

Spring 2009

Deterrence in a Sea of Just Deserts: Are Utilitarian Goals Achievable in a World of Limiting Retributivism

Matthew Haist

Follow this and additional works at: <https://scholarlycommons.law.northwestern.edu/jclc>

 Part of the [Criminal Law Commons](#), [Criminology Commons](#), and the [Criminology and Criminal Justice Commons](#)

Recommended Citation

Matthew Haist, Deterrence in a Sea of Just Deserts: Are Utilitarian Goals Achievable in a World of Limiting Retributivism, 99 J. Crim. L. & Criminology 789 (2008-2009)

This Comment is brought to you for free and open access by Northwestern University School of Law Scholarly Commons. It has been accepted for inclusion in Journal of Criminal Law and Criminology by an authorized editor of Northwestern University School of Law Scholarly Commons.

DETERRENCE IN A SEA OF “JUST DESERTS”: ARE UTILITARIAN GOALS ACHIEVABLE IN A WORLD OF “LIMITING RETRIBUTIVISM”?

MATTHEW HAIST*

Recent scholarship on theories of criminal punishment has increasingly focused on retributivist justifications for punishment. While within this retributivist camp opinions differ as to the particulars of such policies, there is general agreement that criminals getting what they deserve, that is, their “just deserts,” should be the underlying goal and rationale of the criminal justice system. From this point, these scholars argue that a criminal should receive punishment according to what the criminal deserves. Some forms of retributivism, however, have attempted to draw support from other theories of criminal punishment. By borrowing elements of other theories, specifically utilitarian theories, scholars have attempted to bolster support for retributivist policies. A particularly well received form of retributivism and the focus of this Comment, “limiting retributivism,” argues that a range of punishments will fall within the criminal’s just deserts, and that utilitarian concepts can alter the punishment within the aforementioned range. This Comment scrutinizes limiting retributivism’s appeal to utilitarian theories of punishment to determine if such a system of punishment can achieve many of the outcomes sought by utilitarian theories, specifically deterrence.

This Comment argues that the answer to this question, while complex, is ultimately no. I begin my inquiry by expounding on the history of retributive and utilitarian theories of punishment, and the specific concepts of limiting retributivism and what “factors” exist in determining a criminal’s just deserts. Then, I shift focus and analyze the practical effects of these factors in light of recent behavioral psychology and behavioral law and economics research on cognitive biases. Subsequently, I describe why these insights into human psychology and the effects of cognitive biases, as

* J.D., Northwestern University School of Law, 2009; A.B., Davidson College, 2003. I would like to thank the editorial staff of the *Journal of Criminal Law and Criminology* for their comments and corrections, and as always, my parents, David and Louise.

applied to the retributive factors in determining just deserts, will actually lead to under-deterrence of criminal activity. I develop this point by acknowledging that while this research also questions utilitarianism's proposal for deterrence through ex ante incentives, the under-deterrence effect of limiting retributivism is far higher than that of utilitarianism. I also highlight why such under-deterrence is fundamentally contrary to utilitarian goals of punishment. I end by arguing that, given the insights of behavioral psychology, utilitarian goals of punishment are not sufficiently accomplished under limiting retributivism. A more pure form of utilitarianism is required to achieve utilitarian goals with hard and fast criminal rules with no appeal to or use of other theories of punishment. I also propose that utilitarians, in collaboration with behavioral law and economics scholars, can further understanding of how the criminal law may incorporate behavioral psychology insights to create more effective ex ante incentives.

I. INTRODUCTION

The question of why we should punish criminals is age-old, and elicits many more answers than the lay person would imagine. Punishing what society deems wrong or unwise seems to be an integral part of any civilization.¹ Specific guidelines for punishment date back to Hammurabi's Code,² while theoretical and philosophical justifications for punishment can be seen as early as Aristotle.³ Simply put, criminal justice is, and has always been, at the center of public concern, so much so that modern societies often view and compare themselves to one another through the lens of what we now call criminal justice.⁴ Given the importance of criminal justice to obtaining a safe and prosperous society, recent developments in the field must constantly be monitored to ensure that we operate under a theory of criminal punishment that either is optimally in

¹ I have not found any evidence or instance of a society that does not have a form of punishment.

² HAMMURABI'S CODE OF LAWS (c. 1780 B.C.E.) (L.W. King trans.), available at <http://www.fordham.edu/halsall/ancient/hamcode.html> (last visited May 1, 2009).

³ ARISTOTLE, NICOMACHEAN ETHICS 200-01 (Roger Crisp trans., Cambridge Univ. Press 2000).

⁴ One only needs to look toward the Europeans' general view of the death penalty and their detestation of continued U.S. support for such punishment to see evidence of this comparison. Justice Scalia's concurring opinion in *Kansas v. Marsh* highlights this point. 126 S. Ct. 2516, 2532 (2006) (Scalia, J., concurring) ("There exists in some parts of the world sanctimonious criticism of America's death penalty.").

line with our concepts of morality or best serves to provide the needed security.⁵

Fortunately, scholars have spent much time and effort—increasingly so within the past twenty to thirty years—debating this issue.⁶ These debates affect our view of which system of criminal punishment is ideal, and demonstrate that the discussion is never-ending.⁷ Criminal justice has evolved over the millennia and will continue to do so over the next. Under Hammurabi's Code, the punishment for almost any crime was death.⁸ Under the Qur'an, crimes were given specific punishments, yet required a particular level of proof to be established to find guilt.⁹ English common law guaranteed a criminal defendant a right to trial by his or her peers.¹⁰ And our founding fathers added the right against self-incrimination and the prohibition of cruel and unusual punishment, while preserving the English common law right to trial by jury.¹¹

⁵ If one gives credence to social contract theory, security and safety is the principal reason individuals enter society in the first place. Therefore, preservation of individuals' interests, be they bodily integrity, property interests, or general liberties, should be the main priority and obligation of government. See THOMAS HOBBES, *LEVIATHAN* 121 (Richard Tuck ed., rev. student ed., Cambridge Univ. Press 1996) (1651) ("A *Common-wealth* is said to be *Instituted* . . . [so] that Man, or Assembly of men . . . live peacefully amongst themselves, and be protected against other men."); JOHN LOCKE, *TWO TREATISES ON GOVERNMENT* 269-78 (Peter Laslett ed., Cambridge Univ. Press 2002) (1698); JEAN-JACQUES ROUSSEAU, *THE SOCIAL CONTRACT AND THE FIRST AND SECOND DISCOURSES* 163-64 (Susan Dunn trans., Yale Univ. Press 2002).

⁶ See, e.g., Albert W. Alschuler, *The Changing Purposes of Criminal Punishment: A Retrospective on the Past Century and Some Thoughts About the Next*, 70 U. CHI. L. REV. 1 (2003); Russell L. Christopher, *Deterring Retributivism: The Injustice of "Just" Punishment*, 96 NW. U. L. REV. 843 (2002); Richard S. Frase, *Punishment Purposes*, 58 STAN. L. REV. 67 (2005); Norval Morris, *Desert as a Limiting Principle*, in *PRINCIPLED SENTENCING* 201 (Andrew von Hirsch & Andrew Ashworth eds., 1992); Alice Ristroph, *Desert, Democracy, and Sentencing Reform*, 96 J. CRIM. L. & CRIMINOLOGY 1293 (2006); Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453 (1997).

⁷ See *supra* note 6.

⁸ HAMMURABI'S CODE OF LAWS, *supra* note 2 (stating, for example, in Code number six that "[i]f any one steal the property of a temple or of the court, he shall be put to death, and also the one who receives the stolen thing from him shall be put to death," and in Code number fourteen that "[i]f any one steal the minor son of another, he shall be put to death").

⁹ AN INTERPRETATION OF THE QUR'AN [THE KORAN], *Al-mā'idah* 5:38, at 113 (Majid Fakhry trans., bilingual ed., N.Y. Univ. Press 2002) ("As for the thieves, whether male or female, cut off their hands in punishment for what they did . . .").

¹⁰ See 3 WILLIAM BLACKSTONE, *COMMENTARIES ON THE LAWS OF ENGLAND* 349 (Dawsons of Pall Mall 1966) (1768); see also A.E. DICK HOWARD, *MAGNA CARTA: TEXT AND COMMENTARY* § 39 (Univ. Press of Virginia 1964) (1215) ("No free man shall be taken, imprisoned, disseised, outlawed, banished, or in any way destroyed, nor will We proceed against or prosecute him, except by the lawful judgment of his peers and by the law of the land.").

¹¹ U.S. CONST. amend. V, VII-VIII.

This evolution of criminal procedures and customs has been accompanied by rigorous academic debate over the goals and rationales of criminal punishment.¹² The progression can be seen from its beginning under Aristotle,¹³ to the nineteenth century works of Immanuel Kant¹⁴ and Jeremy Bentham,¹⁵ and into the twentieth and now twenty-first centuries.¹⁶ Today, the debate largely involves three theories of criminal punishment: utilitarianism, retributivism, and denunciation—although denunciation tends to take a back seat to the first two frameworks.¹⁷

Some modern scholars have attempted to form theories of punishment by mixing the three concepts into a single theory of punishment, or have argued in favor of one theory through appeals to another.¹⁸ A common argument is to support a retributivist system of punishment, but with attempts to appeal to certain utilitarian concerns.¹⁹ The most widely accepted form of this theory is Norval Morris's "limiting retributivism."²⁰ This Comment attempts to further the debate over theories of criminal punishment by investigating whether such retributivist appeals to utilitarianism are valid and can achieve utilitarian goals. Specifically, this Comment seeks to determine whether Morris's limiting retributivism theory, and its appeal to utilitarian concepts, can achieve utilitarian goals.

I undertake to answer this question in five parts. Part II presents a brief historical overview of theories of criminal punishment in order to provide the reader with sufficient background and context to understand the recent movements in this debate. Part III then describes the current movement of scholarship towards the retributivist camp, the several factors that potentially determine the severity of a criminal's punishment under a retributive theory of punishment, and, finally, this camp's attempted appeal to utilitarianism. Part IV describes the concept of limiting retributivism in

¹² See *supra* note 6.

¹³ ARISTOTLE, *supra* note 3, at 200-01.

