

Kantian Punishment and Ethics: Dichotomy or Manifestation of Impure Ethics?

Julien Lafleur

A Major Research Paper submitted in partial fulfillment of the requirements for the
Master of Arts in Political Science

School of Political Studies
Faculty of Social Science
University of Ottawa

©Julien Lafleur, Ottawa, Canada, 2022

Kantian Punishment and Ethics: Dichotomy or Manifestation of Impure Ethics?

Punishment is not a new phenomenon. Its purpose and methods and the reasoning behind it have evolved throughout history. Parents often use it to educate their children, sports referees to keep the game fair, pet trainers to house-train animals, friends to communicate disapproval, and governments to deter bad behaviour and enforce social norms (Goodwin & Gromet, 2014, pp. 561–562). It is an enduring feature of society that may take many forms, from the application of physical pain, the privation or prohibition of goods, seclusion, and social ostracization, to name only a few. Expectedly, due to the prominence of punishment in society, the desire to rationalize its use and the harsh treatments sometimes associated with it has made it a recurring theme in social and political philosophy. Among leading theories on the subject are (a) consequentialist theories, which are focused on determining a punishment based on its potential future benefits to society; (b) restorative theories, which focus on repairing the harm caused to the victims; and (c) retributive theories, which aim at obtaining justice by making the offender suffer in accordance to the moral blameworthiness of the crime (Goodwin & Gromet, 2014, pp. 561–562). This research paper focuses on the latter.

Retribution is an interesting topic because it pushes into an instinctive human desire to punish those who do us wrong through pain and hardship, which not only appears harsh and destructive to society but also seems to conflict with the moral premise it theoretically relies on because determining the moral blameworthiness of a crime is undoubtedly a very subjective endeavour. Finding an appropriate punishment could also leave the victim displeased or the criminal unjustly punished through a sentence that is too lenient, excessive, or against the wrong person (Wood, 2008, p. 208). Therefore, it is arguable that for a retributive penal system to be just and fair for everyone, it should rely on a robust moral theory, one with strong laws that are

clear and universal, to minimize the chances of undue punishments.

A philosopher who claimed to have written such moral theory is the renowned German philosopher Immanuel Kant. In his work *The Metaphysics of Morals*, Kant clearly states his views on punishment. Punishment is a categorical imperative (Kant, 2017, p. 114 [6:331]). It must be performed unconditionally and serve no other purpose than to punish the wrongdoer for the committed crime, seemingly grounding Kant's views on punishment as retributive and well-anchored within his moral theory (2017, p. 114 [6:331]). However, there are many instances throughout Kant's works where his views on punishment appear to conflict with his moral theory. Some of his claims, such as the necessity for a punishment to involve suffering and be "physically bad" (Kant, 2002, p. 55 [37]), appear contradictory to our modern concept of moral good. His insistence on the death penalty in some instances, its exclusion at other times, and the subjectivity necessary for assessing the moral blameworthiness of a crime also appear to be incompatible with many aspects of the categorical imperative (e.g., the formula of universal law), which is at the source of Kant's moral theory.

Additionally, although Kant claims that punishment should never serve as a deterrent for future crimes, his examples do not always reflect this idea. For example, his consideration for the imposition of higher fines on the upper classes (in contrast to the lower classes) so that they do not "indulge in a verbal insult on some occasion" (Kant, 2017, p. 115 [6:332]) clearly demonstrates some thoughts toward the idea of deterrence, making the punishment not strictly retributive and seemingly incompatible with Kant's categorical imperative. Consequently, this paper focuses on the following question: How does Kant establish punishment as a manifestation of pure ethics, and why do some of his examples appear to demonstrate otherwise? It argues that Kant's views on punishment are theoretically loyal to his moral philosophy grounded in pure

