral judgment and retributive models of punishment. Reliability or consistency in judgment sets an upper limit on validity or accuracy of judgment, but they are not the same. In a later related study, the investigators were able to improve jurors’ ability to make distinctions between the MPC categories of reckless and knowing, but this greater consistency in judging mental state was not translated into commensurate changes or distinctions in punishments provided to defendants (Ginther et al., 2014). As discussed below, even small amounts of systematic bias and error can be magnified when they are brought up to scale and factored into the various decision-making points throughout the criminal justice process.

Research in cognitive psychology by Kahneman (2011) and others suggests that human decision making is influenced by two distinct processes:

fast, implicit, unconscious processes guided by simplifying heuristics, and slower, more rational and deliberative processes that are in part reflected in conscious awareness. Theoretical and empirical research in this area has led to studies of implicit biases and their influence on decision-making outcomes in a variety of real-world contexts. Unfortunately, there is a paucity of work focusing on the impact of implicit biases on judgments of mens rea and criminal responsibility. The work that is available does suggest that implicit racial biases may be relevant to decision making at various stages of the criminal justice system, including judgments of culpability. For example, research suggests that an ambiguous act of aggression (e.g., a shove) is likely to be rated as more violent when performed by a Black person and more likely to be attributed to that person's dispositional characteristics than when a White person does the shoving (Duncan, 1976). Other research suggests that we see Black boys as more adult and less innocent than White boys of the same age (Goff, Jackson, Di Leone, Culotta, & DiTomasso, 2014). Combined, these studies begin to suggest that, at least in some circumstances, jurors may be more likely to implicitly assume a greater degree of culpability when judging a Black defendant than a White defendant. Along these lines, Levinson (2007) found that implicit racial biases impact the type of facts that mock jurors are likely to remember, with more information regarding aggression being remembered about Black defendants, and more information relevant to mitigating factors being remembered about Whites. In a recent study, Levinson, Cai, and Young (2010) found that their jury-eligible study participants implicitly associated photos of Black men with guilty verdicts, calling into question the presumption of innocence. Finally, in a field study of unconscious racial bias among trial judges, the investigators found that trial judges held implicit racial biases that were at least as pronounced as those held by the general public (Rachlinski, Johnson, Wistrich, & Guthrie, 2009). Although this series of studies suggests that implicit biases may influence judgments of mens rea in criminal cases, few investigators have made this distinct connection. More focused and systematic research on this issue is clearly warranted.

Overall, evidence is beginning to accumulate to suggest that racial biases, often implicit, can affect decision making across various levels of the criminal justice system, ranging from interactions with the police (American Civil Liberties Union, 2013; Banks, 2001; Brown & Langan, 2001; Correll et al., 2007; Eberhardt, Goff, Purdie, & Davies, 2004), to decisions about bail and plea bargaining (Kutateladze & Andiloro, 2014;

Schlesinger, 2005), to eyewitness identifications. For example, recent evidence suggests that eyewitnesses to crimes that are stereotypically associated with Blacks, such as violence or mugging, may be more prone to mistaken identification when the suspect has more stereotypically Afrocentric features (e.g., darker skin, thicker lips, and/or wider nose) (Osborne & Davies, 2013). Moreover, once at trial, jurors may become more confident in their guilty verdicts as the defendant's skin darkens. For example, Levinson and Young (2010) found that when mock jurors were presented with photo evidence depicting a suspect whose skin tone was either dark or light, those presented with the photo of the dark-skinned suspect considered it to be stronger evidence of guilt and were more likely to render a guilty verdict. Other evidence suggests that Blacks with more Afrocentric features receive more severe sentences and that those who kill Whites are more likely to be sentenced to death if they are perceived to have more stereotypical Afrocentric features, even after controlling for a wide range of factors (Blair, Judd, & Chapleau, 2004; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006). Finally, at the level of public policy, Hetey and Eberhardt (2014) found in an experimental study that Americans support more punitive criminal justice policies (i.e., longer prison sentences for minor offenses) when a greater proportion of those affected are presented as "more Black" than when they are presented as "more White."