¹⁴ IMMANUEL KANT, THE PHILOSOPHY OF LAW: AN EXPOSITION OF THE FUNDAMENTAL PRINCIPLES OF JURISPRUDENCE AS THE SCIENCE OF RIGHT 194-204 (W. Hastie trans., Edinburgh, I.&T. Clark 1887) (1796).

¹⁵ JEREMY BENTHAM, AN INTRODUCTION TO THE PRINCIPLES OF MORALS AND LEGISLATION (1781), *reprinted in* 1 THE WORKS OF JEREMY BENTHAM 1, 83-96 (John Bowring ed., Thoemmes Press 1995) [hereinafter MORALS AND LEGISLATION]; JEREMY BENTHAM, PRINCIPLES OF PENAL LAW (1843), *reprinted in* 1 THE WORKS OF JEREMY BENTHAM, *supra*, 365, 396-97.

¹⁶ See *supra* note 6.

¹⁷ JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW 14-19 (4th ed. 2006).

¹⁸ *Id.* at 11-24.

¹⁹ See Norval Morris, *Punishment, Desert and Rehabilitation*, in SENTENCING 257 (Hyman Gross & Andrew von Hirsch eds., 1981).

²⁰ See Ristroph, *supra* note 6, at 1301-02.

greater detail. Part V describes why limiting retributivism's appeal to utilitarianism cannot succeed. In doing so, it utilizes recent behavioral psychology research to demonstrate how the structure of retributivist punishment leads to under-deterrence of crime. Specifically, it highlights how the factors involved in determining punishment under retributivism are affected by cognitive biases. Part VI then acknowledges that the same behavioral psychology research challenges utilitarianism's attempt to create sufficient and optimal *ex ante* incentives. However, while utilitarianism without any adjustment for cognitive biases leads to a one-time under-deterrent effect, limiting retributivism essentially creates a "double" under-deterrent effect. Because of this double effect, limiting retributivism's appeal to utilitarianism fails. Thus, based on this behavioral psychology research, utilitarian goals of punishment, and deterrence in particular, are best served through a purely utilitarian theory of punishment. Finally, this Part provides insights into how utilitarian punishment may be improved through the use of behavioral psychology research.

II. BACKGROUND

Retributive and utilitarian theories of criminal punishment are the two predominant theories of punishment discussed today.²¹ The two theories are often thought of as the two opposing concepts of what criminal punishment should be.²² The debate between the two camps is by no means a recent phenomenon. It has evolved over hundreds of years and will probably continue to advance over the next hundred years.²³ But before we address the history and development of this debate, it is necessary here to give a quick synopsis of each theory.

A. RETRIBUTIVISM

Retributivism posits that punishment is necessary because society must engage in some form of retribution against those who violate its laws.²⁴ Its central tenet defines *punishment* as society's response to a criminal action that has occurred in the past.²⁵ The value of the punishment of the crime is

²¹ DRESSLER, *supra* note 17, at 19.

²² *See id.*

²³ *See, e.g.,* Alschuler, *supra* note 6.

²⁴ *See* GEORG WILHELM FRIEDRICH HEGEL, HEGEL'S PHILOSOPHY OF RIGHT §§ 90-104 (T.M. Knox trans., Oxford Univ. Press 1978) (1821); *see also* KANT, *supra* note 14, at 196 (describing retaliation as the only principle capable of determining "the quality and quantity of a just penalty").

²⁵ *See* R.A. DUFF, TRIALS AND PUNISHMENTS 4 (1986) (noting that retributivism "find[s] the sense and justification of punishment in its relation to a past offence").

in the punishment itself;²⁶ when someone has committed a crime, they simply deserve to be punished.²⁷ "Punishment that gives an offender what he or she deserves for a past crime is a valuable end in itself and needs no further justification."²⁸ Furthermore, society is morally obligated to punish wrongdoers.²⁹

B. UTILITARIANISM

Utilitarian theories of criminal punishment stem from general utilitarian concepts. Joshua Dressler suggests that "[t]he purpose of all laws is to maximize the net happiness of society."³⁰ Therefore, punishment should only be administered if it results in an overall benefit to society.³¹ Only when punishment leads to more aggregate pleasure than aggregate pain is punishment justified.³² As John Robinson and John Darley argue, "Punishment for a past offense is [only] justified by the future benefits it provides."³³ In this manner, utilitarian theories are sometimes referred to as "consequentialist" because they are concerned solely with how punishment will affect future actions and with society's future aggregate happiness.³⁴ Punishment can do this in two main ways. First, it can deter future criminals from committing crimes; second, it can either incapacitate criminals or rehabilitate them so that they cannot or will not engage in future crimes.³⁵

C. HISTORY

The history of these two theories of punishment is extensive. Each has enjoyed the support of legal scholars and intellectual giants, and each has had its periods of dominance.³⁶ While the specifics of the theories may have changed over the years, their foundations have remained constant. Retributivism has always been concerned with punishing criminals as

²⁶ See Stanley I. Benn, *Punishment*, in 7 THE ENCYCLOPEDIA OF PHILOSOPHY 29, 30 (Paul Edwards ed., 1967) ("[T]he punishment of crime is right in itself . . .").

²⁷ See *id.*

²⁸ Robinson & Darley, *supra* note 6, at 454.

²⁹ See John Rawls, *Two Concepts of Rules*, 64 PHIL. REV. 3, 5 (1955).

³⁰ DRESSLER, *supra* note 17, at 14.

³¹ See BENTHAM, MORALS AND LEGISLATION, *supra* note 15, at 83-84.

³² See *id.*

³³ Robinson & Darley, *supra* note 6, at 454.

³⁴ See Christopher, *supra* note 6, at 848 ("Consequentialist theories justify punishment . . . on the actual, good consequences that are attained, for example, deterrence of a crime . . .").

³⁵ Robinson & Darley, *supra* note 6, at 453.

³⁶ See, e.g., Alschuler, *supra* note 6.

simple punishment for their crimes,³⁷ while utilitarianism has valued the punishment of criminals because of the benefit society may reap as a result of such punishment.³⁸

Retributivism can trace its origins far into the past.³⁹ From the time of antiquity through the middle-ages, many criminal justice systems were based largely on the concept of retributivism—the criminal getting what he or she deserved.⁴⁰ Indeed, retributivism can be seen in biblical and Talmudic forms of justice.⁴¹ The existence of early forms of retributivism is not surprising when one considers the emotional aspects of this theory—punishment based in large part on a feeling that the perpetrator deserves the punishment.⁴²

But retributivism was not solely a legal or emotional justification for punishment. It was also justified on moral and philosophical grounds.⁴³ Arguing in support of these justifications for punishment, Immanuel Kant wrote that "[p]unishment can never be administered merely as a means for promoting another Good"⁴⁴ On the contrary, "Punishment ought to be pronounced over all criminals proportionate to their internal wickedness."⁴⁵ In this sense retributive punishment could be thought of as a moral or ideological and philosophical need to condemn wrongdoers. Hegel provided an even more comprehensive justification of retributivism.⁴⁶ He concluded that crimes needed to be negated in order to re-establish equivalence in society, and that negation could only be achieved through punishment.⁴⁷ Retribution was the link between the crime and the punishment.⁴⁸

³⁷ See *infra* text accompanying notes 39-48.

³⁸ See BENTHAM, MORALS AND LEGISLATION, *supra* note 15, at 83.

³⁹ See IGOR PRIMORATZ, JUSTIFYING LEGAL PUNISHMENT 12 (1989).

⁴⁰ See *id.*

⁴¹ *Id.* at 13 ("The history of the retributive view of punishment begins with the biblical and [T]almudic ethical and legal ideas.").

⁴² Cf. PRIMORATZ, *supra* note 39, at 13 (describing the moral aspect of retributive punishment, which this Comment posits will often equate to an emotional aspect of retributive punishment).

⁴³ See *id.*

⁴⁴ KANT, *supra* note 14, at 195.

⁴⁵ IMMANUEL KANT, THE SCIENCE OF RIGHT (1790), *reprinted in* 42 GREAT BOOKS OF THE WESTERN WORLD: KANT 447 (Robert Maynard Hutchins ed., 1989).

⁴⁶ See Markus Dirk Dubber, *Rediscovering Hegel's Theory of Crime and Punishment*, 92 MICH. L. REV. 1577, 1577-78 (1994).

⁴⁷ *Id.* at 1581-82.

⁴⁸ HEGEL, *supra* note 24, at § 101.

Utilitarian theories of punishment, while perhaps not as old as retributivism, may also date back several millennia.⁴⁹ While early retributivism may have called for a thief to have his or her hand cut off (also potentially a utilitarian response), it is possible that a decision to imprison the thief instead of, or in addition to, cutting off his or her hand reflected utilitarian values. After all, the utilitarian principle of incapacitation asserts that while in prison, the thief is no threat. Seen in this light, utilitarian concepts may date back almost to the beginnings of civilization. Indeed, utilitarian justifications for punishment can be seen as early as in the work of Plato. In *Protagoras*, Plato argued that punishment for punishment's sake was "taking blind vengeance like a beast."⁵⁰ He continued, "No, punishment is not inflicted by a rational man for the sake of the crime that has been committed—after all one cannot undo what is past"⁵¹

The medieval and early modern eras witnessed the continued development of utilitarian thought. St. Thomas Aquinas argued that the purpose of punishment required that some good emerge from the punishment.⁵² Later in the millennium, Thomas Hobbes indicated support for utilitarian forms and rationales of punishment. He argued that society should be "forbidden to inflict punishment with any other design[], than for the correction of the offender, or direction of others."⁵³ The utilitarian rationale in this statement appears undeniable.

Yet it was not until the writings of Jeremy Bentham in the eighteenth and nineteenth centuries that utilitarianism was most clearly articulated.⁵⁴ Although it was John Stewart Mill who actually coined the term *utilitarianism* to describe the doctrine,⁵⁵ it was Jeremy Bentham who first proposed his theory that society's goal should be to maximize utility, that is, to maximize the amount of aggregate pleasure and minimize the amount

⁴⁹ A more recent example can be seen in the work of John Locke. LOCKE, *supra* note 5, at 273 (noting that a crime may be punished for future restraint).

⁵⁰ PLATO, *PROTAGORAS*, reprinted in *THE COLLECTED DIALOGUES OF PLATO* 308, 321 (Edith Hamilton & Huntington Cairns eds., 1982) (stating that punishment is permissible if "the [punisher's] intention be directed chiefly to some good, to be obtained by means of the punishment of the person who has sinned").

