“The Relevance of Prior Record in the Criminal Law: A Response to the Theory of Professor von Hirsch”

Darcy L MacPherson, *University of Manitoba*
The Relevance of Prior Record in the Criminal Law: A Response to the Theory of Professor Von Hirsch

Darcy L. MacPherson *

Andrew von Hirsch believes that criminal punishment is based on desert. The author agrees with von Hirsch that an offender's blameworthiness is due to a combination of the harm caused by the offence and the culpability of the offender, and that the punishment must be proportional to these two measures. The author then explains two theories of culpability, choice theory and character theory, to which he later returns in explaining his own theory.

The paper continues with an examination of von Hirsch's evolving views on the relationship between prior record and proportional punishment. Von Hirsch maintains that society gives new offenders a "discount," in recognition of the fact that human frailty may result in lapse of judgment. When someone re-offends a sufficient number of times, he or she loses the "discount" (mitigation of punishment). For von Hirsch, a lack of prior record mitigates deserved punishment, but prior record cannot increase punishment.

The author argues that recidivism can increase culpability, and that punishment can thus increase for prior record while still being proportional. Recidivism should affect punishment in small, incremental steps, to reflect the offender's additionally culpable choice to act in the present without regard for society's past censure. This approach respects moral agency as being central to desert theory and accepts the basic assumption that people can choose to act as rational, moral agents. The offender's decision not to react to society's message of blame as a moral agent, by again committing the same crime, is a culpable choice in itself. The author maintains that this model fits with desert theory rather than defiance theory and is thus an extension of Professor von

* Assistant Professor, Faculty of Law, University of Windsor (2000-2002); University of Manitoba (2002-current). I owe thanks to Professors Richard Bronaugh, William Conklin, Brian Etherington, Allan Manson, Andrew Simister and Lorna Turnbull for their insightful comments. I would also like to thank all of the anonymous reviewers who gave liberally of their time and expertise in reviewing this article. Generous support for this article was received from the Law Foundation of Ontario and the Legal Research Institute of the University of Manitoba. Finally, thanks are also owed to my research assistants, Barbara Johnson, Jane Magree, Glen Perinot, Patty Robinet and Lyla Simon.
Hirsch’s work. The article concludes with a suggestion for the practical application of his theory to the sentencing process: a plateau-based system for increasing punishment.

Introduction
I. Von Hirsch’s Theory
   A. Harm
   B. Culpability
      (i) Choice Theory
      (ii) Character Theory
   C. Proportionality
II. Von Hirsch’s View of Prior Record
III. Prior Record and Culpability
IV. Developing a New Theory
Conclusion

Introduction

Professor Andrew von Hirsch has been acknowledged as one of the leading proponents of retributivist\(^1\) theory in criminal legal philosophy.\(^2\) For von Hirsch, punishment has an embedded moral

---

1. Retributivist theories of justice are focused on the act in question (in our discussion, the act said to be criminal) and the actor who performed the act. The consequences or effects on other people of punishing the actor are irrelevant (or, at the very least, reduced to secondary importance). Von Hirsch himself prefers the term “desert” in describing this type of theory. See Andrew von Hirsch, Doing Justice: The Choice of Punishments (Boston: Northeastern University Press, 1976) at 45-46 [Von Hirsch, Doing Justice]. This paper adopts the von Hirsch preference.


178 (2002) 28 Queen’s L.J.
message of wrongful conduct by the offender. Society wants to make this message a public one, by the public nature of the conviction and sentence. Von Hirsch’s theory argues that the amount of punishment that an offender should receive is a function of both the harm caused by the offence and the culpability of the offender.

But what ought to be the relevance of an offender’s prior record when deciding the quantum of punishment? Von Hirsch argues that prior record is relevant to sentencing, in that subsequent offences lead to a “progressive loss of mitigation.” The result of this approach is that after a third or fourth offence, prior record has no relevance to sentencing.

This paper agrees with von Hirsch that prior record should not bring an offender’s punishment outside what would otherwise be deserved for the offence committed by the particular offender. However, I also argue that multiple repetitions of the same offence by the same offender should increase his or her culpability over time. Therefore, a multiple-recidivist deserves more punishment than a first-time offender. This paper adopts von Hirsch’s theory of punishment. Part II is concerned with setting out some key elements of his theory, including the concepts of harm, culpability and proportionality. Part III considers von Hirsch’s writings on the relevance of the prior record in sentencing. In Part IV, I return to the concept of culpability and choice theory, which holds people accountable (or culpable) for bad choices. I will argue that while all offenders make bad choices, repeat offenders make an additional choice that increases the

4. Ibid. at 29.
5. According to von Hirsch, the sentence of a first offender is “mitigated” somewhat on account of his or her status as a first offender. This “mitigation” is lost by subsequent wrongdoing; see infra note 6 at 192.

D.L. MacPherson 179
degree to which they should be held accountable for their actions, and that increased punishment may be therefore justified. Finally, in Part V, I will sketch my own theory, arguing that prior record should lead to successive modest increases to punishment.

I. Von Hirsch’s Theory

To begin, what is punishment? For von Hirsch, punishment is composed of three elements: (a) the imposition of hard treatment; (b) by an authority with the right to do so; and (c) with an embedded message of moral wrongdoing or blame.7 The third element is what von Hirsch refers to as “censure.”8 Censure is society’s attempt to communicate to the offender that he or she was wrong9 and has done something blameworthy.

In von Hirsch’s theory, the concept of desert is based on the blameworthiness of the offender’s conduct. For von Hirsch, blameworthiness in terms of state criminal punishment10 involves two distinct concepts: (a) the harm caused by the offence

---

7. This definition conforms to a desert view of punishment. For definitions of punishment from alternative perspectives, see H.L.A. Hart, “Prolegomenon to the Principles of Punishment” in Punishment and Responsibility: Essays in the Philosophy of Law (New York: Oxford University Press, 1968) 1 at 4-5, and Jeremy Bentham, An Introduction to the Principles of Morals and Legislation (New York: Hafner Publishing Company, 1948) at 168-169. As should be clear from the discussion thus far, however, this paper is concerned with desert theory.
9. Ibid. at 10.
10. To be clear, the focus of this article is criminal punishment imposed by the state. Nonetheless, some of the examples used herein may be extra-legal punishment. The term “extra-legal” is used to describe those punishments which: (a) do not have a basis in the law; and (b) are accepted as not needing any such basis. The most obvious examples of “extra-legal” punishments, as I use the term, are punishments meted out by parents to their children and perhaps punishments meted out by teachers to their students. I choose the term “extra-legal” (as opposed to “non-legal”) punishment because it avoids a connotation of running contrary to the law.
committed, and (b) the culpability of the offender. Expressed as a mathematical equation, one might say that the quantum of punishment received by an offender should be proportional to the harm caused by the offence multiplied by the culpability of the offender. It is necessary to look more closely at the foregoing statement, and in particular, to look at the theories used by von Hirsch to attempt to quantify and explain both harm and culpability.

While this mathematical statement may appear simple and straightforward, it can become complex. First, not all harmful conduct has culpability attached to it. For example, if I break a person’s nose because of an involuntary spasm of my arm, there is harm to that person, but I am not culpable, because it was involuntary. Therefore, notwithstanding that harm was done, no criminal liability can attach because there was no culpability.

Correspondingly, culpability without harm will not suffice for criminal liability. A desire to harm someone, without actually doing anything to bring harm about, is not enough to attach criminal sanctions. For example, assume Person A wishes that

11. Harm to others is the most common result of criminal wrongdoing. For example, assault causes physical harm, theft causes harm to well-being through loss of property, and so on. However, there are prohibitions that are unrelated to harm to others. Certain criminal wrongs (like the failure to wear one’s seat belt) are about harm to the actor him- or herself. For an in-depth discussion of these wrongs, see Joel Feinberg, *Harm to Self*, vol. 3 (Oxford: Oxford University Press, 1986) and John Kleinig, *Paternalism* (Manchester: Manchester University Press, 1983). Similarly, other prohibitions (like a prohibition against nudity in public) combat conduct that is considered “offensive.” For a discussion of the underpinnings of these prohibitions, see Joel Feinberg, *Offence to Others*, vol. 2 (Oxford: Oxford University Press, 1985), and Andrew von Hirsch, “The Offence Principle in Criminal Law: Affront to Sensibility or Wrongdoing?” (2000) 11 King’s College L.J. 78. For the purposes of this paper, it is sufficient to recognize that these other categories exist, and that harm to others is therefore not strictly necessary to justify a criminal prohibition.