Presently, the extent to which racial disparities in rates of incarceration are due to implicit biases is unknown. However, even small biases can have a large effect at the societal level when aggregated across the various stages of the criminal justice process ranging from arrest to sentencing (Kang et al., 2012). For example, as noted in an article by Kang and colleagues (2012), co-author Anthony Greenwald has done some calculations based on simulation data using a very conservative effect size of $r = .1$ across each of the decision points, and found that a Black defendant would serve a sentence that was almost two times longer than the sentence served by a White for a crime with an average sentence of five years. When you raise the effect size to $r = .2$, which is still lower than the average effect size for racial discrimination effects, the Black defendant would serve a sentence that was over three times longer. The cumulative magnitude of the social impact of such decision-making biases becomes quite apparent when aggregated across all defendants subjected to implicit biases and across the nearly 21 million state and federal criminal cases (Kang et al., 2012).

IV. THE PSYCHOSOCIAL CONTEXT AND CONSEQUENCES OF PUNISHMENT

As noted above, punishment is a multidimensional construct that can have both retributive and consequentialist dimensions. This section will focus on the psychosocial context and the consequences of punishment for the punisher and the punished at the individual, community, and societal levels.

A. The Punisher

In addressing the psychosocial consequences for individuals who punish others, the analysis begins by examining the contexts in which punishment is delivered. The effects of punishment on the individual punisher is a topic about which there is more commonsense speculation than empirical evidence or research. However, research suggests that people typically respond to direct provocation with a strong emotional reaction associated with moral outrage and the triggering of a fight-or-flight response (Darley, 2009). In the context of second-party punishment, whether the person immediately retaliates with an aggressive response may depend on a wide variety of factors including the magnitude of the threat, the person's ability to deal with it head on, his or her level of impulse control and behavioral restraint, the range of alternative responses he or she has available to meet the challenge, and his or her ability to execute the response chosen (Crick & Dodge, 1994; McFall, 1982). The immediate goal often involves a combination of diminishing or eliminating the immediate threat and deriving a sense that justice has been served (Darley, 2009). There may also be some secondary expectation that future threats will be deterred (Hoffman, 2014).

In addition to the immediate emotional reaction of the individual punisher, including excitation of some of the reward centers of the brain, people may act aggressively in response to a transgression in an attempt to improve their negative mood (Bushman, Baumeister, & Phillips, 2001; de Quervain et al., 2004; Knutson, 2004). However, over time, those who release anger toward an offending party often ruminate and become angrier as a result. In effect, avoiding the impulse to punish is generally a more effective means of obtaining emotional equilibrium than venting anger—a finding in direct contradiction to theoretical and popular notions of the beneficial cathartic effects of administering punishment (Carlsmith, Wilson, & Gilbert, 2008). Moreover, anger arousal itself can increase the

blame attributed to others in interpersonal conflict situations and enhance judgments of criminal intent in situations involving ambiguous criminal conduct (Ask & Pina, 2011).

In a related vein, strong beliefs in autonomous individualism and free will are associated with more harsh judgments of immoral behavior. In turn, exposure to immoral behavior can itself strengthen belief in free will as a means to justify blaming someone for a moral transgression (Clark et al., 2014). Other individual and demographic characteristics that generally stand out as common in those who endorse particularly punitive forms of punishment include being male, conservative, authoritarian, and older (Applegate, Cullen, & Fisher, 2002; Bowers, 1998; Bray & Noble, 1978; Rossi & Berk, 1997). Overall, one common denominator in most individuals' desire to punish is the sense of a need for retaliation in response to anger, blame, or harm (Bushman et al., 2001). However, notwithstanding the general public endorsement of retributive principles and the assumptions about human behavior that guide them, recent research suggests that the American public may be willing to accept sanctions based on principles of restorative justice, at least when some element of retributive punishment is retained (Gromet & Darley, 2006).

Regarding third-party punishment, some of the same emotional reactions are present that that are evoked by second-party punishment, although in muted form. Here too, there is often some expectation that retributive punishment will not only provide a sense of closure for victims and their families but that it will also contribute to a public sense that just desert has been served. More research on this topic is clearly needed to identify the extent to which public perceptions based on common sense about closure is supported by empirical evidence. However, just as the criminal justice system is a narrow tool for preventing crime, it is likely to be a narrow tool for bringing about the complex cognitive and emotional integration involved in obtaining closure after a threatening and traumatic event (Kanwar, 2002).