⁵¹ *Id.*

⁵² II ST. THOMAS AQUINAS, *SUMMA THEOLOGICA*, in *20 GREAT BOOKS OF THE WESTERN WORLD*, second part, part I, question 92, art. 2, at 215 (Robert Maynard Hutchins ed., founders ed. 1952) ("And it is the fear of punishment that law makes use of in order to ensure obedience . . .").

⁵³ HOBBS, *supra* note 5, at 106.

⁵⁴ See BENTHAM, *MORALS AND LEGISLATION*, *supra* note 15, at 83-86.

⁵⁵ JOHN STUART MILL, *UTILITARIANISM* (Oskar Piest ed., Bobbs-Merrill Co. 1957) (1861).

of aggregate pain in society.⁵⁶ Criminal punishment, therefore, was only acceptable if it increased future pleasure or decreased future pain. Using his theory of utility, Bentham was the first to clearly argue that prevention of future crimes should be the primary goal of punishment. "General prevention ought to be the chief end of punishment, as it is its real justification."⁵⁷ Because criminal acts decrease aggregate happiness in society, prevention of such acts increases aggregate happiness.

The influence of Bentham's and others' works on utilitarianism and their influences on the criminal justice system cannot be overstated. Bentham, in particular, accurately touched on all the issues mentioned by previous academics and philosophers, and weaved those concerns into a comprehensive and conclusive theory that could be applied to life in general.⁵⁸

Following Bentham, we see the emergence of applications of utilitarian principles of punishment by prominent legal scholars.⁵⁹ Several famous American scholars and judges adopted and further developed utilitarian theories of punishment. Oliver Wendall Holmes, Jr., for example, wrote in 1881 that "prevention would . . . seem to be the chief and only universal purpose of punishment."⁶⁰ He openly supported deterrence as a chief principle of punishment:

If I were having a philosophical talk with a man I was going to have hanged . . . I should say, I don't doubt that your act was inevitable for you but to make it more avoidable by others we propose to sacrifice you to the common good. You may regard yourself as a soldier dying for your country if you like. But the law must keep its promises.⁶¹

John H. Wigmore described deterrence as "the kingpin of the criminal law."⁶²

As the twentieth century progressed, the movement towards utilitarian theories of punishment grew in its influence.⁶³ Even the Supreme Court

⁵⁶ See BENTHAM, *MORALS AND LEGISLATION*, *supra* note 15, at 1-13.

⁵⁷ BENTHAM, *PRINCIPLES OF PENAL LAW*, *supra* note 15, at 396.

⁵⁸ See PRIMORATZ, *supra* note 39, at 13.

⁵⁹ See *infra* text accompanying notes 60-62.

⁶⁰ OLIVER WENDELL HOLMES, JR., *THE COMMON LAW* 46 (Little, Brown & Co. 1945) (1881).

⁶¹ Letter from Oliver Wendell Holmes, Jr., to Harold J. Laski (Dec 17, 1925), in 1 *HOLMES-LASKI LETTERS: THE CORRESPONDENCE OF MR. JUSTICE HOLMES AND HAROLD J. LASKI, 1916-1925*, at 806 (Mark DeWolfe Howe ed., 1953).

⁶² Harry Olson, Homer Cummings & John H. Wigmore, *A Symposium of Comments from the Legal Profession: The Loeb-Leopold Case (Concluded)*, 15 J. AM. INST. CRIM. L. & CRIMINOLOGY 395, 401 (1924).

⁶³ However, this period can be differentiated between adoptions of rehabilitation and deterrence as governing theories. See Alschuler, *supra* note 6, at 6-14.

appeared to have found the virtues of the philosophy. In 1949, the Court held that "[r]etribution is no longer the dominant objective of the criminal law."⁶⁴ Three years later the Court spoke of the "tardy and unfinished substitution of deterrence and reformation in place of retaliation and vengeance as the motivation of public prosecution."⁶⁵

The latter half of the twentieth century saw a continued development of utilitarian principles that nicely dovetailed into the philosophy of the emerging law and economics movement.⁶⁶ This movement sought to craft legal rules that created *ex ante* incentives that would lead to efficient outcomes in a variety of situations.⁶⁷ The utilitarian concepts of using criminal punishments to decrease the amount of future crime through deterrence played right into the law and economics movement and its use of *ex ante* incentives to create optimal care and activity levels.⁶⁸ By creating a bright-line rule of harsh punishments, potential criminals are incentivized not to commit a crime.⁶⁹

As discussed above, during this period incapacitation and deterrence became important goals of the criminal justice system.⁷⁰ Professor Albert Alschuler highlights the use of law-and-economics-based deterrence policies during the period with the example of federal income tax evasion penalties.⁷¹ Even though physically stealing \$100,000, for example, from the government would appear to most of us as worse than avoiding paying \$100,000 worth of income tax, income tax evasion was punished much more severely than outright theft.⁷² The justification was that tax evaders were caught less often than thieves; therefore, harsher penalties were necessary to create a sufficient deterrent effect.⁷³

As the twentieth century wore on, utilitarian justifications for punishment seemed to be gaining a foothold in American criminal jurisprudence.⁷⁴ Yet this blossoming of acceptance of utilitarian concepts

⁶⁴ *Williams v. New York*, 337 U.S. 241, 248 (1949).

⁶⁵ *Morissette v. United States*, 342 U.S. 246, 251 (1952).

⁶⁶ *Cf. Alschuler, supra* note 6, at 11 (noting economists' views on criminal punishment).

⁶⁷ *See, e.g., Steven Shavell, Criminal Law and the Optimal Use of Nonmonetary Sanctions as a Deterrent*, 85 COLUM. L. REV. 1232 (1985).

⁶⁸ *See id.*

⁶⁹ *See id.*

⁷⁰ *See Alschuler, supra* note 6, at 11.

⁷¹ *Id.*

⁷² *Id.*

⁷³ *Id.* at 11 n.61 ("Especially in light of the limited number of criminal tax prosecutions relative to the estimated incidence of such violations . . . deterring others from violating the tax laws is the primary consideration underlying these guidelines." (citing U.S. SENTENCING GUIDELINES AND POLICY STATEMENTS § 2T4.1 (April 1987))).

⁷⁴ *Id.* at 9-12.

of punishment was not to last. A "renaissance" of retributivist thought was soon to challenge utilitarian theory.⁷⁵ This recent movement towards retributivism as the predominant school of American criminal jurisprudence is discussed in the next Part.

III. MODERN RETRIBUTIVISM

As recently as the middle to late twentieth century, many scholars and judges were still under the assumption that retributive theories of punishment had either been thoroughly defeated or served little to no purpose in the criminal justice system.⁷⁶ In 1972, Justice Thurgood Marshall declared that "no one has ever seriously advanced retribution as a legitimate goal of our society."⁷⁷ A criminal law scholar in the 1950s claimed that "[p]unishment as retribution belongs to a penal philosophy that is archaic and discredited by history."⁷⁸ Even the drafters of the Model Penal Code seemed to virtually ignore the role of retributivism in criminal punishment.⁷⁹ The 1962 draft included retribution as a limiting principle while seemingly including every other plausible justification as a legitimate and primary rationale and purpose for punishment.⁸⁰ In such a hostile atmosphere, a reemergence of retributivism seemed highly unlikely; yet, exactly such a reemergence has occurred, and with great strength and influence.⁸¹

Over the last quarter of the twentieth century and into the early part of the twenty-first century, retributivism has reestablished itself as the dominant theory behind criminal justice.⁸² That this revolution in thought about criminal punishment has occurred is disputed by few scholars⁸³ and can be seen in a plethora of examples.⁸⁴ One scholar noted that "[t]here has been a steady rise in the popularity of retributivism over the last decade,

⁷⁵ See Martin R. Gardner, *The Renaissance of Retribution—An Examination of Doing Justice*, 1976 WIS. L. REV. 781, 781.

⁷⁶ See *supra* text accompanying notes 63-75; *infra* text accompanying notes 77-80.

⁷⁷ *Furman v. Georgia*, 408 U.S. 238, 363 (1972) (Marshall, J., concurring).

⁷⁸ Austin MacCormick, *The Prison's Role in Crime Prevention*, 41 J. CRIM. L. & CRIMINOLOGY 36, 40 (1951).

⁷⁹ See MODEL PENAL CODE § 1.02 (Proposed Official Draft 1962).

⁸⁰ *Id.*

⁸¹ See Christopher, *supra* note 6, at 845-47.

⁸² *Id.*

⁸³ But see Gerald Leonard, *Towards a Legal History of American Criminal Theory: Culture and Doctrine from Blackstone to the Model Penal Code*, 6 BUFF. CRIM. L. REV. 691, 692 n.6 (2003) (arguing that there has not been a revolution of retributivist thinking).

⁸⁴ See Christopher, *supra* note 6, at 845; Gardener, *supra* note 75, at 784; Ristroph, *supra* note 6, at 1293.

which is surprising given its near death in the 1950s and 1960s.”⁸⁵ Such sentiments are seen through the works of several other criminal law scholars.⁸⁶

These scholarly insights about the reemergence of retributivism were based on strong examples in a number of contexts. In *Spaziano v. Florida*, the United States Supreme Court acknowledged the role of retribution in criminal justice.⁸⁷ The Court stated that retribution “is an element of all punishments society imposes.”⁸⁸ State courts, too, adopted a more retributivist slant when interpreting criminal punishment statutes.⁸⁹ Additionally, several states adopted retributivist principles of punishment into their penal codes.⁹⁰ Pennsylvania’s penal code went so far as openly advocating retribution as the primary purpose of criminal punishment: “The sentencing guidelines provide sanctions proportionate to the severity of the crime [that] . . . establishes a sentencing system with a primary focus on retribution”⁹¹

While the specific event that triggered the reemergence of retributivism is highly debated, the reasons for its reemergence are more agreed upon.⁹² Although several reasons existed for the growth, three are predominant. First, retributivism altered its rhetoric away from talk of retribution and towards criminals getting their just deserts.⁹³ In other words, retributivism no longer spoke of victims or society taking vengeance on the criminal, but instead spoke of the criminal getting what he or she deserved for committing the crime. Second, retributivism received a facelift in the form of much scholarly development of its concepts and philosophical arguments and justifications.⁹⁴ This development helped modernize retributivist thought and moved retributivism away from its

⁸⁵ Jean Hampton, *Correcting Harms Versus Righting Wrongs: The Goal of Retribution*, 39 UCLA L. Rev. 1659, 1659 (1992).