13. This skips over the law of attempts, which does not require actual harm to result as long as the attempt is completed. Nonetheless, we are concerned here

D.L. MacPherson 181
Person B were dead. Person A’s desire (however temporary) for the death of Person B may be sufficient to ground culpability. Yet, the culpable thought is not enough to ground criminal liability (or the punishment attached to it), because no harm is done to Person B. Even if the criminal justice system could tell that a person wished to harm someone else, this is generally not enough, by itself, to attract the attention of the criminal law.\textsuperscript{14}

Yet, this statement does not apply to every instance where there is both harm and culpability. There must also be an \textit{offence}. The criminal law does not prohibit every instance where there is both harm done and culpability for the harm. To use a simple example, the criminal law generally does not prohibit negligent assault.\textsuperscript{15} Assume that Person A inadvertently makes contact with Person B, injuring her. This is not sufficient for criminal sanctions to attach, even though there may be some harm to Person B, and some culpability in Person A, because negligent assault is not prohibited by the criminal law.\textsuperscript{16}

I have thus far considered when the mathematical expression of von Hirsch’s theory does \textit{not} apply. But that is only a negative exercise and does not explain how society should assess what the terms “harm” and “culpability” really mean. These are the things with paradigm criminal harms. Therefore, the issue of attempts is best left to another day.

\textsuperscript{14} This is why the mathematical equation is stated as “harm multiplied by culpability” and not “harm plus culpability.” If the equation were stated as the latter, then punishment could be justified in the absence of one of the elements. To use a numeric example, if the harm were a “10” and the culpability were a “0”, then on the “plus” conception, there is a positive numeric value for punishment of “10”. However, on the “multiplied by” conception, no punishment can result in the absence of either harm or culpability.

\textsuperscript{15} Negligence is, in some instances, a sufficient showing of culpability to warrant the attachment of criminal sanctions. See, for example, the offence of failing to provide the necessaries of life under s. 215 of the \textit{Criminal Code}, R.S.C. 1985, c. C-46, s. 215.

\textsuperscript{16} In Canada, intention or recklessness is generally necessary to ground a conviction for assault. With respect to intentional assault, see s. 265 of the \textit{Criminal Code}, \textit{ibid}. With respect to the availability of reckless assault, see also \textit{R. v. Emans} (2000), 35 C.R. (5th) 386 at 393-394 (Ont. C.A.).
to which punishment must be proportional and once understood, we will be much closer to understanding von Hirsch’s theory of proportionality.

A. Harm

How should society judge harm? Von Hirsch and Nils Jareborg have developed a harms analysis that is based on a generalized concept of the effect of the harm done on the “living-standard” of the objective victim.17 Von Hirsch and Jareborg interpret the victim’s living standard to include at least the following four types of protected areas: (a) physical integrity; (b) material support and amenities; (c) freedom from humiliation; and (d) autonomy and privacy.18 Violations in those areas affect four levels of comfort: (a) subsistence; (b) minimal well-being; (c) adequate well-being; and (d) enhanced well-being.19 Attached to each of these four levels of comfort is a general idea of the harm done when it is invaded. If the violation strikes to the subsistence level, then the harm is grave; if it attacks minimal well-being, then the harm is serious, and so on.20

17. See Andrew von Hirsch and Nils Jareborg, “Gauging Criminal Harm: A Living-Standard Analysis” (1991) 11 Oxford J. Leg. Stud. 1. There are other ways to judge and quantify harm as well. See Joel Feinberg, Harm to Others (Oxford: Oxford University Press, 1984) [Feinberg, Harm to Others]. For Feinberg, harm is to be judged in accordance with the “interests” of the individual harmed. Once an “interest” is at stake, the harm is then quantified by reference to the degree to which the choices (or liberty) of the victim are restricted as a result of the violation of the interest in question. See Feinburg, Harm to Others at 35, 206-212.

It should be noted as well that even though von Hirsch now uses the living-standard approach to harm, in his earlier work he had used Feinberg’s approach for determining the seriousness of harm. See Andrew von Hirsch, Past or Future Crimes: Deservedness and Dangerousness in the Sentencing of Criminals (Manchester: Manchester University Press, 1985) at 66-74 [Von Hirsch, Past or Future Crimes].
19. Ibid. at 17.
20. Ibid. at 29.
While other approaches to the quantification of harm are at least partially centred on the subjective loss of choice of the victim, the living-standard approach is wholly objective. It examines “standard” harms and is solely concerned with the harm that the “reasonable” victim would experience from the particular violation.

B. Culpability

Culpability is the state of mind that the actor must possess in order for criminal liability to attach. The American Model Penal Code sets up a fourfold typology of states of mind that can qualify as culpable: (a) purposeful wrongdoing; (b) knowing wrongdoing; (c) reckless wrongdoing; and (d) negligent wrongdoing.

But the more interesting question is: what underpins the attribution of culpability to a particular person? To respond, I will examine two different theories in the philosophy of culpability. The first is based on the morally defective choices of the actor; the second is based on character that is morally deficient. With respect to the first, I will examine the theory of Michael Moore. The work of Michael D. Bayles will provide an example of the second.

21. See Feinberg, Harm to Others, supra note 17 at 192-193.
22. See von Hirsch & Jareborg, supra note 17 at 5.
25. This discussion will focus on the polar opposites of this debate. Nonetheless, there are theories that attempt to bridge the gulf between these extremes. For example, see Peter Arenella, “Character, Choice and Moral Agency: The Relevance of Character to Our Moral Culpability Judgements” (1990) 7:2 Social Philosophy & Policy 59.
(i) Choice Theory

Moore's theory centres on the actor's decision to do something that is morally reprehensible, or not to do what is morally required. In other words, the actor's choices are the focus of the discussion of culpability. If the actor chooses to do something that is immoral (that is, something harmful to others), society may punish him for that choice, if the choice is blameworthy and results in harm.

The first benefit of this theory is that it accords with a layperson's understanding of the way blame works in everyday life. If someone says "Person A is to blame for the fact that my car was driven over a cliff," the response might be "Person A was not driving: Person B was. Therefore, Person B is to blame." In such circumstances, one would expect that reasonable people would not hold Person A to blame for the car going over the cliff.

But why is this so? The easy answer is the absence of choice in the scenario presented. Generally, the driver is expected to exercise decision-making power over the direction and speed of an automobile. Absent special circumstances, the driver makes the choices about how the car should be driven. Therefore, when harm (the loss of the car) results from these choices, liability may attach. If Person B (the driver) is responsible for the car going over the cliff, then the separateness of persons means that Person A cannot and (more importantly for our purposes) should not be

---

26. See Michael S. Moore, "Choice, Character, and Excuse" (1990) 7:2 Social Philosophy & Policy 29. There are many different variants of theories of culpability that involve some element of choice. See H.L.A. Hart "Negligence, Mens Rea and Criminal Responsibility" in Hart, supra note 7, 136 at 150-151 (which focuses on the actor's capacity to make a choice rather than the choice actually made by the actor). See also Jean Hampton, "Mens Rea" (1990) 7:2 Social Philosophy & Policy 1 (which focuses on the actor's choice to defy the dictates of the law). I concentrate on Moore's contribution to the debate, however, because even those who disagree with him have been known to complement the completeness and subtlety of his argument. See Arenella, ibid. at 63, n.10.

27. Moore, ibid. at 35.
held liable for the resulting harm. So a choice conception of culpability accords with an everyday idea of when people should blame others for harmful behaviour.

The second advantage of a choice theory of culpability is that it promotes respect for the institution of the law and thereby promotes respect for the rule of law. The certainty principle requires that a person have the opportunity to know what will subject him or her to the sanction of the criminal law. In brief, choice theory respects the certainty principle because if a person chooses to do nothing, he or she largely avoids criminal sanction.

If the law were to adopt a theory of culpability that did not involve any choice, then the people governed by that law would not be able to order their lives so as to avoid criminal liability. If people can order their lives to avoid sanction, they are more likely to respect the judgments of the criminal law and adopt them as their own. Let us assume that we lived in a system

28. We are skipping over the area of accomplice liability. Most sophisticated legal systems provide for liability for persons who assist or encourage others in the commission of crimes. See, for example, s. 21 of the Criminal Code, supra note 15. The special issues presented by s. 21 (and its equivalents in other advanced legal systems) are better left to another day.

29. For a brief comment about the importance of correspondence between normative philosophical discussion and everyday thought and practice, see infra note 78.

30. This can be considered an offshoot of the “principle of legality,” and in its Latin nomenclature, “nullum crimen sine lege, nulla poena sine lege” (no crime without law; no punishment without law). See Canada, Law Reform Commission, Criminal Law — The General Part: Liability and Defences, Working Paper 29 (Ottawa: Queen’s Printer, 1982) at 12.


32. This is one of the criticisms traditionally levelled at utilitarian analysis by desert theorists. Assume that desert side-constraints are added into utilitarian
where culpability did not involve choice. When an offender did something for which he or she was punished, the offender could legitimately say, “I did nothing wrong. Therefore, I do not accept the law’s right to punish me. The law is not worthy of respect.” I will return later to discuss the importance of respect for the criminal law when I discuss defiance theory.

Von Hirsch’s writings show an obvious commitment to the choice theory of culpability. In *Censure and Sanctions*, von Hirsch begins by assuming that the criminal law is designed to force a person to examine and regulate his or her conduct according to its norms, and that all persons can be tempted to commit crimes. One can only regulate one’s conduct through choices, by choosing to act in one way as opposed to another. Furthermore, von Hirsch is committed to the idea of people as responsible moral agents, capable of being persuaded by moral appeals. This is evidence that a person’s choices and subsequent responsibility for them are key elements of von Hirsch’s theory.