B. The Punished

The consequences of punishment for the punished turn in part on the nature and duration of punishment imposed. The punishment of choice in the American legal system is incarceration. The potential consequences of incarceration for the individual include exposure to institutional violence,

separation from family and community, removal from the workforce and related difficulties reentering it, association with criminal peers, increased risk of developing psychological disorders, loss of a stake in the future, and resentment (Haney, 2006; Travis & Waul, 2003; Western, 2002). There are also some presumed benefits and anecdotal examples about the positive impact of prison for some individuals. Some of the social costs of incarceration include the emotional and financial impact on the children and families of the prisoner, the tax loss and burden of incarceration, and the additional victims of crime created by the recidivism of prisoners who have been punished but not treated effectively or properly reintegrated back into society (Durose, Cooper, & Snyder, 2014; Travis, McBride, & Solomon, 2005).

If one of the goals of incarceration is to prevent future crime and violence, it is a dismal failure, as reflected in the high rates of violence within prisons and the high recidivism rates upon release (Chen & Shapiro, 2007). The threat of harsh punishment is not effectively preventing crime because individuals do not often contemplate the probability of being caught and punished when acting on an impulse to commit a crime (MacCoun, 1993). Drawing on the conceptualization of Keller and colleagues (2010) suggesting that general deterrence can have both negative and positive features, Tyler's (1990; 2001; 2005) research on procedural justice seems particularly relevant here. His work suggests that compliance with the law is enhanced when citizens, including minorities, have a voice in the process, are treated fairly by legal authorities, and view legal institutions as legitimate sources of authority that reflect their interests. On the one hand, negative interactions with police and increasing rates of incarceration within minority communities may undermine the vital sense of trust and legitimacy of legal institutions, and contribute to resentment and legal non-compliance. Alternatively, better relations between minority communities and all arms of the criminal justice system may be important ingredients in the kind of positive, forward-looking general deterrence inspired by Keller and colleagues' (2010) comprehensive framework (Eitle, Stolzenberg, & D'Alessio, 2005; Tyler, 2001; 2005). Even the most ardent retributivists need to consider the range of personal and social costs of incarceration noted above in their attempts to calibrate truly just and proportionate punishments.

One of the primary reasons for the advocacy of long prison sentences as the dominant response to criminal behavior is the belief that "nothing works" in rehabilitating offenders (Martinson, 1974). In fact, this view was

in part responsible for the shift in juvenile justice policy from rehabilitation to punishment. However, based on rapid advances over the past few decades, there is considerable evidence that rehabilitation does work, with recidivism rates being brought down from a baseline of about 70 percent to just over 20 percent for juvenile offenders, even very serious ones, when they are provided with community-based, multisystemic interventions that target the youth's social skills and family, peer, school, and community environments (Borduin et al., 1995). This progress reflects a shift in frame of reference from a primary focus on intrapsychic factors and character traits of the delinquent youth to a more social ecological perspective.

Similar community-based risk management interventions also show promise with adult offenders, although more systematic research in this area is needed (Gendreau, Cullen, & Bonta, 1994). Back in 1998, the National Institute of Justice initiated partnerships with several state governments in an effort to systematically review evidence on what works in reducing recidivism (Sherman et al., 1998). Among their most prominent findings, Sherman and colleagues found that vocational training, risk-focused treatment programs, and community corrections can all be effective in reducing recidivism for various crime-involved groups.

More recently, in a meta-analysis of treatment studies of the effectiveness of cognitive-behavioral interventions with adult and juvenile offenders in both correctional and community settings, Landenberger and Lipsey (2005) found that cognitive-behavioral therapy (CBT) was indeed effective at reducing recidivism. They noted that the effective ingredients of the intervention were treatment of higher-risk offenders, effective implementation, inclusion of an anger management component, and training in interpersonal problem solving. Overall, CBT produced a 25 percent reduction in recidivism, and programs that included the best combination of the identified effective CBT ingredients led to a 50 percent reduction in recidivism in comparison to the average control group. Landenberger and Lipsey (2005) concluded, "The central issue for research on CBT with offender populations at this juncture is not to determine if it has positive effects, but to determine when and why it has the most positive effects" (p. 472). Other research has demonstrated that restorative justice programs may have some positive benefits and help to reduce recidivism, at least in first-time offenders involved in less serious crimes (Rodriguez, 2007). More broadly, the value of restorative justice may go beyond any impact on recidivism if it is incorporated as a component into other evidence-based

interventions and shown to satisfy at least in part the public's demand for some restitution and "payback" by the offender.