⁸⁶ See, e.g., Richard S. Frase, *Limiting Retributivism*, in THE FUTURE OF IMPRISONMENT 83, 90-104 (Michael Tonry ed., 2004); Gardener, *supra* note 75, at 784; Kyron Huigens, *The Dead End of Deterrence, and Beyond*, 41 WM. & MARY L. REV. 943, 980 n.152 (2000); Sanford H. Kadish, *Fifty Years of Criminal Law: An Opinionated Review*, 87 CAL. L. REV. 943, 978 (1999).

⁸⁷ 468 U.S. 447, 462 (1984).

⁸⁸ *Id.*

⁸⁹ See Michelle Cotton, *Back with a Vengeance: The Resilience of Retribution as an Articulated Purpose of Criminal Punishment*, 37 AM. CRIM. L. REV. 1313, 1326-27 (2000) (citing examples of state courts adopting retributivist rationales for criminal punishment).

⁹⁰ See, e.g., Act Operative July 1, 1977, ch. 4.5, 1976 Cal. Stat. 5140 (codified as amended at CAL. PENAL CODE § 1170(a)(1) (West 2004 & Supp. 2009)).

⁹¹ 204 PA. CODE § 303.11 (2008).

⁹² Christopher, *supra* note 6, at 846 n.6.

⁹³ See Ristroph, *supra* note 6, at 1299-1301.

⁹⁴ See *id.* at 1299-1300.

archaic image of an "eye for an eye" punishment.⁹⁵ Lastly, many retributivists developed and adopted new systems of retributivism that attempted to appeal to other theories of punishment.⁹⁶ For purposes of this Comment, I focus on limiting retributivism's appeal to utilitarianism, and will now discuss the three developments in great detail.

The rhetoric of retributivism had always been one of its weakest points.⁹⁷ A system that speaks of the basis of punishment in terms of vengeance is bound to evoke bad connotations similar to reactions towards vigilante justice.⁹⁸ This rhetoric leads to retributivism being "portray[ed] . . . as glorified vengeance, [a] celebrat[ed] and codif[ied] . . . brutal and inhumane impulse to harm those who harm us."⁹⁹ In response to these negative images associated with the term retribution, retributivism scholars have adopted less severe and more egalitarian-sounding rhetoric.¹⁰⁰ The newly chosen motto was "just deserts." When a person commits a crime, his or her actions deserve punishment in accordance with the severity of the crime and his or her moral culpability.¹⁰¹

This change in rhetoric has largely achieved its goals.¹⁰² "The emphasis on desert seems to have helped disassociate retribution from revenge, for it allows punishment theorists to draw on a concept that has more neutral philosophical status."¹⁰³ Thus, instead of being punished because of revenge or hatred, the criminal receives his or her just deserts as punishment.¹⁰⁴

This change in rhetoric accompanied a change in retributivism's underlying philosophy as well.¹⁰⁵ While opinions differ within the philosophy of modern retributivism, there is consistency in the emergence of the use of egalitarian principles within those theories.¹⁰⁶

⁹⁵ See *id.* at 1299-1306.

⁹⁶ See *id.*

⁹⁷ See *id.*

⁹⁸ See *id.* at 1298-99.

⁹⁹ *Id.* at 1299.

¹⁰⁰ *Id.*

¹⁰¹ See Frase, *supra* note 6, at 73 (citing ANDREW VON HIRSCH, *CENSURE AND SANCTIONS* 29-33 (1993); Richard S. Frase, *Excessive Prison Sentences, Punishment Goals, and the Eighth Amendment: "Proportionality" Relative to What?*, 89 MINN. L. REV. 571, 590 (2005)).

¹⁰² See *supra* text accompanying notes 81-86.

¹⁰³ Ristroph, *supra* note 6, at 1300.

¹⁰⁴ See *id.*

¹⁰⁵ See Herbert Morris, *Persons and Punishment*, in *SENTENCING*, *supra* note 19, at 93.

¹⁰⁶ See Ristroph, *supra* note 6, at 1300 ("These egalitarian arguments have won much support.").

Herbert Morris was one of the pioneering supporters of the new theoretical basis for retributivism.¹⁰⁷ Morris argued that the law should create equal rights and responsibilities for all members of society.¹⁰⁸ When an individual commits a crime, the perpetrator violates this equal distribution.¹⁰⁹ The criminal receives the benefit of others not committing crimes, but violates his responsibility not to commit crimes against others.¹¹⁰ Punishment, therefore, is simply society restoring the balances of benefits and burdens of the law.¹¹¹

Other retributivists consider a criminal act to consist of the criminal declaring he is more important than the victim.¹¹² By delivering just deserts as punishment, society signals that the criminal is not, in fact, more valuable than the victim, demonstrating the victim's equality through the punishment of the criminal.¹¹³

Lastly, retributivists attempted to reconcile retributivism with utilitarianism to gain adherents among traditional supporters of utilitarian theories of punishment.¹¹⁴ While, historically, retributivism and utilitarianism were viewed as diametrically opposed and mutually exclusive theories of punishment, late twentieth century scholars were able to create a system that bridged the gap.¹¹⁵

The most influential of these scholars was Norval Morris, not to be confused with Herbert Morris.¹¹⁶ Morris established the limiting retributivism model of criminal punishment.¹¹⁷ Limiting retributivism did not follow the typical retributivist concept that punishment should be based and predicated on a criminal getting his or her just deserts.¹¹⁸ Instead, desert and moral blameworthiness served as a limiting principle, "a principle that, though it would rarely tell us the exact sanction to be

¹⁰⁷ See Morris, *supra* note 105, at 93-95.

¹⁰⁸ *Id.*

¹⁰⁹ *Id.*

¹¹⁰ *Id.* at 95 ("[I]t is just to punish those who have violated the rules and caused the unfair distribution of benefits and burdens. A person who violates the rules has something others have—the benefits of the system—but by renouncing what others have assumed, the burdens of self-restraint, he has acquired an unfair advantage.").

¹¹¹ See WOJCIECH SADURSKI, *GIVING DESERT ITS DUE: SOCIAL JUSTICE AND LEGAL THEORY* 229 (1985).

¹¹² See, e.g., Jean Hampton, *An Expressive Theory of Retribution*, in *RETRIBUTIVISM AND ITS CRITICS* 1, 5-6 (Wesley Cragg ed., 1992).

¹¹³ *Id.* at 6-11.

¹¹⁴ See, e.g., Morris, *supra* note 19, at 257-59.

¹¹⁵ See DRESSLER, *supra* note 17, at 19.

¹¹⁶ See Morris, *supra* note 19, at 257-59.

¹¹⁷ *Id.*

¹¹⁸ *Id.*

imposed . . . would nevertheless give us the outer limits of leniency and severity which should not be excluded."¹¹⁹ Despite the fact that Morris used retributivism as a limiting principle, his theory can still be seen as located squarely within the retributivist camp of theories of punishment.

Morris acknowledged the fundamental difficulties in determining with precision a prisoner's just deserts. Instead, retributivism would often lead to a range of possible penalties.¹²⁰ While we can be sure that some penalty is either too little or too much, we cannot always know exactly what the correct retributivist penalty should be.¹²¹ This leads to Morris's use of retributivism as a limiting principle instead of a guiding principle, as mentioned above.¹²²

Within the range of penalties dictated by retributivism's limiting principles, other theories of criminal punishment are utilized to increase or decrease the sentence to achieve those theories' goals.¹²³ One such theory, supported strongly by Morris himself, was "parsimony," which essentially held that a criminal's punishment should be no more severe than needed.¹²⁴ This served as a further limiting principle that had to be addressed before any other considerations could result in an increase or decrease in punishment within the acceptable range. For this Comment, however, the most important theory to which Morris's limited retributivism attempted to appeal was utilitarianism.

Limiting retributivism appealed to utilitarian concerns in two main ways. First, and of lesser concern, it created increased uniformity in sentences, which was believed to be of utilitarian value.¹²⁵ More uniform sentences would create a clear indication of the level of punishment that a criminal would receive and thus was thought to have a deterrent effect.¹²⁶ Second, and more importantly, within sentence parameters a judge could factor in utilitarian concerns such as incapacitation or deterrence to increase or decrease the severity of the sentence.¹²⁷ Thus, "desert permits the reasonable pursuit of utilitarian aims even as it forestalls the dangers of

¹¹⁹ *Id.* at 259.

¹²⁰ *Id.*

¹²¹ See Ristroph, *supra* note 6, at 1301-02 (noting that "[t]he 'limiting retributivism' model attributed to Norval Morris begins with the recognition that our intuitions about how much punishment a given offender deserves are often imprecise").

¹²² See, e.g., HERBERT L. PACKER, *THE LIMITS OF THE CRIMINAL SANCTION* 66 (1968) ("I see an important limiting principle in the criminal law's traditional emphasis on blameworthiness But it is a *limiting* principle, not a justification for action.").

¹²³ See Morris, *supra* note 19, at 264.

¹²⁴ See NORVAL MORRIS, *THE FUTURE OF IMPRISONMENT* 60-62 (1974).

¹²⁵ See Frase, *supra* note 6, at 74-75.

¹²⁶ *Id.*

¹²⁷ *Id.* at 76-77.

excessive utilitarianism.”¹²⁸ In light of these developments, Norval Morris’s theory of limiting retributivism seemed to succeed at bridging the gap between retributivism and utilitarianism.

As a result of its ability to appeal to both retributivism and utilitarianism, limiting retributivism became a widely accepted theory of criminal punishment.¹²⁹ Some sentencing guidelines, where theories of criminal punishment are put to practice, are evidence of such acceptance. Some of the guidelines, Minnesota’s for example, seem to be adopting limiting retributivism as their guiding principles.¹³⁰

IV. LIMITING RETRIBUTIVISM

Because of its extreme success,¹³¹ limiting retributivism is the model of retributivism I will focus on in the remaining portions of this Comment. Here, a more detailed description of limiting retributivism is necessary to highlight how it functions in practice, how it specifically appeals to utilitarianism, and what factors are involved in determining the upper and lower bounds of acceptable levels of punishment.