(ii) Character Theory

A counterpoint to Moore’s choice theory is character theory; one striking example is the work of Michael D. Bayles, particularly his article “Character, Purpose and Criminal Responsibility.” Bayles argues that society should punish people for actions that are symptomatic or indicative of bad character. In other words, society should punish people for who they are, and not for what they do.

---

analysis (for an example of a theory which explicitly does this, see Hart, supra note 7). Proponents of desert would argue that once people figure out that what they do is not the primary determinant of (and reason for) the sanction against them, they are less likely to accept the punishment meted out by the criminal law as being legitimately imposed. If this occurs, the efficiency of the system decreases. See, for example, von Hirsch, *Censure and Sanctions*, supra note 3 at 41-42.

35. (1982) 1 Law & Phil. 5.

D.L. MacPherson
Bad character, demonstrated through actions or omissions, forms the basis for a judgement of culpability. The bad character trait shown through purposeful (or intentional) wrongdoing is the desire to harm others. For knowing conduct, the character trait is disregarding harm to others that the actor knows will occur (although the actor does not intend it). Bad character also shows itself in reckless actions because of the actor’s disregard for a substantial risk of harm, and in negligence through the actor’s carelessness.\textsuperscript{36}

Therefore, one of the first advantages of character theory is that it conforms to the four states of culpability put forward by the American Model Penal Code. A related strength is that if the theory is sound, it could explain criminal liability for negligence, which choice theory cannot.\textsuperscript{37} While Bayles’ theory allows for liability for negligence, “accidents” are excluded.\textsuperscript{38}

Another advantage of character theory is that it can be used to justify extremely long sentences for violent offenders. After all, if people are punished for who they are rather than what they have done, it becomes much easier to justify long sentences for offenders (such as psychopaths) whose character traits indicate a

\textsuperscript{36} \textit{Ibid.} at 8-10. \textsuperscript{37} In any act-based theory, liability for negligence will present its own theoretical problems, including those common to omissions generally. Negligence is, after all, an omission to satisfy the requisite standard of care in a given situation. See supra note 31 and accompanying text. Therefore, there is no specific act \textit{per se} to which one can point as being culpable. Rather, it is the absence of a particular act that results in liability.

Bayles admits that “the basis for blame [in negligence] is much less sure [than that for recklessness]”; see supra note 35 at 11. However, Bayles maintains that the inclusion of negligence in the analysis strengthens his conclusions with respect to the importance of character to culpability. (See \textit{Ibid.} at 10).

\textsuperscript{38} \textit{Ibid.} at 7. Despite providing this distinction, the essay provides no meaningful analysis as to what it includes. I must therefore assume that the meaning of this distinction is to be provided by the substantive law. Even in the substantive law, however, such a distinction can only be drawn in the vaguest of terms.
high probability of recidivism. While character theory would allow for people with psychopathic tendencies to be punished for their actions, choice theory excludes such people from criminal liability because of their inability to make choices. Finally, character theory avoids the need for a voluntariness requirement: all that is required is an action which demonstrates the blameworthy character trait. In determining punishment, character theory looks for the continuing existence of the character trait.

Now that I have presented the concepts of harm and culpability, I will discuss von Hirsch’s theory of proportional punishments.

C. Proportionality

Proportionality, for von Hirsch, is purely backward-looking. In other words, punishment cannot be determined by looking forward at the offender’s future prospects. At least in von Hirsch’s model, proportionality requires that a sentence be focused on the harm from the offence and the culpability of the

39. This is the basis of “selective incapacitation” theory. See James Q. Wilson, “Selective Incapacitation” in von Hirsch & Ashworth, supra note 6, 113 at 115-116.
40. Bayles, supra note 35 at 17.
41. This is because they objectively lack the “fair opportunity to do otherwise” required by Moore’s theory; see Moore, supra note 26 at 52.
42. Bayles, Supra note 35 at 7.
43. Moore argues that Bayles is making an interpretive analysis of how well character theory fits with existing American law on excuses (see Moore, supra note 26 at 49). Bayles views character theory (which he calls “Humean theory”) as a plausible one. See Bayles, supra note 35 at 5. Moore attributes character theory to David Hume (see Moore, supra note 26 at 40, n. 38). Therefore, Bayles clearly subscribes to the normative basis of character theory, although he does use compatibility with the current law as a justification for the application of the theory. In other words, in my view, Moore confuses Bayles’ thesis (the defence of Humean theory) with the means by which Bayles attempts to defend that thesis (through its compatibility with the then-current American law).
offender at the time of the offence.\footnote{In other words, von Hirsch’s approach does not fit with character theory. In fact, as referred to above, character theory can be used to justify sentences that would be disproportionate to the harm caused by the offence. See Wilson, \textit{supra} note 39.} This temporal perspective is one of the keys to von Hirsch’s theory of proportionality.\footnote{Von Hirsch, \textit{Censure and Sanctions}, \textit{supra} note 3 c. 2. Von Hirsch and Jareborg also say that a temporal perspective is essential to a determination of the level of harm, see von Hirsch \& Jareborg, \textit{supra} note 17 at 21-22.}

What does von Hirsch mean by “proportionality?” This term actually refers to two distinct but interrelated concepts: ordinal proportionality and cardinal proportionality.\footnote{Von Hirsch, \textit{Censure and Sanctions}, \textit{ibid.} at 17-19; Andrew von Hirsch, “Ordinal and Cardinal Desert” in Andrew von Hirsch \& Andrew Ashworth, eds., \textit{Principled Sentencing} (Boston: Northeastern University Press, 1992) 207 [Von Hirsch, “Ordinal and Cardinal”]; von Hirsch, “Proportionality”, \textit{infra} note 60.} Ordinal proportionality is concerned with developing a list of all crimes according to their seriousness (for our purposes, their blameworthiness).\footnote{This listing is what von Hirsch calls “rank-ordering.” See von Hirsch, \textit{Censure and Sanctions}, \textit{ibid.} at 18.} Ordinal proportionality is an exercise in determining the seriousness of crimes \textit{relative to one another}. This task is not as simple as it may appear. Crime-seriousness is not just a function of the harm caused by the offence,\footnote{\textit{Supra} note 11 and accompanying text.} but also of the culpability of the offender. Is an offence with more harm but less culpability more serious than an offence where there is less harm but more culpability? Perhaps, but perhaps not.\footnote{For a detailed discussion of the problems associated with this, see Bottoms, \textit{supra} note 2 (in particular “Puzzle 1: A Rickety Ladder?”) My object here is not to resolve this debate, but to point out that determining crime-seriousness would be, in and of itself, a difficult and time-consuming exercise. This challenge will not be taken up here.}

Ordinal proportionality is more than just a list of offences ranked in order of seriousness. There must also be appropriate distances between all of the offences to reflect both differences and similarities between the levels of seriousness for each

\footnote{(2002) 28 Queen's L.J.}
offence. In the end, one would ideally have a list of offences containing all conceivable combinations of harm and culpability, with appropriate distances between them.

But ordinal proportionality only does half of the job. It ranks offences relative to one another. Cardinal proportionality, on the other hand, is concerned with the overall punitive nature of the punishment scale. Put another way, cardinal proportionality sets the outer limits of the system of punishment.

Both ordinal and cardinal proportionality are essential to von Hirsch’s theory. The role of cardinal proportionality is to make sure that the overall severity of the punishment is not too high. If it is too high (for example, the death penalty for a first assault), it treats offenders as if their lives have no value whatsoever. If it is too low (probation for a pedophile), it indicates that the system does not care about those who are victims of these crimes. Ordinal proportionality allows the system to make judgements and generalizations about the relative seriousness of crimes.

51. There are two sub-requirements of proportionality used to determine the distance between offences. “Parity” means that like offences should be treated alike. When the offences are not alike, “spacing” determines how much of a difference there should between the penalty for each of the offences. See von Hirsch, Censure and Sanctions, supra note 3 at 18-19.

52. Ibid. c. 5.

53. Ibid. at 17-19. Note that not all writers who believe in cardinal proportionality necessarily subscribe to the idea of ordinal proportionality put forward by von Hirsch. See Norval Morris, “Desert as a Limiting Principle”, in von Hirsch & Ashworth, supra note 6 at 180. Also, if one believes in proportional punishments, this does not necessarily mean that one accepts desert is the primary determinant of punishment. See Hart, supra note 7 at 25.


55. Von Hirsch, Doing Justice, supra note 1 at 126; von Hirsch, Censure and Sanctions, supra note 3 at 10-13. Nonetheless, von Hirsch says that there may come a time in the future when crime has become so rare as to reduce the need for this lower boundary. Therefore, technically, this lower boundary is not required as a matter of proportionality. For a discussion of this, see von Hirsch, Censure and Sanctions, ibid. at 37-38.

56. See von Hirsch, Censure and Sanctions, ibid. c.5, where von Hirsch discusses another use of the term “anchoring the penalty scale” than that used in our discussion here. There, he discusses the desire to bring overall punishment levels
us assume that society decides that a relatively major theft warrants a $50 fine. Then, ordinal proportionality takes over and sets the punishment for every other offence in the punishment scale relative to the seriousness of the crime that warrants a $50 fine.