In recent years, as society has slowly begun to recognize the harmful impacts of overreliance on incarceration, state governments have begun to endorse (and in some cases, even mandate) more of a push toward evidence-based community corrections (Burrell & Rhine, 2013). Although juveniles are currently more likely to benefit from these policy shifts than adults, this move is slowly beginning to impact adults as well. Overall, there is a need for rigorous analysis of the effective ingredients of a wide range of interventions with adult offenders. This effort should be directed at identifying which types of interventions work best for which type of offenders and under what circumstances within both institutional and community-based settings. Moving forward, the old mantra of "nothing works" should be replaced by a new one proclaiming that "evidence-based interventions do work with both juvenile and adult offenders."

V. DIRECTIONS FOR FUTURE RESEARCH AND POLICY REFORM

Based on the literature reviewed above, the following list of suggestions for future research and policy reform is proposed:

1. Examine the full scope of the public conception of retribution and just desert. To what extent do facets of this construct go beyond proportionate punishment?
2. Broaden the focus of basic research on consequentialist responses to crime, and identify the circumstances under which the public may be willing to embrace these broader approaches.
3. Conduct systematic research contrasting folk psychology notions of *mens rea* and criminal responsibility with evidence from empirical research.
4. The very resistance to studying factors that challenge folk psychology notions of autonomous individualism, in and of itself, should become the focus of empirical inquiry.
5. Study discrepancies between public perceptions of the consequences of incarceration and direct empirical evidence along such dimensions as the level of psychological distress and abuse prisoners do and should feel while incarcerated.

6. Identify what factors, both positive and negative, influence the connection between juvenile delinquency and later adult incarceration.
7. Investigate implicit biases at all stages of the criminal justice system, including the guilt and sentencing phases of the trial. Do the calculations for different types of crimes and the extent to which any implicit biases when brought up to scale contribute to disproportionate minority confinement.
8. Empirically assess and facilitate the implementation, evaluation, and refinement of alternative responses to crime such as community-based corrections and probation.
9. Identify the parameters of moral panic, and adopt policies that will stem strong public over-reaction to real or exaggerated crises.
10. Investigate the role of media in influencing public preferences for alternative sanctions for criminal behavior.
11. Examine the implications of analogizing from the social-cognitive problem-solving models that psychologists have successfully developed to address retaliatory and instrumental aggression at the individual level (Crick & Dodge, 1994), and bringing them up to scale at the policy level to help diminish the unreflective demand for unduly harsh punishment in the American criminal justice system.
12. Encourage researchers from a wide range of neighboring disciplines to help contextualize human behavior for purposes of better informing both moral intuitions and criminal justice policies.

CONCLUSION

This article was written to provide some insight about the causes and consequences of punishment from a psychological perspective, and more importantly, to answer the call of the National Academy of Sciences report to draw on interdisciplinary research to help guide legal and policy reform aimed at reducing mass incarceration. To date, the criminal law has been anchored in an intuitively appealing, but empirically discredited, folk psychology model of human behavior that narrowly focuses culpability analysis on difficult- if not impossible-to-judge past mental states—a process that fuels unreflective retributive urges that provide moral cover for unduly harsh sanctions for criminal behavior, punitive

urges that are legitimized and encouraged at individual and systemic levels. In the policy realm, well-intentioned scholars and policy makers called for the rebirth of retribution during the closing decades of the twentieth century as a limit on government overreach and in response to the empirical science of the day, suggesting that nothing could be done to rehabilitate criminal offenders.

Now, more than a decade into the twenty-first century, the field of psychology and its neighboring disciplines have something new to contribute to criminal law and criminal justice policy reform aimed at reducing mass incarceration. To reduce the number of citizens incarcerated in the American criminal justice system, it will be necessary to take full advantage of consequentialist conceptualizations of and responses to crime that make the reduction and prevention of criminal behavior and rehabilitation of the offender explicit and measurable policy goals—goals that can and should be achieved in the least restrictive setting necessary to ensure public safety. This will require a shift away from intuitive, folk psychology models of criminal responsibility that focus on divining atomistic past mental states that decontextualize human behavior and help justify unduly harsh criminal sanctions. Human behavior is much more complex, better understood, and more fairly judged when analyzed within a social ecological framework that takes into account the complex interplay of dynamic biological, psychological, and social factors. Such biopsychosocial factors not only contribute to criminal behavior but also serve as levers that can help promote law abiding behavior among citizens from diverse ethnic and racial backgrounds and affirm the human dignity of each individual in American society.