As previously mentioned, limiting retributivism seeks to create a system of criminal punishment in which retributivist principles determine upper and lower limits of a criminal’s potential range of punishment.¹³² The criminal’s culpability and moral blameworthiness are the relevant factors in assessing the upper and lower limits of punishment.¹³³ While determining an appropriate, specific punishment may create difficulties for scholars and practitioners, creating a range of possible punishments is easier, although not without its own difficulties.

Professor Martin Redish claims that when interpreting constitutional text, you may not always know when your interpretation is the correct one, but you do know when it is the wrong one.¹³⁴ The text provides a range of

¹²⁸ Ristroph, *supra* note 6, at 1302.

¹²⁹ One scholar noted that limiting retributivism has become a “widely endorsed and adopted model.” Frase, *supra* note 6, at 76. Another stated that it is “[o]ne of the most widely followed reconciliations of desert with utilitarian aims.” Ristroph, *supra* note 6, at 1301.

¹³⁰ Minnesota’s sentencing guidelines seem to have adopted limiting retributivism to some degree. See Richard S. Frase, *Sentencing Guidelines in Minnesota: 1978–2003*, 31 CRIME & JUST. 131, 149 (2005) (“Thus, for most defendants, the guidelines essentially retain the traditional indeterminate sentencing system and its utilitarian values, subject only to retributive ‘caps’ set by the presumptive duration of stayed prison terms.”).

¹³¹ See Frase, *supra* note 6, at 76; Ristroph, *supra* note 6, at 1301.

¹³² See *supra* text accompanying notes 116–24.

¹³³ ANDREW VON HIRSCH, *DOING JUSTICE: THE CHOICE OF PUNISHMENTS* 69 (1976).

¹³⁴ See Martin H. Redish & Karen L. Drizin, *Constitutional Federalism and Judicial Review: The Role of Textual Analysis*, 62 N.Y.U. L. REV. 1, 22 (1987) (citing Frederick

possible interpretations, and it should be quite clear when you have exceeded that range.¹³⁵ This same idea can apply to creating an acceptable range of punishments for a criminal. You cannot always know if a certain punishment is correct, but you can know if it is either too severe or too lenient. If a criminal steals a car, we intuitively know that a one-day prison sentence is far too low, while a fifty-year sentence is far too high. Yet we may not know if a ten-year or a fifteen-year sentence is more appropriate.¹³⁶

While upper and lower limits can be very roughly determined through intuitive retributivist beliefs given a specific crime, how do we determine different ranges of punishments from one type of crime to another? In other words, how do limiting retributivists determine that an appropriate range of punishments for theft is one to five years, for example, while an appropriate range for rape is ten to twenty? The answer is much the same as how limiting retributivists determine the appropriate range for a single crime: there are intuitive sensibilities about punishments. There is a general belief that some crimes are more morally reprehensible than others.¹³⁷ People tend to agree that rape is a far more reprehensible crime than mere theft. These general beliefs about the varying degrees of reprehensibility of crimes are supported by empirical evidence.¹³⁸ Almost all people, across all groups, when asked to rank crimes from least-worst to worst, rank them in similar manners.¹³⁹ This evidence bolsters limiting retributivism's appeal to intuition to create upper and lower limits by demonstrating that such intuition is inherent in human nature and, therefore, in some sense, not arbitrary. Although the fact that differing crimes receive different levels of punishment seems obvious, it is still important to note that the range of punishment from crime *A* to crime *B* is justified in a similar manner as that of ranges of potential punishments within a single type of crime. Both differences are based on moral culpability. This fact can be seen in the Model Penal Code, which ranks crimes and states of required knowledge, implying differing levels of moral culpability for each mens rea.¹⁴⁰

Once we have accepted that limiting retributivism can successfully create a range of acceptable punishments, we are faced with how utilitarian

Schauer, *An Essay on Constitutional Language*, 29 UCLA L. REV. 797, 828 (1982)).

¹³⁵ See *id.*

¹³⁶ This example is taken from Ristroph, *supra* note 6, at 1302.

¹³⁷ Cf. VON HIRSCH, *supra* note 133, at 66-71 (arguing that the existence of differing punishments for different crimes implies that certain crimes are worse than others).

¹³⁸ Joseph E. Jacoby & Francis T. Cullen, *The Structure of Punishment Norms: Applying the Rossi-Berk Model*, 89 J. CRIM. L. & CRIMINOLOGY 245 (1998).

¹³⁹ *Id.*

¹⁴⁰ See Kevin R. Reitz, *Reporter's Introduction to MODEL PENAL CODE: SENTENCING* 1, 3-4 (Discussion Draft 2006).

principles may be met and goals achieved within those ranges. Here, utilitarian concepts such as incapacitation or deterrence are utilized.¹⁴¹ The judge or sentencing guidelines determine the range of punishments available for certain crimes, given the type of crime and the criminal's culpability and moral blameworthiness, and then utilitarian concerns are used to determine the adequate punishment within that range.¹⁴²

According to supporters of limiting retributivism, these adoptions of utilitarian principles can have demonstrable effects and thus serve utilitarian purposes.¹⁴³ Limiting retributivism's use of utilitarian principles can serve utilitarianism's twin goals of incapacitation and deterrence; examples of each follow.

Suppose a person has robbed a grocery store. The upper and lower bounds for this theft are one to ten years. The trial judge discovers that this individual has been previously convicted of theft, assault, and attempted robbery, and received jail terms of one year, six months, and one year, respectively. Given this individual's demonstrated propensity to commit crime, utilitarians consider this criminal a threat to society—that is to say, the criminal's demonstrated propensity to commit crimes indicates he is likely to commit future crimes and thus is bound to cause pain to others. Therefore, society will be better off if this individual is put in jail for a long time where he cannot hurt others. Thus, with these considerations in mind, a judge may sentence this individual to ten years in prison, the maximum allowed. This same logic is behind the "three strikes and you're out" rule of many state penal codes.¹⁴⁴ This example is a clear demonstration of an appeal to incapacitation principles of utilitarianism.

Now let us consider another hypothetical. An individual is convicted of attempted rape. The range of acceptable punishments for attempted rape is two to twenty years. Let us also assume that this crime was committed in a state or county where it was known that a very low percentage of rape or attempted rape victims took actions against their perpetrators and pursued criminal charges. In this regard, a very low percentage of people committing rape or attempted rape end up being charged with the crime, and an even lower percentage will be found guilty and sentenced to a term of imprisonment. In light of this low rate, a very high level of punishment may be required to create sufficiently strong incentives that will convince potential rapists that they should not commit the crime. Simple economic calculations support this statement. If there is a 10% chance of being found

¹⁴¹ See MORRIS, *supra* note 124, at 60-62.

¹⁴² See *id.*; see also Frase, *supra* note 6, at 76-77 (explaining the concept of limiting retributivism and Morris's role in its development).

¹⁴³ See MORRIS, *supra* note 124, at 60-62.

¹⁴⁴ See, e.g., CAL. PENAL CODE § 667(e)(2) (West 1999).

guilty and an expected jail time of five years, then the expected harm to the rapist is a six-month jail sentence. On the other hand, if the expected jail time is twenty years, the expected harm is a two-year sentence. According to simple law and economic principles, the higher expected harm will create greater incentives not to commit the crime.¹⁴⁵ Given this information, the judge decides to award twenty years in an attempt to demonstrate to the community that one found guilty of attempted rape will face twenty years imprisonment.¹⁴⁶ This decision to sentence one found guilty of rape or attempted rape to the highest possible prison term, according to the law and economics theory previously mentioned, creates stronger incentives not to commit rapes, and therefore serves as a deterrent, the other main goal of utilitarianism.¹⁴⁷ This example demonstrates limiting retributivism's appeal to deterrence.

But this leaves us with our last, and, as we will see, most important question. Retributivists claim that the level of culpability and moral blameworthiness of the offender determines the upper and lower bounds of permissible punishment.¹⁴⁸ This seems simple enough, yet how are culpability and moral blameworthiness quantified? I cannot imagine a retributivist would ever claim that a man with Down syndrome was as culpable or morally abject for committing assault as a full-grown man with average intellect. Neither, I imagine, would a retributivist say the permissible range of punishments should be the same for a young child who commits theft as for an adult. The question is then: what factors are utilized when determining the upper and lower bounds of permissible punishment, given the fact that people committing the same crime may have varying degrees of culpability and moral blameworthiness?

Scholars and judges have dealt with this problem by viewing desert on a case-specific basis.¹⁴⁹ By looking at the specifics of each case, the judge can determine the most appropriate range of punishment in line with the defendant's culpability and moral blameworthiness, and the severity of the crime.¹⁵⁰ "Desert is typically understood to be a highly individual matter, in the sense that to be deserving or not is dependent on particular traits of the individual."¹⁵¹ One scholar goes as far as describing desert as based "as

¹⁴⁵ See Shavell, *supra* note 67, at 1235.

¹⁴⁶ This example is similar to the federal income tax evasion example mentioned *supra*, in the text accompanying notes 70-73.

¹⁴⁷ See Shavell, *supra* note 67, at 1235.

¹⁴⁸ See generally Ristorph, *supra* note 6.

¹⁴⁹ *Id.* at 1321.

¹⁵⁰ *Id.*

¹⁵¹ *Id.*

much or more on circumstances and personal characteristics as on physical actions and harm."¹⁵²

Several commentators have acknowledged that each individual's desert may differ greatly.¹⁵³ Various conditions or situations exist that may increase or decrease a criminal's culpability and moral blameworthiness, and hence alter his or her desert.¹⁵⁴ Such factors range from mental ability to age.¹⁵⁵ I refer to these issues of personal characteristics and circumstances as "aggravating" and "mitigating" factors. An aggravating factor is any factor that increases a criminal's desert, while a mitigating factor is any factor that decreases a criminal's desert. While the academic discussion of the list of possible aggravating and mitigating factors is nowhere near complete, I posit that based on the underlying structure of retributivism and its dependency on moral blameworthiness, that list will be lengthy.