In addition to its use as a theoretical concept, proportionality also finds its way into statutory provisions used in actual sentencing decisions. For example, Canada’s Criminal Code provides that the sentence should be proportionate to both the gravity of the offence and the degree of responsibility of the offender.\textsuperscript{57} Von Hirsch views this provision as one of the most positive pieces of Canadian sentencing reform.\textsuperscript{58} Other commentators suggest, however, that despite this statutory directive, the exercise of judicial discretion in sentencing vacillates between proportionality and other contrary principles.\textsuperscript{59} As a result, it is far from clear that proportionality actually rules the sentencing landscape.

\section*{II. Von Hirsch’s View of Prior Record}

While von Hirsch has remained constant in his commitment to the use of proportionality in punishment,\textsuperscript{60} his explanation of the
down over time based on a scale of reducing penalties as far as crime prevention concerns will allow. As can be seen in that chapter, this is what von Hirsch calls a “decremental strategy.”

\footnote{57. Criminal Code, supra note 15, s. 718.1.}


\footnote{59. See Allan Manson, Essentials of Canadian Law: The Law of Sentencing (Toronto: Irwin Law, 2001), at 74-81 [Manson, The Law of Sentencing]. In terms of “contrary principles,” individualization is said to have “gained ground” (at 81) in the Canadian sentencing context. However, individualization was one of primary components of rehabilitationism. Von Hirsch is in favour of proportionality and opposed to individualization. See von Hirsch, Doing Justice, supra note 1 c. 4.}

\footnote{60. See, for example, von Hirsch, Doing Justice, ibid. c. 8; von Hirsch, Past or Future Crimes, supra note 17 c.3; supra note 3 c.2; Andrew von Hirsch, “Proportionality in the Philosophy of Punishment: From “Why Punish?” to
relationship between prior record and the quantum of punishment has evolved over time, and his current thinking involves a rejection of his earlier accounts. Before I respond to these rejections, it is necessary to outline this evolution in some detail.

In *Doing Justice*, the argument for assigning less punishment to first-time offenders (as opposed to recidivists) rests on two different footings. The first relates to culpability of the offender, and the second to institutional limitations of the criminal law itself. Von Hirsch put the first argument as follows:

In assessing a first offender’s culpability, it ought to be borne in mind that he was, at the time he committed the crime, only one of a large audience to whom the law impersonally addressed its prohibitions. His first conviction, however, should call dramatically and personally to his attention that the behaviour is condemned. A repetition of the offence following that conviction may be regarded as more culpable since he persisted in the behaviour after having been forcefully censured for it through his prior punishment.

The second argument is premised on the idea that it is very difficult to judge a person’s culpability. Von Hirsch argues that the sentencing judge cannot really be certain of the culpability of a first-time offender, but that each subsequent offence committed by that offender should lead to increased confidence in the judgement of culpability.

In *Past or Future Crimes*, von Hirsch steps away from the rationale for first-offender leniency offered in *Doing Justice*. He argues that less punishment of first offenders is still justified, but on a different basis. The offender is said to be making a “plea in self-extenuation”: the offender is saying that the behaviour is not like him, and therefore, he should be granted a partial reprieve.

64. *Ibid.* at 86.

D.L. MacPherson
from the consequences of his or her actions. Society's acceptance of this plea is based on the recognition of the frailty of the human psyche. Society "tolerates" a first offence (without excusing it entirely), because anyone can suffer from a lapse in moral judgement, however, the differential between punishment for first-time and subsequent offenders should be modest. Interestingly, in Past or Future Crimes, von Hirsch indicates that it is unclear whether the differential constitutes a discount for first-time offenders, or a premium for recidivists.

In Censure and Sanctions, von Hirsch did not deal with the issue of prior record in any detail. In "Desert and Previous Convictions," he once again draws on the arguments made in Past or Future Crimes, and "Criminal Record Rides Again", however, at the very least, he clarifies that the differential treatment is a discount for early offenders, and not a premium for recidivists.

66. Ibid. at 81-82.
67. Ibid. at 83-84.
68. Ibid. at 89-90.
69. Ibid. at 90-91.
70. This is interesting because, in a 1981 article, von Hirsch put forward a convincing case that the differential treatment was, in fact, a discount for first-time offenders, and not a premium for recidivists. See Andrew von Hirsch, "Desert and Previous Convictions in Sentencing" (1981) 65 Minn. L. Rev. 591 at 613-614.
71. In each of his first two books, von Hirsch dedicated an entire chapter to a discussion of the relevance of previous convictions. In Censure and Sanctions, supra note 3, on the other hand, von Hirsch stood pat on the position put forward in Past or Future Crimes, supra note 17 and "Criminal Record Rides Again", infra note 73. (See von Hirsch, Censure and Sanctions, ibid. at 18, n. ****, 60).
73. In fact, most of the substantive arguments made in "Criminal Record Rides Again" (1991) 10:2 Criminal Justice Ethics 2 at 2, 55-57, are reiterated in "Desert and Previous Convictions", supra note 6. Some of those ideas are expanded upon, while some are simply put forward a second time (often in very similar language). In the interests of preventing unnecessary duplication, I will summarize only the later article.
74. Von Hirsch, "Desert and Previous Convictions", supra note 6 at 192.
The explanation of tolerance first put forward in Past or Future Crimes is also expanded upon in “Desert and Previous Convictions.” In essence, this explanation runs as follows: lapses in judgement happen to the best of people. The fact that the lapse leads to the breach of a criminal prohibition makes the lapse no less a matter of the human condition. Therefore, society provides a discount in recognition of this human frailty. In other words, to a modest extent, tolerance mitigates the punishment that is otherwise deserved. The discount is withdrawn gradually for repetitions of the offence because the offender has stepped further from what society accepts as a product of human frailty and looks more like someone who is not willing to make an effort to avoid criminal conduct. After the third or fourth offence, the discount is withdrawn entirely, and there can be no increases for the prior record of the offender beyond the loss of the discount.

75. Ibid. at 193-194.
76. Ibid. at 195.
77. Ibid. at 197. Although this article responds only to the von Hirsch view of prior record, this is not to suggest that von Hirsch’s view stands alone on this issue. Other theorists have also presented theories on the relevance of prior record. For example:

(a) George P. Fletcher – Fletcher agrees with von Hirsch that the offender’s desert should determine the quantum of punishment, and that desert should be determined by both harm (which Fletcher generically calls “wrongdoing”) and culpability. However, in Fletcher’s view, prior criminal record should not influence sentencing at all, whether positively or negatively. See George P. Fletcher, Rethinking Criminal Law (Boston: Little Brown & Company, 1978) at 461-466. Therefore, even though von Hirsch’s conclusions on the relevance of prior record are driven in part by his desert rationale for punishment generally, Fletcher’s views demonstrate that accepting a desert rationale for punishment does not lead inexorably to the acceptance of von Hirsch’s account of prior record;

(b) James Q. Wilson – Wilson views prior record as highly relevant to his “selective incapacitation” theory of punishment. Selective incapacitation theory operates on the assumption that recidivism among criminals can be predicted on the basis of several factors. The criminal justice system can then prevent recidivism by incapacitating those criminals who show a high risk of re-offending if released. While Wilson agrees that prior record by itself is not necessarily an accurate predictor of a high risk offender, all seven of his factors must or could potentially be affected by the presence or absence of a
As I noted at the end of the previous section, the principle of proportionality (and its components of harm and culpability) has received strong legislative approval. However, von Hirsch's views on the relevance of prior record have not been as influential in the practical setting. When examining any normative theory, consideration should be given to how far it deviates from current practice. In my view, if the theory is otherwise logically defensible, then a high degree of consistency with current practice is a positive attribute of a theory, and should be recognized as such. Courts and commentators on sentencing in Canada have consistently held that a long record of similar offences does aggravate the deserved sentence. As Professor Manson explains: "... a long list of related offences is always a seriously aggravating factor which cannot be ignored, although it may, in rare circumstances, be rebutted." There has been little suggestion that prior record ceases to be relevant as soon as it is no longer a criminal record. See Wilson, supra note 39 at 113. Therefore, criminal record is very relevant to Wilson's theory.

(c) Julian V. Roberts - Roberts (along with von Hirsch and others) occupies something of a middle ground between these two extremes of Fletcher and Wilson. Roberts maintains a desert-based rationale for punishment. He argues that the justification for differential treatment of recidivists could be put on three different footings: (i) flat recidivist premiums; (ii) treating the increase as part of a notional "suspended sentence" for the prior offending; and (iii) connecting the relevance of prior criminality to parole eligibility dates for the current offence. See Julian V. Roberts, "The Role of Criminal Record in the Sentencing Process" in Michael Tonry, ed., Crime and Justice: A Review of Research, vol. 22 (Chicago: The University of Chicago Press, 1997) 303 at 351-355. While there are interesting conceptual issues raised by Roberts' analysis, space does not permit me to respond fully to his argument here.

78. I do not mean to suggest that a normative theory on sentencing (such as that put forward by von Hirsch or the one presented here) need necessarily conform to the practice "on the ground." However, if two theories are equally defensible on a normative basis, but one would require less of an overhaul of current practice than the other to apply the theory in practice, this is an indicator that the one theory may be more acceptable than the other.

mitigating factor, as von Hirsch argues. In Part V, we will return to consider whether my theory of prior record comes any closer to current practice.