Recent psychological research has indeed helped to illuminate the strong, ingrained, and often overwhelming human impulse toward retribution and retaliation in response to wrongdoing. However, that same body of research has also helped to identify the human capacity for more advanced, higher-order rational and deliberative processes that can help reign in those urges and consider alternative responses and their consequences. One of the exemplars of this highly developed if not unique human capacity is the ability to create and use tools to achieve outcomes that would be otherwise impossible. The scientific method and systematic empirical research is one such potential tool. The question is, will the criminal law and criminal justice policies take full advantage of this tool?

REFERENCES

- American Civil Liberties Union. (2013). *The war on marijuana in black and white*. New York, NY: ACLU Foundation. Retrieved from <https://www.aclu.org/files/assets/III4413-mj-report-rfs-reli.pdf>
- Applegate, B. K., Cullen, F. T., & Fisher, B. S. (2002). Public views toward crime and correctional policies: Is there a gender gap? *Journal of Criminal Justice*, 30(2), 89–100. doi:10.1016/S0047-2352(01)00127-1
- Ask, K., & Pina, A. (2011). On being angry and punitive: How anger alters perception of criminal intent. *Social Psychological and Personality Science*, 2(5), 494–499. doi:10.1177/1948550611398415
- Banks, R. R. (2001). Race-based suspect selection and colorblind equal protection doctrine and discourse. *UCLA Law Review*, 48, 1075–1124.
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, 54(7), 462–479. doi:10.1037/0003-066X.54.7.462
- Baumeister, R. F., Masicampo, E. J., & Vohs, K. D. (2011). Do conscious thoughts cause behavior? *Annual Review of Psychology*, 62(1), 331–361. doi:10.1146/annurev.psych.093008.131126
- Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science*, 15(10), 674–679. doi:10.1111/j.0956-7976.2004.00739.x
- Borduin, C. M., Mann, B. J., Cone, L. T., Henggeler, S. W., Fucci, B. R., Blaske, D. M., & Williams, R. A. (1995). Multisystemic treatment of serious juvenile offenders: Long-term prevention of criminality and violence. *Journal of Consulting and Clinical Psychology*, 63(4), 569–578.
- Bowers, D. A. (1998). Giving people what they want: An exploratory analysis of felony sentencing in 49 states. *International Journal of Comparative and Applied Criminal Justice*, 22(1), 119–130. doi:10.1080/01924036.1998.9678612
- Bray, R. M., & Noble, A. M. (1978). Authoritarianism and decisions of mock juries: Evidence of jury bias and group polarization. *Journal of Personality and Social Psychology*, 36(12), 1424–1430. doi:10.1037/0022-3514.36.12.1424
- Brickman, P., Folger, R., Goode, E., & Schul, Y. (1981). Microjustice and macrojustice. In M. J. Lerner & S. C. Lerner (Eds.), *The Justice Motive in Social Behavior* (pp. 173–202). New York, NY: Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4899-0429-4_9
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Brown, J. M., & Langan, P. A. (2001). *Policing and homicide, 1976–98: Justifiable homicide of felons by police and murder of police by felons* (No. NCJ 180987). Washington, DC: Bureau of Justice Statistics. Retrieved from <http://www.bjs.gov/content/pub/pdf/ph98.pdf>
- Buckholtz, J. W., Asplund, C. L., Dux, P. E., Zald, D. H., Gore, J. C., Jones, O. D., & Marois, R. (2008). The neural correlates of third-party punishment. *Neuron*, 60(5), 930–940. doi:10.1016/j.neuron.2008.10.016
- Burrell, W. D., & Rhine, E. E. (2013). Implementing evidence-based practices in community corrections. *Justice Research and Policy*, 15(1), 143–158. doi:10.3818/JRP.15.1.2013.143