Aggravating and mitigating factors play a large role in determining a criminal's just desert. For example, how blameworthy is a poor baker stealing bread from his employer to feed his family? What about a situation in which a man thinks his actions are saving the lives of many, yet actually result in several deaths? Or a ten-year-old child who accidentally shoots and kills his sister? Most would not argue that these three people are just as morally culpable and blameworthy as those who stole or murdered in cold blood, for fun or revenge. I believe these three situations illustrate how many factors could be used to alter the range of potential punishments under a limiting retributivism theory of punishment.¹⁵⁶

In conclusion, limiting retributivism, in its current form, consists of determining a range of potential punishments by utilizing all the various aggravating and mitigating factors that affect the defendant's moral culpability, and then altering the punishment within those parameters to serve utilitarian purposes. As we will see in the following Part, this system

¹⁵² Alschuler, *supra* note 6, at 19.

¹⁵³ See, e.g., JOEL FEINBERG, *Justice and Personal Desert*, in *DOING AND DESERVING: ESSAYS IN THE THEORY OF RESPONSIBILITY* 55, 61 (1970).

¹⁵⁴ See, e.g., Kadish, *supra* note 86.

¹⁵⁵ See Ristroph, *supra* note 6, at 1314-27.

¹⁵⁶ Although I have not seen any scholarly writings that have argued this point, I believe it is a natural and logical progression of the concept of retributivism in general. Given that moral blameworthiness and culpability are what defines an individual's desert, I cannot see how the examples I illustrated would not affect one's level of desert. In order to preclude such factors from entering the debate of what the individual's just deserts are, retributivists will need to create a different rationale or measuring stick for desert. A possibility could be having some sort of free will requirement, in other words, describing moral blameworthiness as being able to understand the import of your decision, its likely outcome, yet nonetheless going forward with the action.

is questionable because the existence of so many aggravating and mitigating factors itself will lead to severe under-deterrence, due to the effects of cognitive biases. Regardless of the resulting adjustments for utilitarian concerns within the acceptable range of punishments, the overall deterrent effect of this system is inadequate. The following Part describes why utilitarian concerns about deterrence and *ex ante* incentives are not adequately addressed by limiting retributivism.

V. LIMITING RETRIBUTIVISM, UTILITARIANISM, AND COGNITIVE BIASES

Over the past several decades, behavioral psychologists have made numerous discoveries that have questioned previous assumptions about human behavior.¹⁵⁷ These discoveries have questioned the neo-classical image of human beings as rational beings capable of making well-informed, reasoned decisions.¹⁵⁸ In contrast to the optimal rational behavior humans were imagined to possess, they "exhibit bounded rationality, bounded self-interest, and bounded willpower."¹⁵⁹ In other words, human psychology and behavior prevent people from acting or behaving as entirely rational beings, as they have often been thought to be. With these new insights into the true nature of human psychology, several old theories, specifically theories of criminal punishment, must be adjusted to accommodate this new information.

In essence, behavioral psychologists have discovered over the past several decades what many already thought to be true: human beings simply are not rational.¹⁶⁰ Yet research has shown that the extent to which human beings are irrational, and the impact of this irrational behavior, is far greater than previously predicted.¹⁶¹ Errors in human judgment are not simply the effect of emotion creeping into the decision-making process and temporarily overriding the individual's rationality; rather, they are the effect

¹⁵⁷ Ironically, the decision to engage in such research was, in large part, a reaction to the law and economics movement's description of human beings as rational decision-makers. See William M. Goldstein & Robin M. Hogarth, *Judgment and Decision Research: Some Historical Context*, in RESEARCH ON JUDGMENT AND DECISION MAKING 3, 5 (William M. Goldstein & Robin M. Hogarth eds., 1997); Robin M. Hogarth & Melvin W. Reder, *Introduction* to RATIONAL CHOICE: THE CONTRAST BETWEEN ECONOMICS AND PSYCHOLOGY 1, 5 (Robin M. Hogarth & Melvin W. Reder eds., 1987).

¹⁵⁸ See *supra* note 157.

¹⁵⁹ Christine Jolls, Cass R. Sunstein & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1471 (1998).

¹⁶⁰ *Id.* at 1473.

¹⁶¹ Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: The Problem of Market Manipulation*, 74 N.Y.U. L. REV. 630, 633 (1999).

of systematic cognitive biases that affect all of human nature.¹⁶² Studies have demonstrated that human beings are, in many ways, fundamentally irrational. By demonstrating the existence of such cognitive biases as the hindsight bias,¹⁶³ the entity effect,¹⁶⁴ and the representative heuristic effect, behavioral psychologists have demonstrated just how irrational we are.¹⁶⁵

Hindsight bias refers to the phenomenon in which human beings, once knowledgeable of the outcome of some situation, develop a sincere belief that they would have predicted the actual outcome.¹⁶⁶ This may seem like common sense—hindsight is 20/20—yet the extent of this bias is remarkable. Suppose a patient comes into a hospital with a set of symptoms. Doctors know that given these symptoms, there is a 33% chance the patient has disease *A*, a 33% chance that the patient has disease *B*, and a 33% chance that the patient has disease *C*. Before actual diagnosis, several doctors, knowledgeable about all three diseases and their symptoms, are asked the probability that a patient with this set of symptoms has each of the three diseases. Virtually every doctor will say there is a 33% chance that the patient has each disease. The doctors are then presented with a description of a patient with the exact same symptoms, but told the patient had disease *B*. They are then asked, if they had seen this patient before he was diagnosed, what would be the predicted probability of the patient having each of the three diseases. In one study of the hindsight bias, the doctors claimed they would have predicted disease *B* was more than 50% likely to be the cause of the patient's symptoms. Clearly, the responses in both situations should be the same—33% for each disease—yet the results were not so.¹⁶⁷ The doctors all exhibited hindsight biases. The practical effect of this bias on the legal world is huge. How should we view an expert doctor's testimony at a medical malpractice suit with the effect of hindsight bias in mind?¹⁶⁸

¹⁶² J. St B.T. Evans, *Bias and Rationality*, in RATIONALITY: PSYCHOLOGICAL AND PHILOSOPHICAL PERSPECTIVES 6, 6 (K.I. Manktelow & D.E. Over eds., 1993) (noting that humans exhibit "a whole range of systematic errors and biases").

¹⁶³ See, e.g., Hal R. Arkes et al., *Eliminating the Hindsight Bias*, 73 J. APPLIED PSYCHOL. 302, 305 (1988).

¹⁶⁴ See, e.g., Lee Ross et al., *Perseverance in Self-Perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm*, 32 J. PERSONALITY & SOC. PSYCHOL. 880, 882-83 (1975) (referring to the "entity effect" as the perseverance phenomenon).

¹⁶⁵ See, e.g., SUSAN T. FISKE & SHELLEY E. TAYLOR, SOCIAL COGNITION 142-79, 245-94 (2d ed. 1991) (1984).

¹⁶⁶ See Arkes et al., *supra* note 163, at 305.

¹⁶⁷ See *id.* at 305, 306 tbl.1.

¹⁶⁸ See, e.g., Susan J. LaBine & Gary LaBine, *Determinations of Negligence and the Hindsight Bias*, 20 LAW & HUM. BEHAV. 501, 510-11 (1996).

The entity effect is another example of a cognitive bias that affects human decision-making.¹⁶⁹ Research in this field demonstrates that if humans are presented with a hypothetical situation and a set of related facts, they will continue to maintain their belief in a conclusion drawn based on the facts even after they have learned that those facts are false. Thus, if given a set of facts and then asked to create a hypothesis based on those facts, even if those facts are later proved to be entirely false, people still hold on to their initial hypotheses and fundamentally believe it is correct. This tendency is called the entity effect.¹⁷⁰

Another example of a cognitive bias that affects decision-making is the availability bias. This refers to the phenomena in which individuals, faced with a decision, give far too much credence to anecdotal evidence that is particularly salient to them, instead of depending on relevant, known statistical evidence.¹⁷¹ For example, suppose that a person lives in city *X*. Recently, two of the person's friends have had their cars stolen. The person is asked by another friend whether car theft is a big problem in city *X*. Previously the person was provided with statistical evidence demonstrating that, compared to other cities, city *X* has a very low rate of car thefts. Nonetheless, the person will probably still say, and genuinely believe, that car thefts are a big problem in city *X*, despite the evidence to the contrary.

With these examples of cognitive biases in mind, we move to those biases that affect our issue specifically. That is, we will examine which cognitive biases will be triggered by the adoption of limiting retributivism with its aggravating and mitigating factors and what the effects of those biases will be. The biases are motivated reasoning,¹⁷² optimistic bias,¹⁷³ cognitive dissonance,¹⁷⁴ and control illusion.¹⁷⁵

A. MOTIVATED REASONING

Motivated reasoning is a combination of several cognitive biases and describes the tendency in humans to reach desired conclusions in an

¹⁶⁹ See Ross et al., *supra* note 164, at 882-83.

¹⁷⁰ *Id.* at 883-84.

¹⁷¹ See Matthew Rabin, *Psychology and Economics*, 36 J. ECON. LITERATURE 11, 30 (1998).

¹⁷² See, e.g., Ziva Kunda, *The Case for Motivated Reasoning*, 108 PSYCHOL. BULL. 480 (1990).

¹⁷³ See, e.g., Neil D. Weinstein, *Unrealistic Optimism About Future Life Events*, 39 J. PERSONALITY & SOC. PSYCHOL. 806, 806 (1980).

¹⁷⁴ See, e.g., Neil D. Weinstein & William M. Klein, *Resistance of Personal Risk Perceptions to Debiasing Interventions*, 14 HEALTH PSYCHOL. 132, 138-39 (1995).

¹⁷⁵ See, e.g., Ellen J. Langer, *The Illusion of Control*, 32 J. PERSONALITY & SOC. PSYCHOL. 311, 311-13 (1975).

apparently unbiased manner.¹⁷⁶ Specifically, given a set of facts, humans are more likely to reach the conclusion that they want, without actually being aware of this predisposition. When making decisions, “one [naturally] accesses only a biased subset of the relevant beliefs and rules.”¹⁷⁷ Thus people “perceive and interpret evidence in a manner designed to confirm initial hypotheses, but [they] construct[] the initial hypotheses themselves through biased cognitive processes designed to ‘reason’ toward a desired conclusion.”¹⁷⁸ This bias has significant, real-world importance. For example, people may often convince themselves they have not had too much to drink, and that even if they are above the legal limit, they can still handle a car. These individuals honestly believe in such a statement despite the fact that the majority of people in this situation cannot, in fact, drive safely.¹⁷⁹ Unfortunately, “the practice of motivated reasoning appears to be a universal and, perhaps, immutable characteristic of human nature.”¹⁸⁰

B. OPTIMISM BIAS

Optimism bias describes how human beings are simply overly optimistic about factors affecting their lives.¹⁸¹ Despite the fact that many people purport to believe that they are unlucky or that bad things always happen to them, deep-down people are extremely optimistic.¹⁸² Regardless of how aware of statistics or actuarial calculations people are, they still often believe bad things will not happen to them while good things will. For example, people are twice as likely to think their children will be gifted than is statistically probable,¹⁸³ while they are six times more likely to think they will own a home than statistics suggest.¹⁸⁴ “One particular manifestation of this bias is the tendency of people to underestimate their own chance of suffering some adverse outcome even when they accurately state or even overstate everyone else’s chance of suffering the same outcome.”¹⁸⁵

¹⁷⁶ See Kunda, *supra* note 172, at 480.