III. Prior Record and Culpability

Before continuing, I should clarify one point. The rest of this paper assumes that the offender’s prior record is for the same offence as that for which he or she is currently being sentenced. If different offences were to be considered, we would have to determine how they should be treated. To do so would require addressing difficult questions. First, assume that the offender’s crimes are of increasing violence (for example, from assault to aggravated assault). Should the simple assault be considered as an aggravating factor in sentencing for the aggravated assault? Does the same result follow if the prior record is for aggravated assault and the current offence is simple assault? Perhaps it should, but there are arguments to be made on both sides of this issue. Second, should previous offences need to be “similar” to the current offence in order to be considered as part of prior record, and if so, how do we determine “similarity” in this context?

By not dealing with these issues, I do not deny their importance. On the contrary, I recognize that to deal with them would greatly increase the scope of my undertaking, and that they are sufficiently important to warrant independent treatment. Instead, I seek only to address three rather modest questions. First, is recidivist offending relevant to an offender’s desert? In my view, the answer to this question is yes. Von Hirsch objects to this answer, and I will attempt to respond to his objections. Second, if prior record is relevant to desert, how large should increases in punishment for prior record be? Third, when should these increases be included in punishment? If I were to consider any

---

80. For a Canadian case which accepts the idea of never increasing the deserved sentence for an offence due to prior record, but only using a lack of prior record as a mitigating factor, see R. v. Hastings (1985), 36 Alta. L. R. (2d) 193 at 198-199 (C.A.).
other issues here, it might obscure the answers to these three questions.

The first question relates to von Hirsch’s assertion that the existence of a prior record is not relevant to the offender’s desert. To examine this, let us return to the principles of harm and culpability. Do prior convictions add to the harm occasioned by the current offence? I agree with von Hirsch that, on the “living-standard” account of harm, the answer to this question is no. A victim’s living-standard does not suffer more because the offender previously committed the same offence against other victims. Therefore, prior record does not affect the harm done.

Von Hirsch also contends that culpability is unaffected by the offender’s prior record. In the process of reaching this conclusion, he suggests that the seriousness of the offence could be increased by a subsequent offence if one subscribes to defiance theory. I will consider defiance theory in more detail below. Does the idea that culpability remains constant in the face of subsequent misdeeds accord with our everyday understanding of culpability? I think not. Imagine that a parent is disciplining his or her child for taking a cookie without permission. The first time this occurs, the parent may choose to calmly explain to the child that an adult must give permission before a cookie may be taken. On the second occasion when the child transgresses the rule, the same warning may be given, with a much more stern tone. For the third violation, the child may be sent to his or her room. For the fourth transgression, other privileges (such as watching television) might be removed.

---

81. Von Hirsch, Doing Justice, supra note 1 at 87; von Hirsch, Past or Future Crimes, supra note 17 at 79. Likewise, on the Feinberg account, discussed in Harm to Others, supra note 17, the victim’s choices in life are no more restricted by the crime simply because the perpetrator of the crime has committed similar offences in the past against other victims.

82. This issue is raised in the context of increasing the harm element of the offence, because the “defiance” of the offender is an additional harm to be considered. However, the argument could apply equally to both harm and culpability. Therefore, I will respond to this objection as well.
Although this is an example of extra-legal punishment, it illustrates how ordinary people deal with repetitive misbehaviour on a common-sense basis. However, an appeal to common sense is only a first step. Can this intuitive response be justified as a matter of principle? Earlier, I introduced two different theories of culpability: character theory and choice theory. Let us look at each of them in relation to prior record.

Von Hirsch has said that character theory allows for prior criminal record to be considered highly relevant to the punishment of the offender. I agree with this assertion. After all, if culpability is based on the existence of bad character traits, which show themselves through bad acts, the more the bad acts are repeated, the more severe the judgment of culpability can be.

Character theory is very much concerned with the connection between the act and the ability to infer a bad character trait from that act. In fact, the ability to establish such a connection is a necessary pre-condition to the acceptance of character theory. In his discussion of criminal liability for negligence, Bayles indicates that the basis for blame (that is, the connection between the act and the bad character trait that it usually demonstrates) is much less certain in negligence than for other mental states. When the connection cannot be established, the actor is not blameworthy, so no punishment should be imposed.

If one were looking to strengthen the evidence supporting the ordinary inference of bad character, repetition of the act would seem a logical place to begin. If one is unsure of the connection between the act and bad character after the first offence, surely a

83. For a discussion of what is meant by “extra-legal” punishments in this context, see supra note 10.
84. Von Hirsch, Doing Justice, supra note 1 at 87; von Hirsch, Past or Future Crimes, supra note 1 at 77-78.
85. See Bayles, supra note 35.
86. Ibid. at 8.
87. Ibid. at 10-11.
88. Ibid. at 7-8. Excuses are designed to cover situations in which the ordinary inference of a connection between the act and character is not warranted, ibid. at 10.
repetition of the offence would make the state more confident of its culpability judgment. The more confident the state becomes in its assessment of the offender's culpability, the more blame can be attributed and the greater the punishment that is justified. In other words, if character theory is to be believed, recidivism more strongly demonstrates the applicable character flaw with each subsequent offence, and therefore, indicates that more punishment is justified.

Can the same be said of choice theory? Does the common-sense example given above find a principled basis in Moore's conception of culpability for choices? In my view, it does. Any offender has made a culpable choice to violate a specific prohibition of the law, but a repeat offender has made an additional choice not to react to the previous criminal conviction as society has the right to expect that a rational moral agent would.

The concept of moral agency is central to desert theory. If a person chooses to act as if he has not been persuaded by moral communication, the concept of censure has had insufficient effect on that person at that time. I do not mean to suggest that this is a flaw in von Hirsch's theory. My point is more basic. One of the basic assumptions of desert theory is that all people\textsuperscript{89} (whether offenders or not) can, if they so choose, act as rational moral agents.\textsuperscript{90} The opposite side of the coin is that people can, if they so choose, not act as rational moral agents. Is this a culpable choice? I think the answer is yes. A choice not to exercise one's capacity to do something good can be just as culpable as a choice to exercise one's capacity to do something wicked.\textsuperscript{91} For present purposes, it is sufficient to maintain that it is a culpable choice to not respond to censure as would a rational moral agent. I will return to consider how much culpability this choice demonstrates.

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{89} Assuming they are not affected by mental illness or some other cause which renders them incapable of making rational moral judgements.
\item \textsuperscript{90} Von Hirsch, \textit{Censure and Sanctions}, supra note 3 at 5.
\end{itemize}
\end{footnotesize}
From von Hirsch’s theory of proportionality, it was seen that culpability in the absence of harm and/or an offence would not be sufficient to warrant criminal punishment. So the question must be asked: where does one find both the harm and the offence? The simple answer is that the harm is supplied by the current offence for which the accused has been convicted. The increased culpability is an adjunct to the harm, and the offence for which the offender is now being punished.

Thus, I have arrived at the conclusion accepted by von Hirsch in Doing Justice: that is, prior record affects culpability for the current offence; however, I do so for different reasons. First, von Hirsch talked about “persistence in behaviour” despite previously “being forcefully censured” for that behaviour. Secondly, he expressed a concern about the accuracy and reliability of our culpability judgements with respect to first-time offenders.

The Doing Justice account of prior record appears to be primarily concerned with the past. Von Hirsch saw persistent behaviour and previous censuring attempts as justifications for saying that recidivist offending entails increased culpability.

These concerns do not drive my conclusion that prior record can increase culpability. Rather, my focus is on the present choice of the actor and not on the previous censure. To be clear, the previous censure does inform the justice system’s view of the offender’s current choices. However, the current choice of the actor (not to react as would a rational moral agent) remains primary. Choice, as we have seen, is a sufficient basis to ground culpability.

Put another way, in my account, the previous censure does entitle society to expect an offender to react as a rational moral agent. It triggers the right of society to expect a reaction by the offender, but that ends its role. Beyond that point, the concern is solely with the actor’s current choices. Therefore, while the conclusion drawn here (of increased culpability for recidivist offending) is in the result the same as that in Doing Justice, the reasoning used to arrive there is quite different.
In *Past or Future Crimes*, von Hirsch rejects the above conclusion. His objections are twofold: (a) recidivism is evidence of increased culpability, but not proof of it; and (b) such an approach involves an acceptance of defiance theory, which is not an appropriate approach to the issue of deserts. I will call the first of these the "evidence, not proof" objection; the second will be referred to as the "defiance" objection. Let us look at each of these objections in turn.

The "evidence, not proof" objection can in my view be accepted without really affecting my conclusion. It is true that recidivism may not in every case prove an increase in culpability. But this does not mean that recidivism never increases culpability. Von Hirsch's objection appears to run essentially as follows: since a single act of recidivism does not always prove an increase in culpability, it is not relevant to culpability at all. It is one thing to say that the fifth time an offender commits a particular offence, the evidence of recidivism is not sufficient to say that his or her culpability has increased since the fourth time. This is most likely true. But what about the seventh time that the offender commits the offence? Or the 10th? Or the 15th? After that many repetitions, most reasonable people would say that the offender was more culpable than after the third or fourth commission of the offence. The offender has done something more blameworthy by breaking the same law on that many occasions. He or she has not reacted appropriately to any of the previous "chances" to alter his or her behaviour. Therefore, to address this part of the objection, it must be conceded that each recurrence of the offence need not necessarily, in itself, lead to increased culpability. But sufficient repetitions may increase culpability.