- Bushman, B. J., Baumeister, R. F., & Phillips, C. M. (2001). Do people aggress to improve their mood?: Catharsis beliefs, affect regulation opportunity, and aggressive responding. *Journal of Personality and Social Psychology*, *81*(1), 17–32. doi:10.1037/0022-3514.81.1.17
- Carlsmith, K. M. (2006). The roles of retribution and utility in determining punishment. *Journal of Experimental Social Psychology*, *42*(4), 437–451. doi:10.1016/j.jesp.2005.06.007
- Carlsmith, K. M., & Darley, J. M. (2008). Psychological aspects of retributive justice. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 40, pp. 193–236). Waltham, MA: Academic Press. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0065260107000044>
- 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. doi:10.1037/0022-3514.83.2.284
- Carlsmith, K. M., Wilson, T. D., & Gilbert, D. T. (2008). The paradoxical consequences of revenge. *Journal of Personality and Social Psychology*, *95*(6), 1316–1324. doi:10.1037/a0012165
- Chen, M. K., & Shapiro, J. M. (2007). Do harsher prison conditions reduce recidivism?: A discontinuity-based approach. *American Law and Economics Review*, *9*(1), 1–29. doi:10.1093/aler/ahm006
- Clark, C. J., Luguri, J. B., Ditto, P. H., Knobe, J., Shariff, A. F., & Baumeister, R. F. (2014). Free to punish: A motivated account of free will belief. *Journal of Personality and Social Psychology*, *106*(4), 501–513. doi:10.1037/a0035880
- Correll, J., Park, B., Judd, C. M., Wittenbrink, B., Sadler, M. S., & Keesee, T. (2007). Across the thin blue line: Police officers and racial bias in the decision to shoot. *Journal of Personality and Social Psychology*, *92*(6), 1006–1023. doi:10.1037/0022-3514.92.6.1006
- Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin*, *115*(1), 74–101. doi:10.1037/0033-2909.115.1.74
- Darley, J., Fulero, S., Haney, C., & Tyler, T. (2002). Psychological jurisprudence. In *Taking Psychology and Law into the Twenty-First Century* (pp. 35–59). New York, NY: Springer US. Retrieved from http://link.springer.com/chapter/10.1007/0-306-47944-3_2
- Darley, J. M. (2009). Morality in the law: The psychological foundations of citizens' desires to punish transgressions. *Annual Review of Law and Social Science*, *5*(1), 1–23. doi:10.1146/annurev.lawsocsci.4.110707.172335
- Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2000). Incapacitation and just deserts as motives for punishment. *Law and Human Behavior*, *24*(6), 659–683. doi:10.1023/A:100552203727
- De Quervain, D. J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, *305*(5688), 1254–1258. doi: 10.1126/science.1100735
- Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology*, *34*(4), 590–598. doi:10.1037/0022-3514.34.4.590
- Durose, M. R., Cooper, A. D., & Snyder, H. N. (2014). *Recidivism of prisoners released in 30 states in 2005: Patterns from 2005 to 2010* (No. ncj 244205). Washington, DC: Bureau of Justice Statistics.

- Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science*, 17(5), 383–386.
- Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing black: Race, crime, and visual processing. *Journal of Personality and Social Psychology*, 87(6), 876–893. doi:10.1037/0022-3514.87.6.876
- Eitle, D., Stolzenberg, L., & D'Alessio, S. J. (2005). Police organizational factors, the racial composition of the police, and the probability of arrest. *Justice Quarterly*, 22(1), 30–57. doi:10.1080/0741882042000333636
- Fondacaro, M. R. (2000). Toward an ecological jurisprudence rooted in concepts of justice and empirical research. *University of Missouri-Kansas City School of Law*, 69, 179–196.
- Fondacaro, M. R. (2011). The injustice of retribution: Toward a multi-systemic risk management model of juvenile justice. *Journal of Law and Policy*, 20, 145–165.
- Fondacaro, M. R. (2014). Rethinking the scientific and legal implications of developmental differences research in juvenile justice. *New Criminal Law Review: An International and Interdisciplinary Journal*, 17(3), 407–441. doi:10.1525/nclr.2014.17.3.407
- Fondacaro, M. R., & Weinberg, D. (2002). Concepts of social justice in community psychology: Toward a social ecological epistemology. *American Journal of Community Psychology*, 30(4), 473–492. doi:10.1023/A:1015803817117
- Freeman, M. (Ed.). (2011). *Law and neuroscience: Current legal issues* (Vol 13). New York, NY: Oxford University Press.
- Gazzaniga, M. S. (2011). *Who's in charge?: Free will and the science of the brain*. New York, NY: HarperCollins Publishers.
- Gendreau, P., Cullen, F. T., & Bonta, J. (1994). Intensive rehabilitation supervision: The next generation in community corrections? *Federal Probation*, 58(1), 72–78.
- Ginther, M. R., Shen, F. X., Bonnie, R. J., Hoffman, M. B., Jones, O. D., Marois, R., & Simons, K. (2014). The language of mens rea. *Vanderbilt Law Review*, 67, 1327–1372.
- Goff, P. A., Jackson, M. C., Di Leone, B. A. L., Culotta, C. M., & DiTomasso, N. A. (2014). The essence of innocence: Consequences of dehumanizing Black children. *Journal of Personality and Social Psychology*, 106(4), 526–545. doi:10.1037/a0035663
- Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 359(1451), 1775–1785. doi:10.1098/rstb.2004.1546
- Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, 6(12), 517–523. doi:10.1016/S1364-6613(02)02011-9
- Gromet, D. M., & Darley, J. M. (2006). Restoration and retribution: How including retributive components affects the acceptability of restorative justice procedures. *Social Justice Research*, 19(4), 395–432. doi:10.1007/s11211-006-0023-7
- Haney, C. (2006). *Reforming punishment: Psychological limits to the pains of imprisonment* (1st ed.). Washington, DC: American Psychological Association. Retrieved from <http://www.apa.org/pubs/books/4316061.aspx>
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57(2), 243–259. doi:10.2307/1416950
- Hetey, R. C., & Eberhardt, J. L. (2014). Racial disparities in incarceration increase acceptance of punitive policies. *Psychological Science*, 25, 1949–1954.