¹⁷⁷ *Id.* at 493.

¹⁷⁸ Hanson & Kysar, *supra* note 161, at 654.

¹⁷⁹ Cf. Kunda, *supra* note 172, at 495-96 (discussing examples of the motivated reasoning bias; the example given here, while not explicitly supported by Kunda’s article, is a logical progression).

¹⁸⁰ Hanson & Kysar, *supra* note 161, at 654.

¹⁸¹ See Weinstein, *supra* note 173, at 806.

¹⁸² See Neil D. Weinstein, *Optimistic Biases About Personal Risks*, 246 SCI. 1232, 1232 (1989).

¹⁸³ See Weinstein, *supra* note 173, at 810.

¹⁸⁴ *Id.*

¹⁸⁵ Hanson & Kysar, *supra* note 161, at 656.

Optimism bias stems from several sources, chief among them the belief that risk is avoidable through individual action. In other words, the individual believes that while something bad may happen to others, he or she knows how to handle the situation and, as a result, will not suffer the same outcome.¹⁸⁶ Optimism bias is one of the most prevalent biases and, unfortunately, also one of the most "indiscriminate and indefatigable" ones.¹⁸⁷

C. COGNITIVE DISSONANCE

Cognitive dissonance is somewhat of a subset of both optimism bias and motivated reasoning.¹⁸⁸ It describes the general tendency of people to ignore or undervalue information that contradicts more positive information about oneself.¹⁸⁹ Behavioral psychologists believe cognitive dissonance stems from individuals seriously believing themselves to be "smart, nice people."¹⁹⁰ Anything that contradicts this self-image is ignored or downplayed.¹⁹¹ The effect of this cognitive feature is that "[i]n the face of a known risk . . . individuals come readily to the opinion that they themselves—unlike the average person—are relatively immune, and they hold onto these optimistic assessments tenaciously."¹⁹² Thus, "[p]eople prefer to believe that they are intelligent and are not subjecting themselves to a substantial risk."¹⁹³ The impact of this type of bias should be clear: people think more highly of themselves and their abilities than they should.

D. CONTROL ILLUSION

The illusion of control is one which leads individuals to attribute random events to their own actions; in effect, they believe they can control, to some extent, random events.¹⁹⁴ Several studies have demonstrated this bias's effects in a variety of situations.¹⁹⁵ The important point here is that

¹⁸⁶ See, e.g., Laurie J. Bauman & Karolynn Siegel, *Misperception Among Gay Men of the Risk for AIDS Associated with Their Sexual Behavior*, 17 J. APPLIED SOC. PSYCHOL. 329, 344-45 (1987).

¹⁸⁷ Hanson & Kysar, *supra* note 161, at 657.

¹⁸⁸ Cf. Weinstein & Klein, *supra* note 174, at 138-39.

¹⁸⁹ *Id.*

¹⁹⁰ George A. Akerlof & William T. Dickens, *The Economic Consequences of Cognitive Dissonance*, 72 AM. ECON. REV. 307, 308-09 (1982).

¹⁹¹ *Id.*

¹⁹² Hanson & Kysar, *supra* note 161, at 658 (footnotes omitted).

¹⁹³ *Id.*

¹⁹⁴ See Langer, *supra* note 175, at 311-13.

¹⁹⁵ See, e.g., Ellen J. Langer & Jane Roth, *Heads I Win, Tails It's Chance: The Illusion of Control as a Function of the Sequence of Outcomes in a Purely Chance Task*, 32 J.

people believe that their control of a situation is merited by the actual situation.

E. LIMITING RETRIBUTIVISM

With these biases in mind, we now look to limiting retributivism doctrines and their use of aggravating and mitigating factors to determine the appropriate range of punishment. The salient issue here is the theory's open and public support of the use of these aggravating and mitigating factors. The entire concept of a defendant's moral culpability rests on determining how immoral or blameworthy the defendant and his or her actions are. Unfortunately, when this theory openly advocates the use of these aggravating and mitigating factors, it links directly into cognitive biases, where it becomes problematic.

If the potential criminal is aware of the use of aggravating and mitigating factors in determining sentencing ranges, his or her decision-making will be affected by the interaction of his or her cognitive biases and his or her knowledge of the factors in the present situation that may be considered in sentencing. Before a person's decision to commit or not to commit a crime, all four of the aforementioned biases will play their respective roles in his or her judgment process. Although these biases will be present whether or not there are aggravating and mitigating factors, the fact that these factors may play a role in any punishment received will be shown to exacerbate the effects of these biases. How will these cognitive biases affect a criminal's decision-making process, given that the criminal knows that aggravating and mitigating factors may play a role in sentencing?

Because of cognitive biases, a person debating whether or not to commit a crime may be comforted by several facts that the criminal believes to be true. First, the criminal probably does not believe that he or she will be caught in the first place. Second, if he or she is caught, he or she will believe that he will receive a sentence or punishment on the lower end of the range of permissible punishments. As a result of his or her cognitive biases, the criminal is likely to genuinely believe that the jury or the judge will find, given the particular facts of the case, that he or she does not have a high degree of moral culpability. For example, if the person has a history of abuse as a child, or has a poor family that needs to be fed, or if the crime contemplated is non-violent in nature, a person may believe that these or other factors will lead the jury or the judge to find that the

defendant's just desert is on the low end of the range of permissible punishments.

Given the existence of motivated reasoning, we can see that all of the above statements contribute to the hypothesis about potential punishment for the crime contemplated that a person would likely create, given the set of facts as he or she believes them to be, when engaging in the decision-making process of whether to commit a crime. The crux is that the criminal's hypotheses or predictions about any aggravating or mitigating factors that will affect his or her punishment are probably not accurate, but instead are highly influenced by cognitive biases. In other words, because of inherent cognitive biases, a person will conclude that, if caught and found guilty, he or she will not receive a harsh sentence even though a more balanced and rational interpretation of the facts would actually lead to a different conclusion. As mentioned above, the potential criminal's thought process is an example of motivated reasoning. The jury or judge will likely have a more balanced and rational interpretation of the facts and circumstances surrounding the crime.

The same person will also exhibit the optimism bias. If his or her cognitive processes are being filtered through the optimism bias, a person will believe that the judge will not be harsh on him because he or she honestly thinks he or she, and his or her situation, is different. The person will think the judge will be far more likely to find mitigating factors and far less likely to find aggravating factors. As a result, the individual will believe that the judge will determine that the appropriate punishment is on the very low end of the scale.

In addition, the person may suffer from cognitive dissonance in the decision-making process. This person may anticipate a lenient judge, for example, because he or she believes that he or she is a nice person. The judge will see that he or she is a good person who may have simply fallen on some bad luck, as he himself believes. Alternatively, he or she may think he or she is smart enough to outwit the system or the judge. In the mind of the would-be criminal, all of the potential aggravating factors will not be important to the judge, who will see through all of the district attorney's ploys to cast the defendant in a bad light and see him or her for who he or she really is—a good person who made a mistake and had lots of forces out of his or her control pushing him or her towards committing the crime. The criminal believes that he or she is either not very morally culpable, or can convince the judge that he or she is not morally culpable. Therefore, the criminal believes that the judge will grant him or her a range of sentences on the lower end of the spectrum. Of course, the criminal's confidence in his or her good nature or ability is likely unfounded.

Lastly, the would-be criminal may believe that he or she has a way to convince the jury or the judge to be lenient on him. He or she may believe that he or she has a secret method to put the judge or the jury in a good mood or to make them sympathetic to his or her pleas. While there is a chance the defendant may actually have a method to achieve these goals, it is more likely he or she is simply demonstrating control illusion and, in fact, will be unable to control the judge or jury.

While the examples given above are not examples of purely one type of cognitive bias or another, they combine to demonstrate the collective effects of these biases on a criminal's psychology and, thus, how the criminal's decision-making process may be affected. Together, these illustrations demonstrate how a person may anticipate a sentence far lower than he or she rationally should. The open use of aggravating and mitigating factors in determining sentences plays strongly into the effects of these cognitive biases. By openly admitting that they are used in determining sentences, limiting retributivism is, in essence, advertising directly to these biases. Reliance on them in determining sentencing simply adds more factors into which cognitive biases can play, with the effect of altering the criminal's expectations.

This may have a disastrous effect if one abides by a utilitarian theory of punishment. Under limiting retributivism, the use of these aggravating and mitigating factors in determining punishment will lead to would-be criminals anticipating punishments extremely below those which they may actually receive. Here, another law and economics example is useful. Suppose the existence of aggravating and mitigating factors leads criminals to believe that the average sentence will be one year, while the average sentence is in fact five years. The deterrence effect of the punishment, then, is five times lower than it should be. In essence, the very existence of these factors in determining punishment may lead to a person's anticipation of less prison time and hence under-deterrence.

Utilitarianism cannot be reconciled with a system that leads to such under-deterrence. Regardless of whether within the permissible punishment ranges utilitarian principles are utilized, the primary basis on which the range of punishment is decided is contrary to the goals of utilitarianism. Such a retributivist appeal to utilitarianism is similar to a store attempting to appeal to penny-wise shoppers by increasing the price of their goods by five dollars and then giving two-dollar coupons. When the system of determining punishment leads to under-deterrence, the subsequent use of utilitarian concerns is of lesser to no value.