---

94. Von Hirsch uses this rationale to explain the withdrawal of the "discount" for early offenders when they commit subsequent offences. I use it here for a different reason. The non-reaction of the offender as a rational moral agent is the choice that grounds culpability in the actor for subsequent offences. It is important to keep the two arguments conceptually distinct.
95. I will discuss this issue in more depth in Part V, below.
This point raises an additional normative issue that merits some consideration. Any time one creates a theory that considers prior record as anything other than a mitigating factor in sentencing, it looks as though society is punishing the offender again for his first offence. This is what I will refer to as the “double punishment trap.” Essentially, the fact of recidivism alone cannot morally justify increased punishment.\(^6\) I agree with this assertion. Therefore, the question becomes what is it about the recidivist offending (other than the fact of recidivism) that justifies the increased punishment.

As explained above, it is only if the additional offence demonstrates an additional culpable choice by the actor (the choice not to respond as one would expect of a rational moral agent) that the theory kicks in. It is this additional choice (and not the simple fact of repeat offending) that is used to justify the additional punishment. Only previous offenders will have previous experience with censure and have triggered society’s right to a response. Nonetheless, as described above, this does not necessarily mean that every case of recidivism should increase punishment. Therefore, this does not fall into the double punishment trap.

But there is a more fundamental point to be made by von Hirsch’s “evidence not proof” objection. Von Hirsch contends that if we allow prior record to be used as evidence of increased culpability, there is no reason to stop there. Anything that sheds light on a possible increase in the offender’s culpability would thus, according to von Hirsch, be fair game.\(^7\)

There are two inter-related answers to this contention. The first lies in a principle of the criminal law to which I have already briefly referred: certainty.\(^8\) The certainty principle requires that an offender know what factors will be taken into account in determining punishment. If the justice system were to take into account factors other than prior criminal record in determining

---

96. See Ruby, supra note 79 at §7.6.
97. Von Hirsch, Past or Future Crimes, supra note 17 at 79.
98. See supra note 30 and accompanying text.
culpability, this might very well undermine the certainty principle. For example, let us assume that a first-time assault offender was disciplined in school for fighting. Does this increase his or her culpability for a later criminal assault? I would argue that it does not. Only those acts that have brought the actor before the criminal justice system are relevant to the culpability of the actor. To suggest that everything in the actor’s past is relevant to his or her culpability for the current offence goes against the certainty principle. In my view, prior acts that were not the subject of criminal sanction should be excluded from the determination of culpability.

But prior record stands apart from that general exclusion. By the public nature of the conviction and sentence for the prior offence, society indicates that criminal wrongdoing will be considered in judging the culpability of any future criminal act. It should be clear to the offender, before he or she commits a subsequent offence, that his or her prior criminal convictions, unlike other factors, may have an impact in determining the punishment for subsequent conduct.

The second answer to von Hirsch’s concern about the use of evidence other than prior record to increase culpability lies in the distinction between legal and non-legal contexts. Let us return to our fighting in school example. In the non-legal context, there may not be a level of common understanding to allow laypeople to communicate to one another effectively the level of harm done by the act and/or the culpability of the actor. In our example, was the victim hospitalized due to serious injury? Or was the fight simply a shoving match in which no one was hurt? In the non-legal context, both could legitimately be described as a “fight.” Third-party observers might not seek sufficient detail to differentiate between these two meanings of the term “fight.” Most likely, an assumption would be made in the absence of the requisite information.

But in the legal process, the context is different, in at least three significant ways. The first relates to offence descriptions. There is a level of common understanding among lawyers and judges as to what is meant by “assault” as opposed to “assault causing bodily
harm” and “aggravated assault.” Offence descriptions tend to be relatively precise and specific, as compared to everyday non-legal language. This reduces the likelihood that the language used by lawyers and judges in the offender’s previous sentencing will be misunderstood by those involved with the offender’s current sentencing.

Second, the sentence imposed for previous offences adds to the judge’s confidence in his or her assessment of previous legal judgments of harm and culpability. For example, let us assume that the previous sentence imposed was minimal compared to that allowed by statute. This is an indication that the harm caused by the previous offence and the culpability of the offender were at the lower end of the scale. Conversely, if the previous sentence was at the maximum level, it is likely that both the harm caused and the culpability of the offender were very high. In order to draw such a conclusion, a comparison must be made between the offender before the court and other offenders convicted of the same crime. Statutory minimum and maximum sentences, as well as statistical averages, are or could be available to help judges in making this determination. However, since there is generally no code of punishments available in non-legal contexts, such an assessment is not usually possible. Therefore, the availability of the sentence for the offender’s previous crimes (and relevant comparison groups) also distinguishes judgements of harm and culpability in the legal context from other similar judgements.

Third, the legal context tells us, through its laws on evidence and procedure, and in other ways, what are the legally relevant facts, how they may be proved, and what are the issues to be decided. The law of sentencing considers certain facts not to be relevant to the sentencing decision. For example, the fact that an accused person forced the Crown to prove its criminal allegations at trial, as opposed to pleading guilty, is not an aggravating factor in sentencing.99

On the other hand, in everyday situations, the fact that a person refuses to admit and acknowledge his or her wrongdoing is often held against him or her in imposing punishment. When a parent asks a child about a misdeed the child might say nothing until the parent offers proof of the misdeed. The parent may then proceed to punish the child both for the transgression and for the failure to admit the mistake when first confronted. In other words, the rules that are applied in the legal context do not always apply in other situations. Therefore, it is inappropriate to import instances of non-legal punishment into the criminal process. Only prior criminal record can justifiably increase culpability.

Therefore, my response to von Hirsch’s “evidence, not proof” objection is that the cause-effect relationship between recidivism and increased culpability is not automatic. However, this does not necessarily lead to the conclusion that the relationship cannot exist. It simply tells us that the relationship must be examined with some care, to ensure that the conclusion of increased culpability is warranted in the circumstances. If it is, then increased punishment should result. Further, the fact that one accepts prior record as evidence of increased culpability does not mean that every indicator of increased culpability need necessarily be considered in determining punishment. Certainty requires that the factors taken into account in making that determination in the criminal sphere be known in advance. A prior conviction puts the offender on notice that his or her prior record will be used in determining culpability for subsequent offences. Also, in the legal context, culpability judgements are supported by offence descriptions, sentence comparison groups and rules regarding what is legally relevant: these provide a degree of confidence in assessing the meaning to be ascribed to previous convictions. The same cannot be said of the non-legal context, so evidence other than prior record should not be used to increase the offender’s culpability.

This responds to the first of von Hirsch’s concerns. But how can one respond to his “defiance” objection? To explain this
objection, in *Past or Future Crimes*, von Hirsch relies on the following quotation from George Fletcher:

I would argue that, in a liberal society, defiance should not constitute a wrong that justifiably enhances the punishment a recidivist deserves . . . . In some social contexts, such as the family, the military, or a football team, those in authority might regard it as justified to respond to threatened defiance with a penalty greater than that otherwise thought appropriate for the offense already committed . . . . But do we want to conceptualize civil society in this way? A society that respects individuality and diverse purposes hardly stands in analogy to organizations devoted to specific goals, such as winning a war or a game.¹⁰¹

There are two interrelated answers to the “defiance” objection. First, the state is an authority that must be respected. Many crimes can only be explained by reference to the authoritative nature of the state. Common examples would include failing to appear after release from pre-trial custody¹⁰² and breaching a probation order.¹⁰³ Another example is the more severe treatment of the intentional killing of a peace officer, as compared to the intentional killing of other people.¹⁰⁴ The rationale for these offences presupposes that the criminal law and its agents should be respected. So the first answer to the “defiance” objection is that defiance may be relevant to culpability for certain criminal offences. This is not to suggest that it explains culpability generally, but rather that its use as a factor in labelling certain

---

¹⁰². See Criminal Code, supra note 15, s. 145(2).
¹⁰⁴. See Penal Code, Cal. Stat. Ann. §190 (a), (c) (West 1988). Paragraph (a) indicates that the standard penalty for second-degree murder is one of 15 years to life. However, paragraph (c) says that if the victim is a peace officer who was engaged in the performance of his or her duties (and if the defendant knew or reasonably should have known that the peace officer was engaged in the performance of his or her duties), the mandatory penalty for second-degree murder is one of life imprisonment without possibility of parole.
behaviour as criminal is not always objectionable or as problematic as von Hirsch and Fletcher suggest.

Choosing to defy authority can be relevant to criminalization. To suggest otherwise could cripple the operation of the criminal justice system. Let us assume, for example, that the failure to appear after being released on bail was not an offence. Then the criminal justice system would have to choose not to have a bail system at all or to have one with no sanctions attached. If the former were chosen, the jails would become clogged with offenders who might not need to be detained prior to trial. If the latter were chosen, those on pre-trial release could violate the conditions of that release with impunity. Obviously, neither of these alternatives is practicable. Therefore, society needs to have offences like failure to appear and breach of probation. These have an underpinning of defiance that helps to justify criminalization.