- Hoffman, M. B. (2014). *The punisher's brain: The evolution of judge and jury*. Cambridge, England: Cambridge University Press. Retrieved from <http://www.cambridge.org/us/academic/subjects/law/socio-legal-studies/punishers-brain-evolution-judge-and-jury>
- Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.
- Kang, J., Bennett, M. W., Carbado, D. W., Casey, P., Dasgupta, N., Faigman, D. L., . . . Mnookin, J. (2012). Implicit bias in the courtroom. *UCLA Law Review*, 59, 1122–1186.
- Kanwar, V. (Vik). (2002). Capital punishment as closure: The limits of a victim-centered jurisprudence. *NYU Review of Law and Social Change*, 27(2), 215–255.
- Kaplan, J., Weisberg, R., & Binder, G. (2000). *Criminal law: Cases and materials*. New York, NY: Aspen Publishers.
- 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(2–3), 99–116. doi:10.1007/s11211-010-0113-4
- Knutson, B. (2004). Sweet revenge? *Science*, 305(5688), 1246–1247. doi:10.1126/science.1102822
- Kolber, A. J. (2012). Unintentional punishment. *Legal Theory*, 18(01), 1–29. doi:10.1017/S1352325211000218
- Kutateladze, B. L., & Andiloro, N. R. (2014). *Prosecution and racial justice in New York county* (National Institute of Justice No. 2011-DJ-BX-0038). New York, NY: Vera Institute of Justice. Retrieved from <http://www.vera.org/sites/default/files/resources/downloads/race-and-prosecution-manhattan-technical.pdf>
- Landenberger, N. A., & Lipsey, M. W. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, 1(4), 451–476. doi:10.1007/s11292-005-3541-7
- Levinson, J. D. (2007). Forgotten racial equality: Implicit bias, decisionmaking, and misremembering. *Duke Law Journal*, 57(345), 345–392.
- Levinson, J. D., Cai, H., & Young, D. (2010). Guilty by implicit racial bias: The guilty/not guilty implicit association test. *Ohio State Journal of Criminal Law*, 8(187), 187–208.
- Levinson, J. D., & Young, D. (2010). Different shades of bias: Skin tone, implicit racial bias, and judgments of ambiguous evidence. *West Virginia Law Review*, 112, 307–350.
- Libet, B. W. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8(4), 529–66.
- MacCoun, R. J. (1993). Drugs and the law: A psychological analysis of drug prohibition. *Psychological Bulletin*, 113(3), 497–512.
- Martinson, R. (1974). What works? Questions and answers about prison reform. *The Public Interest*, 35(3), 22–54.
- McFall, R. M. (1982). A review and reformulation of the concept of social skills. *Behavioral Assessment*, 4(1), 1–33. doi:10.1007/BF01321377
- Miller, D. T., & Vidmar, N. (1981). The social psychology of punishment reactions. In M. J. Lerner & S. C. Lerner (Eds.), *The Justice Motive in Social Behavior* (pp. 145–172). New York, NY, US: Academic Press.
- Moos, R. H. (1973). Conceptualizations of human environments. *American Psychologist*, 28(8), 652–665. doi:10.1037/h0035722
- Morse, S. J. (2000). Rationality and responsibility. *Southern California Law Review*, 74, 251–268.

- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings in social judgement*. Englewood Cliffs, NJ: Prentice Hall.
- Osborne, D., & Davies, P. G. (2013). Eyewitness identifications are affected by stereotypes about a suspect's level of perceived stereotypicality. *Group Processes & Intergroup Relations*, 16(4), 488–504.
- Rachlinski, J. J., Johnson, S. L., Wistrich, A. J., & Guthrie, C. (2009). Does unconscious racial bias affect trial judges? *Norte Dame Law Review*, 84(3), 1196–1246.
- Robinson, P. H., & Darley, J. M. (2007). Intuitions of justice: Implications for criminal law and justice policy. *Southern California Law Review*, 81(1), 1–68.
- Rodriguez, N. (2007). Restorative justice at work: Examining the impact of restorative justice resolutions on juvenile recidivism. *Crime & Delinquency*, 53(3), 355–379. doi:10.1177/001128705285983
- Rossi, P. H., & Berk, R. A. (1997). *Just punishments: Federal guidelines and public views compared*. Hawthorne, NY: Walter De Gruyter, Inc.
- Schlesinger, T. (2005). Racial and ethnic disparity in pretrial criminal processing. *Justice Quarterly*, 22(2), 170–192. doi:10.1080/07418820500088929
- Severance, L. J., & Loftus, E. F. (1982). Improving the ability of jurors to comprehend and apply criminal jury instructions. *Law & Society Review*, 17(1), 153–197. doi:10.2307/3053535
- Shen, F. X., Hoffman, M. B., Jones, O. D., Greene, J. D., & Marois, R. (2011). Sorting guilty minds. *NYU Law Review*, 86(5), 1306–1360.
- Sherman, L. W., Gottfredson, D. C., MacKenzie, D. L., Eck, J., Reuter, P., & Bushway, S. D. (1998). *Preventing crime: What works, what doesn't, what's promising* (No. NCJ 171676). New York, NY: National Institute of Justice.
- Skinner, B. F. (1965). *Science and human behavior*. New York, NY: Free Press.
- Slobogin, C., & Fondacaro, M. R. (2011). *Juveniles at risk: A plea for preventive justice*. New York, NY: Oxford University Press.
- Slobogin, S., & Brinkley-Rubinstein, L. (2013). Putting desert in its place. *Stanford Law Review*, 65(1), 77–136.
- Strobel, A., Zimmermann, J., Schmitz, A., Reuter, M., Lis, S., Windmann, S., & Kirsch, P. (2011). Beyond revenge: Neural and genetic bases of altruistic punishment. *NeuroImage*, 54(1), 671–680. doi:10.1016/j.neuroimage.2010.07.051
- Travis, J., McBride, E. C., & Solomon, A. L. (2005). *Families left behind: The hidden costs of incarceration and reentry* (No. CPR03 0105). New York, NY: Urban Institute: Justice Policy Center.
- Travis, J., & Waul, M. (2003). *Prisoners once removed: The impact of incarceration and reentry on children, families, and communities*. Washington, DC: The Urban Institute.
- Travis, J., Western, B., & Redburn, S. (2014). *The growth of incarceration in the United States: Exploring causes and consequences*. National Academy of Sciences. Washington, D.C.: National Academies Press.
- Tyler, T. R. (1990). *Why people obey the law*. New Haven, CT: Yale University Press.
- Tyler, T. R. (2001). Public trust and confidence in legal authorities: What do majority and minority members want from the law and legal institutions? *Behavioral Sciences & the Law*, 19, 215–235. doi: 10.1002/bsl.438
- Tyler, T. R. (2005). Policing in black and white: Ethnic group differences in trust and confidence in the police. *Police Quarterly*, 8(3), 322–342. doi:10.1177/109861104271105

- Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *The American Psychologist*, *54*(7), 480–492.
- Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. New York, NY: The Guilford Press.
- Weiner, B., Graham, S., & Reyna, C. (1997). An attributional examination of retributive versus utilitarian philosophies of punishment. *Social Justice Research*, *10*(4), 431–452. doi:10.1007/BF02683293
- Western, B. (2002). The impact of incarceration on wage mobility and inequality. *American Sociological Review*, *67*(4), 526–546. doi:10.2307/3088944