VI. ACHIEVING UTILITARIAN GOALS

Limiting retributivism's use of aggravating and mitigating factors in determining punishment will likely result in strong under-deterrence of crime. Insights into human behavioral psychology demonstrate this fact.¹⁹⁶ Therefore, limiting retributivism proponents' attempts to appeal to supporters of utilitarian theories of criminal punishment should be rebuffed as contrary to utilitarian aims. However, establishing that limiting retributivism does not sufficiently serve utilitarian goals does not end the debate. A careful reader will notice that the cognitive biases identified by behavioral psychologists, and used here to suggest the failure of limiting a retributivist system of punishment to achieve utilitarian goals, may also be used to challenge utilitarian systems of punishment as well. After all, the same cognitive biases that lead criminals to irrationally believe they will either not be caught or will be found innocent may also affect deterrent levels in a purely utilitarian system. In this section, I defend the argument that utilitarians should reject limiting retributivism appeals to utilitarian goals because cognitive biases create a double under-deterrent effect under a limiting retributivism system while they create only a single under-deterrent effect under a utilitarian system. Furthermore, I highlight possible methods to adjust utilitarian policies to accommodate previously mentioned research in behavioral psychology and to better achieve utilitarian goals.

Given the research in behavioral psychology, utilitarians face a difficult task in arguing in further support for utilitarian theories of punishment, specifically deterrence-based punishment. Because behavioral psychologists have shown that people do not rationally react or respond to facts, it could be argued that all of the deterrence efforts of utilitarianism will ultimately fail. Critics of utilitarianism will argue that this research proves that utilitarian goals of deterrence are unreachable and therefore utilitarianism is an unworkable system of punishment. Yet, these arguments that utilitarianism is an unworkable system fail under scrutiny.

The evidence that people do not act as rational individuals, the basis of much of utilitarianism and of law and economics, weakens the allure of utilitarianism; yet it does not destroy the theory altogether. Optimism bias, the illusion of control, motivated reasoning, and cognitive dissonance all serve to decrease the overall deterrent effect of criminal punishment; however, they do not destroy all deterrent effect. Whether because criminals think they have less of a chance of getting caught or judges will be lenient with them if caught, the overall expected harm that a criminal will anticipate will be lower than it should be. The reason behind this sub-optimal deterrent level is the effect of cognitive biases.

¹⁹⁶ See *supra* Part IV.

But does this mean that punishment has no deterrent effect? The answer is *no*. While the deterrent effect is not optimal, it still exists, albeit in a weaker form. At this point, one may wonder, is this not the same argument used to discredit retributivism's appeal to utilitarianism? Was not the fact that retributivism led to under-deterrence the very reason it was held to be contrary to utilitarian goals? The answer is *yes*, but only to a degree. The real question then becomes which system creates a stronger deterrent effect?

To answer this question, one needs to determine which system is more affected by these cognitive biases: limiting retributivism or utilitarianism. In a limiting retributivism system, two factors exist that will lead to under-deterrence. First, the use of basic utilitarian principles to alter a sentence within the acceptable range will not have as much deterrent effect as one would hope. Second, the open use of mitigating and aggravating factors in determining punishment will create further under-deterrence.¹⁹⁷ Both steps in the determination of punishment under limiting retributivism, creating the range of permissible punishment and altering the specific punishment within that range, deter at a sub-optimal level. Utilitarianism does not face this double under-deterrence effect. Under a utilitarian system, the general deterrent effect of the punishment will be less than ideal, but no other factors will further reduce the deterrent effect. In other words, there is likely only one point where cognitive biases create sub-optimal deterrence levels.

An example will demonstrate this difference appropriately. The crime at issue is auto theft. Suppose it is an ideal world, where people act according to the classic economic view of human beings as rational and the optimal level of deterrence is created by a punishment of twenty years. In other words, this punishment provides the ideal deterrent effect. In a utilitarian system, one found guilty of this crime will face a twenty-year sentence if caught and convicted. Now, let us assume that human nature, as defined by behavioral psychology, will result in a 20% decrease in the deterrent effect, either because criminals expect not to get caught, or because they expect leniency in sentencing, or for whatever other reason. Whereas under the previous scenario of completely rational human behavior, a stated punishment of twenty years in jail has an expected punishment of twenty years, now there is an expected punishment of sixteen years. Thus, with cognitive biases in mind, a twenty-year sentence no longer creates the optimal deterrence level because criminals will expect a sixteen-year sentence. The utilitarian system has two options: (1) increase the punishment by 25% to twenty-five years, so the aggregate effect is

¹⁹⁷ See *supra* Part III.

twenty years of expected punishment, resulting in the optimal deterrence level or (2) accept less-than-ideal deterrent levels—an 80% effective deterrent effect.

Now analyze this situation under a limiting retributivism system of punishment. Under a limiting retributivism system, assume that the range of punishment for someone convicted of stealing a car is fifteen to twenty-five years. Because of the open use of aggravating and mitigating factors in determining sentencing, the potential car thief will expect a 20% decrease in the applied range. Now, he or she expects a range of twelve to twenty years instead of fifteen to twenty-five. The anticipation of a decreased range of punishment, which as mentioned above stems from cognitive biases, is the first way in which limiting retributivism under-deters.

Next, the criminal will further expect his or her sentence to be 20% less than is probable within the range. This occurs because the criminal will anticipate this second level of decision-making by the judge—he or she will know that the judge has discretion to sentence within a given range. At this point, cognitive biases may lead the criminal to expect or anticipate a specific level of punishment within the range that is less than what is actually likely, given the facts of the case. This is where the second level of under-deterrence occurs in limiting retributivism. Within the eight-year range, the facts of the case may indicate four years of added punishment, yet the criminal's prediction of this number is affected by his or her biases, and as a result he or she will predict 3.2 years of added punishment. Thus the criminal's expected punishment will be 15.2 years (12 years plus 3.2 years).

Thus, the aggregate deterrence level will be closer to the optimal level under a utilitarian system. Under utilitarianism the potential criminal will expect a 16 year sentence, while under limiting retributivism he or she will expect a 15.2 year sentence. Because deterrence levels are higher under a utilitarian system of punishment than under a limiting retributivism system, adherents of utilitarian theory must reject limiting retributivism. Using a utilitarian system will lead to greater deterrence and decrease the likelihood of future crimes. Accordingly, utilitarianism provides the greatest level of deterrent effects. Thus, for those who hope the law can have a deterrent effect by creating *ex ante* incentives for individuals not to commit crimes, utilitarianism is the only system and theory of punishment that is acceptable. If one is steadfast in the belief of the value of using *ex ante* incentives, the only system permissible is that which comes closest to creating optimal levels, which is utilitarianism.

Yet this does not mean that utilitarians have won and should pack up and go home happy. As has been demonstrated, utilitarian systems of punishment may still not result in optimal levels of deterrence. Human

psychology undoubtedly decreases the deterrent effect of criminal laws. Anyone who believes the law can and should create ideal levels of incentives to create optimal outcomes, including most proponents of utilitarian theories of punishment, has much more work to do.

As the recent developments in the fields of behavioral law and economics indicate, the goal of optimal outcomes may still be reached if we tailor laws in recognition of cognitive biases or create laws—or, more specifically, systems of punishment—that minimize the effects of these biases.¹⁹⁸ In fact, unbeknownst to lawyers at the time of the rules' creation, several legal rules can be seen as methods to overcome cognitive biases.¹⁹⁹ An excellent example of such a law is the “business judgment rule” in corporate law. This rule essentially states that when challenging a business decision of a company executive, all that is required to defend the decision—absent any evidence of self-dealing—is that it was in some way rational, even if not the best decision.²⁰⁰ In essence, this rule can be seen as compensating for the existence of hindsight bias by declaring that the outcome is not important at all, and saying that if there was any reasonable rationale for making the decision that was ultimately made, then the decision is legitimate.²⁰¹

While the example of the business judgment rule demonstrates how behavioral psychology insights can help craft better laws, or justify previously existing ones, these laws are examples of *ex post* reasoning. What concerns utilitarians are *ex ante* incentives. The cognitive biases that affect *ex ante* incentives are very different than those that affect *ex post* decisions. In other words, the hindsight effect, highlighted by the business judgment rule example, focuses more on a person's view of a situation after that situation has occurred, and thus is unlikely to have effects on *ex ante* decision-making and hence *ex ante* incentives. Unfortunately, biases that affect such decision-making, for example, optimism bias and motivated reasoning, have been demonstrated to be extremely difficult to adjust for or to alter.²⁰²

Utilitarians face a tough challenge, but one that may not be insurmountable. Utilitarianism has been demonstrated to have at least some effect on deterring criminals. Further investigations may prove that cognitive biases are, in fact, correctable. If not, other, more correctable biases may be further investigated and applied. Insight into such biases

¹⁹⁸ See Jolls, Sunstein & Thaler, *supra* note 159, at 1522-41.

¹⁹⁹ *Id.*

²⁰⁰ Smith v. Van Gorkom, 488 A.2d 858, 872 (Del. 1985).

²⁰¹ Jolls, Sunstein & Thaler, *supra* note 159.

²⁰² Hanson & Kysar, *supra* note 161, at 657.

may not lead to a system of perfect deterrence, but it may assist utilitarians in crafting laws that are better able to deter criminals. Through the research, use, and adoption of laws tailored to accommodate for, or to alter, cognitive biases, deterrence levels may be increased.

VII. CONCLUSION

Recent developments in retributivist theory have resulted in the creation of a limiting retributivism system of criminal punishment. Such a system attempts to appeal to utilitarian concerns in an effort to garner support from those who desire the criminal law to serve utilitarian purposes. Utilitarians must reject this appeal, however, because it leads to lesser deterrent levels compared to those expected under a properly utilitarian system of punishment.

Recent developments in behavioral psychology suggest reasons why the limited retributivist system of punishment may lead to sub-optimal deterrence levels. Limiting retributivist systems of punishment are affected by cognitive biases on two separate occasions, leading to a double under-deterrence effect, while utilitarian systems of punishment only face a single under-deterrence effect. Simply put, cognitive biases suggest that limiting retributivism systems of punishment will lead to higher levels of under-deterrence than will utilitarian systems of punishment. In light of these inferences, utilitarians must resist the appeals of proponents of limiting retributivism and insist on a more rigid utilitarian system of punishment.

Utilitarians must engage in further research in an effort to mitigate the effects of cognitive biases on deterrence levels. The *ex ante* incentives that utilitarians wish to create through the law may be improved if laws are drafted to account for—and overcome—demonstrated cognitive biases.