The second answer is slightly different. I would argue, in fact, that I have not slipped into defiance theory at all to justify the increase in culpability for recidivist offending. Jean Hampton’s essay, “Mens Rea”, is a good example of defiance theory, and a closer look will help to resolve whether my account is based on defiance and not on desert.

For Hampton, the justification for punishment is dependent on blame, but the culpability required for punishment is defined

105. Von Hirsch, Doing Justice, supra note 1 at 130-131, n.96. Von Hirsch suggests that there is a distinction between the general system of sanctions (based on desert) and the instrumental sanctions necessary to operate the system (based on deterrence). He specifically mentions pre-trial confinement in this discussion.

106. Supra note 26.

107. To be fair, I find much about Hampton’s analysis to be problematic. As I said above, while defiance may have some role with respect to the criminalization of certain offences, I fundamentally disagree with Hampton as to its general applicability. My objections to the Hampton account deserve a separate discussion which space does not permit here. I mention the Hampton account only to show that (despite the protests of Fletcher and von Hirsch to the contrary), one need not rely on defiance theory in order to justify increased culpability for repeat offenders.

108. Hampton, supra note 26 at 21.
by a particular choice: the choice to defy the law. This necessarily entails three points: (a) the actor must make a decision to choose the action; (b) at the time of the choice, the actor must know that the action was prohibited; and (c) the prohibition is made by an authority that the actor knows is binding in the circumstances. The actor then chooses to supplant the actual authority with another one that allows the actor to do what he or she wants.

An example will assist here. Imagine that Person A wants to hit Person B because Person B made a derogatory remark about Person A’s brother. Person A knows that the law does not allow him to hit Person B, but he does it nonetheless. Person A is asked why he did it. The response might well be, “I had to do it to defend my younger brother. The law had nothing to do with this.” In other words, Person A knew the better thing to do (not to assault Person B) but did the worse (by assaulting Person B). This is what Hampton calls defiance.

Does my account fall under defiance theory as presented by Hampton? In my view, it does not. In arguing that prior record increases culpability, my account is not concerned with the ability of the rational moral agent to substitute a valid authority figure (the law) with a different reason or authority that will allow the action the agent wishes to perform. I am concerned instead with the agent’s choice not to exercise his capacity to react as should a rational moral agent. This account is focused on the morally culpable reasons for which the act was done. This is the essence of choice theory.

Put another way, the Hampton account is concerned with the process of decision-making. The primary question in her account seems to be: how does one justify to oneself the decision to take an action which one knows is otherwise culpable? The account presented here is more concerned with the person making the decision, on the one hand (is the person a rational moral agent?),

109. Ibid.
110. Ibid. at 15.
111. Ibid. at 16.
and on the other, the *product* of that decision-making (was the choice a culpable one?).

Therefore, in sum, neither the "evidence, not proof" objection nor the "defiance" objection stands up to close scrutiny as an objection to the theory I am advancing. A skeptic might argue that, despite my concentration on a desert-based theory of punishment, I am really using a disguised form of either rehabilitation or character theory to justify my approach to prior record. The skeptic might say that I focus on how many more offences the offender has committed in order to determine the prospects of the offender redeeming him- or herself, or to determine the character of the offender.

Rehabilitation was the predominant refrain of punishment theory in the early 1970s. It was in response to rehabilitationism that von Hirsch's first book was written. This approach is inherently forward-looking. If the goal is to rehabilitate the offender, this can only be done by altering current patterns of behaviour and replacing them with more desirable ones.

As should be obvious from the discussion above, I view prior record from a backward-facing perspective. At the time of conviction and sentence, the court judges the culpability of the offender *as of the time of the offence*. The court should not look into the future and attempt to (at best) guess as to what the future prospects of the offender may be.

As noted above, a basic tenet of the criminal law is certainty. Desert theory defines the factors relevant to the sentencing decision with a reasonable amount of precision. The harm caused by the offence, the culpability of the offender at the time of the offence, and the prior record of the offender are all relevant to the punishment meted out. Rehabilitation theory would seem to

113. Ibid.
114. Ibid.
115. See supra note 30 and accompanying text.
adopt a much broader stance on the factors relevant to the decision to punish. Presumably the court would have to examine all aspects of the offender’s situation to determine whether he or she is susceptible to rehabilitation. Even assuming that the court is capable of carrying out such a broad-ranging inquiry, this would still leave the offender uncertain as to how deeply his or her life would be probed for this purpose, and as to how a court might assess his or her likely response to attempts at rehabilitation. This violates the certainty requirement of the criminal law.

The same problem presents itself when one deals with culpability for character as opposed to acts. In the theory presented here, acts are what matter, not any character flaw that the act demonstrates. The theory presented here is concerned with specific choices of the offender resulting in morally reprehensible conduct. Certainty, in my view, could be violated if one were to accept a character-based rationale. If bad acts demonstrate character flaws,\textsuperscript{117} and bad character is what justifies culpability, then the search for bad acts (and the bad character that they demonstrate) is presumably not restricted to those acts that have been proven in the criminal justice system. Like rehabilitation theory, character theory, in my view, leaves the offender too uncertain of what will be held against him or her by the criminal justice system the next time he or she commits an offence.

In sum, the theory I am presenting does not lapse into rehabilitation under the guise of blame: it looks backward; rehabilitation looks forward. Similarly, this theory is not about the character of the offender; it is focused instead on acts that were committed by the offender. His or her character is not in issue.

Having explained the normative basis of my theory in this section, I move on to look at questions of application, and in particular, at how the increased culpability of the recidivist offender will affect the punishment for the current offence.

\textsuperscript{117} See supra note 37 and accompanying text.
IV. Developing a New Theory

I have tried to explain von Hirsch’s theory of the relevance of prior record, and I have found it in need of revision in some respects. Therefore, I have suggested a theory justifying increased culpability of repeat offenders. However, I am not suggesting that von Hirsch’s “progressive loss of mitigation” idea be rejected. On the contrary, I believe that this idea captures the thought process that ought to occur when punishing those who have committed a small number of offences.

I agree with von Hirsch that it is unreasonable to expect anyone never to have a lapse in judgement, including a lapse that leads to criminal behaviour. In short, to err is human. To place full blame on a person for something that is part of the human condition seems less than fair and just. In other words, I find the reasoning behind von Hirsch’s idea that lapses in judgment are to be expected from people (and that lapses in judgement that lead to criminal behaviour are no less so) is compelling.

Nonetheless, I cannot agree with the conclusion to which von Hirsch’s “loss of mitigation” theory drives him: that once the discount evaporates for subsequent offending, there is no further increase in culpability. The von Hirsch theory on progressive loss of mitigation for early offences should instead work together with the “increased culpability” theory outlined below.

If the choice not to react as a rational moral agent is a culpable choice, the next question is: how culpable is it? How much of an increase in culpability does it represent? The answer, I believe, lies in the same rationale as the discount for early offending. If society provides a modest discount for early offending, then modest increases in culpability are called for by offences outside the discount range (in von Hirsch’s theory, the first three or four offences). There are three reasons for this. First, I have already said that the theory sketched here should work together with the “discount” theory offered by von Hirsch. Both theories involve only one factor in punishment: the presence or absence of prior record for the offender. The two should form a coherent whole. If
so, then the theories must be internally consistent. To have minor increases in punishment at the bottom of the scale and very large increases thereafter does not integrate them sufficiently. One would have two theories operating, but they would never combine into a rational, consistent whole. With modest increases, on the other hand, the two theories could be brought together seamlessly, providing a consistent framework for considering prior record in the allocation of punishment.

Second, it is important to remember that the rationale for increases in culpability is the fact that a person has chosen not to react as a rational moral agent should. Therefore, any increase in culpability resulting from that choice must respect the offender’s rational moral agency. This idea supports a minor increase in culpability for recidivism. Otherwise, punishment threatens to become more about the future incapacitation of the offender rather than about his or her deserts. Desert theory is about respecting the ability of a person to react to punishment as a rational moral agent. If the sentence is too long, society is treating people as animals to be controlled rather than as rational moral agents who can be reasoned with on a moral level.\footnote{For a discussion of this, see von Hirsch, \textit{Censure and Sanctions}, supra note 3 at 11.} If a society treats certain of its members as animals, this is exactly the way that those people will react.

Third, and most important, although I argue that prior record increases culpability, this is only one part of the culpability calculation. The concepts of purposeful (or intentional), knowing, reckless and negligent wrongdoing are more important factors in that calculation. The idea of prior record should not lead us to ignore these key concepts. If one were to accept large increases for prior record, not reacting as a rational moral agent would take primacy. Such an approach, it seems to me, would be misconceived. The current offence is the reason that the offender finds himself or herself before the criminal justice system. His or her increased culpability for failing to react as a rational moral agent is a necessary consequence of being before the criminal
justice system for the subsequent offence. To make prior record the primary determinant of culpability would be to make the effect (the increased culpability of the recidivist as compared to other offenders) more important than the cause (the fact that the offender, recidivist or not, has committed an offence).

Another way to look at the same argument is by examining the censure element of punishment. It was asserted earlier that censure is about moral communication. Censure is designed to convey moral blameworthiness. What is more important to our societal assessment of blameworthiness: the recidivism of the offender or the offence that was committed? In my view, it must be the latter. In other words, the more relevant thing is that the offender committed the culpable act that formed the basis of the offence. If that is so, the censure (the message of blame), and thus the punishment must be expressed more in terms of the present offence than for prior record. If large increases were permitted for prior record, the majority of the moral communication would be focussed on the less blameworthy of the two elements.

The increases should therefore be minor. But should the increases occur with each subsequent offence? As may be obvious from the earlier discussion, the answer is no. This is how the criminal justice system can respect the "evidence not proof" objection raised by von Hirsch. The link between subsequent offending and increased culpability is not automatic. For the link to be sufficiently certain to warrant increased punishment, a number of subsequent offences would have to be committed. For example, society may be uncertain that the culpability for a fourth offence is greater than that for a third offence. But can the same uncertainty exist with respect to an increase in culpability between a third offence and a sixth offence? Surely, in that scenario, society would be justified in saying that the choice not to react as a rational moral agent is definitely a culpable one, and that the actor can be held accountable for it.

If culpability does not necessarily increase with each subsequent offence, then how should the criminal justice system take account of prior record? This concept is best described as a series of plateaus. The first plateau could be found, perhaps, after the third
offence, consistent with the "progressive loss of mitigation" approach put forward by von Hirsch. 119 This would represent crime-seriousness absent any consideration of prior offending (or the lack thereof). The second offence would be lower on the punishment scale than the third offence; the first offence would be lower than the second. However, the fourth or fifth offence may not result in increased punishment. Therefore, between the third offence and the fifth offence, there is a plateau, in the sense that punishment remains constant. However, at the sixth offence, there would be a modest increase in punishment to reflect the actor's culpable choice not to respond appropriately (that is, as a rational moral agent should react) and to alter his or her behaviour. There would then be further modest increases to reflect further culpable choices at, perhaps, the ninth offence, and possibly at the twelfth offence as well. A precise formula is not critical. It is sufficient to say that an increase in punishment after every subsequent offence is not warranted and to leave the remaining decisions for further discussion.

Are there any limits to the theory of punishment I have just described? Since prior record is part of culpability, these increases in punishment do not violate ordinal proportionality, because culpability is one of the elements to which punishment must be proportional. In other words, increased culpability justifies increased punishment. Therefore, there are no ordinal proportionality constraints on these increases.

Cardinal proportionality is another story, however. It sets the limits on how punitive the punishment scale is allowed to be. So cardinal proportionality does limit the increases in punishment attributable to the theory of prior record presented here: society cannot allow the increases attributable to prior record to become so severe as to raise concerns that the entire punishment scale is too punitive for a moral society to use.

However, is this sufficient? Individual increases in punishment due to prior record should be modest. For the same reasons presented earlier, the modest individual increases must not be

119. See supra note 6.
allowed to become too large in the aggregate: otherwise, the benefits of the modesty of the increases at each plateau will be lost when the number of increases becomes too high. For example, five or six modest increases may be roughly equivalent to one or two very large increases. If so, the modesty of the increases at each plateau may not be sufficient to avoid the pitfalls referred to earlier.

The theory presented here needs to meet this concern about the aggregation of penalties. One possible solution is to say that only a specific number of individual increases will be allowed. For example, after three or four increases, the punishment scale would hit a ceiling of some sort, and there could be no further increases on account of prior record. However, I do not favour such an answer to this dilemma. It seems too simplistic and mechanical a solution to apply to an institution as varied and complex as the criminal justice system.

Perhaps the solution could include fixing punishments within a certain percentage of the punishment that would otherwise be meted out in the absence of prior criminality. For example, one suggestion might be that the punishment at the second plateau could be only 1.25 times the maximum punishment available at the first plateau. Further increases might have to be asymptotic, that is, if the maximum difference on account of prior record at the second plateau were 1.25 times the punishment imposed at the first plateau, then the punishment imposed at the third plateau could be no more than 1.20 times that imposed at the second plateau, and so on.

There would still be a ceiling of some sort, or else aggregation could become a serious problem. However, judges would have more opportunity to arrange the punishment so that the ceiling

---

120. Von Hirsch uses such a concept of percentage increase maximums to deal with aggravating or mitigating factors (other than harm and culpability) in punishment. See supra note 1 at 99-101. In discussing aggravation and mitigation, von Hirsch leaves aside prior record, see ibid. at 101. However, here I use this concept not to discuss aggravation or mitigation, but to help determine the seriousness of the crime committed. This is an essential part of desert, and not an aggravation or mitigation of it.
would be less likely reached. The mechanics of deciding which multiples to use are relatively unimportant, but it is important to recognize that large aggregate increases raise the same potential difficulties (although possibly in a less acute form) as large individual increases.

At the end of Part III, I indicated that, in my view, von Hirsch’s theory on prior record generally does not accord with Canadian practice.\textsuperscript{121} Does the theory presented here do any better in this respect?

Let us return to Professor Manson’s assertion that “a long list of related offences is always a seriously aggravating factor” in sentencing. My theory does allow for increased punishment if the offender has a long record of committing a particular offence. In that sense, this theory does come closer to current Canadian sentencing practice than von Hirsch’s theory.\textsuperscript{122}

There is an additional point to be made about the congruence of this theory with Canadian sentencing practice as well. This congruence is found in a phenomenon that one commentator calls the “jump effect.”\textsuperscript{123} This effect occurs when an offender has been sentenced to similar penalties for similar offences, and the court that is now sentencing the offender decides to “jump” the penalty to a much higher level than the offender has previously experienced. Courts of appeal are generally opposed to such a jump. As Clayton Ruby explains:

This idea [of a “jump effect”] has never been held to justify increasing the sentence beyond that imposed previously; rather, it is a rule that tends to limit

\begin{itemize}
\item \textsuperscript{121} See supra notes 78-80 and accompanying text.
\item \textsuperscript{122} My theory is that prior record is an essential part of desert, and in particular, of the culpability discussion. Therefore, the term “aggravating factor” is not technically part of my theory of prior record. The terms “aggravating” and “mitigating” factors imply that one finds a fit sentence and then adds to or subtracts from the fit sentence based on these outside factors. In my theory, prior record helps in determining a fit sentence and does not add to and subtract from it. Also, see supra note 120.
\item \textsuperscript{123} Ruby, supra note 79 at §7.71.
\end{itemize}
such an increase—where otherwise appropriate and necessary—to one imposed in an incremental manner.\textsuperscript{124}

In my view, my theory complies with the rule to which Ruby refers. Ensuring that the increases for prior record are minor should mean that there are no drastic jumps of the type that the rule is designed to prevent.\textsuperscript{125} Furthermore, the asymptotic nature of the increases will help to ensure that the rule about the "jump effect" is respected.

Therefore, the theory presented here does manage to remain reasonably "in step" with current Canadian sentencing practice. However, the congruence of theory and practice is far from perfect. For example, I have argued in this paper that other offences of which the offender has been convicted are the only acts prior to the current offence that are relevant to culpability\textsuperscript{126}. Acts without a criminal conviction should not matter. This is not in accord with Canadian sentencing practice. Any previous conduct which amounts to an offence for which the offender could have been charged, but was not, can be either proved by the Crown or admitted by the offender at the sentencing hearing, and may be considered by the sentencing judge.\textsuperscript{127} Nonetheless, the judge must sentence only for the offence before the court.\textsuperscript{128} Therefore, although my theory comes closer to current Canadian practice than does von Hirsch's theory, there are still significant gaps between my theory and current practice.\textsuperscript{129}

\begin{enumerate}
\item \textit{Ibid.} at §7.72.
\item \textit{Ibid.} at §7.71.
\item See my response to von Hirsch's "evidence, not proof" objection, in Part IV, above.
\item Manson, \textit{The Law of Sentencing}, \textit{supra} note 59 at 152.
\item \textit{Ibid.}
\item See \textit{supra} note 78.
\end{enumerate}
Conclusion

Professor Andrew von Hirsch has developed a cogent theory of punishment based on the desert of the offender. However, in my view, the desert rationale offered in von Hirsch’s theory need not drive his conclusion with respect to prior record.

The argument has been made in this essay that a desert rationale for punishment can still leave a role for prior record in determining the desert of the offender. Specifically, prior record can affect the culpability of the offender. The choice of the offender not to react as a rational moral agent, by making the decision to offend again, increases culpability. However, care must be taken to ensure that the link between subsequent offending and increased culpability is justified in the circumstances.

Increases in punishment for prior record, even if justified in the circumstances, should be modest. The increases should not occur with every offence. Rather, there should be a series of plateaus designed to account for increases in culpability caused by prior record. Finally, the aggregate increase to punishment for prior record must also be kept within limits. While there are still many questions left open by this analysis, these questions can be taken up in the future. This essay has simply attempted to set the stage for further discussion.