reason and that his disputed examples reflect a pragmatic morality that was increasingly present in Kant's mature works, even though it was never put into a formal system. It also argues that Kant's controversial examples of punishments can be established as manifestations of what Louden (2000) calls Kant's impure ethics, pragmatic moral examples which serve as an imperfect roadmap to pure morality rather than showing its embodiment. Although the debate surrounding Kant's retributivist views and his moral theory is not new, it is hoped that this fresh perspective can provide additional insights while simultaneously challenging the supremacy of metaphysics in Kant's moral work by allowing room for pragmatic moral considerations. It should be noted that the term pragmatic does not refer to the philosophical doctrine of 'pragmatism' but merely to the term in its everyday sense. In particular, it refers to moral practices based on experience rather than pure reason.

In the spirit of keeping this research focused, this paper assumes that Kant's views on punishment are mostly retributive and concentrates strictly on the tensions between his views, some of his examples, and his moral theory. The methodology used consists mainly of a textual interpretation. A textualist interpretation aims to capture the complexities and frictions within an author's work by focusing highly on their texts and with minimal constraints on the interpretation (Beiner, 2013, pp. 37, 41). It gives minimal regard to historical circumstances other than a need to acknowledge that the authors might not have always been concise in their writings for various reasons, such as fear of persecution or ostracism (Beiner, 2013, p. 40, Note 14). Therefore, it is an appropriate methodology for this paper and is precisely how the argument is constructed, relying heavily on Kant's primary texts and using scholarly literature to support its claims. It uses some historical and contextual elements to demonstrate the impact of religion on Kant's life, as well as his interest in anthropology and the development of a pragmatic moral

system. However, these elements are kept to a minimum, and the primary approach remains a textual interpretation focused on explaining the tensions between Kant's views on punishment and his moral theory.

This paper is broken down into four parts. The first part consists of an overview of the existing scholarly literature. The second part establishes the main theoretical framework and affirms Kant's views on punishment as compatible with his moral theory by deliberately leaving his examples aside and focusing strictly on its core theoretical concepts. The third part focuses on some problematic examples, using Louden's work on impure ethics to show how they can support Kant's pragmatic moral thoughts without necessarily contradicting his work on pure ethics. Finally, as a supporting argument, the last part demonstrates how Kant's critique of religion is also a pragmatic moral project that directly influenced his views on punishment and his examples in *The Metaphysics of Morals*.

Literature Review

Scholars do not express unanimous consent on the connection between Kant's views on punishment and his moral theory. The literature reveals many arguments on the subject that can be divided into two central positions. On the one side, some scholars maintain that Kant's views on punishment are retributive and fit within his moral theory (Fleischacker, 1988; O'Connell, 2014; Pincoffs, 1966). On the other side, some scholars view a detachment between Kant's moral theory and his views on punishment (Brooks, 2003; Byrd, 1989; Hill, 1999; Wood, 2008). It is important to mention that no scholar in the latter group really argues against some form of retribution in Kant's views on punishment because his explanations and examples are explicit on the matter, implying that Kant at least thought of himself as a retributivist. For this reason, this literature review adopts a similar structure but leaves the challenges to the retributive nature of

his views aside. First, we will explore the positions and arguments of some scholars defending the consistency between Kant's retributivism and his moral theory, followed by those of some critics embracing a more nuanced or mixed Kantian approach to punishment.

O'Connell (2014) is one of the scholars who defend Kant's retributivist views as consistent with his moral theory. O'Connell mentions that some moral debates surrounding Kant's retributivism revolve around intentionally making a wrongdoer suffer for his crime (2014, p. 478). In other words, how can one morally justify suffering? O'Connell argues that our modern understanding of suffering as intrinsically immoral and our view that making someone suffer is generally wrong are very distant from Kant's understanding of suffering (2014, p. 478). In Kant's position, someone can be a good person and unhappy (through suffering) simultaneously, as personal well-being and morality are two different things (O'Connell, 2014, p. 494). What is essential is to be morally good as a person. The act of suffering has no intrinsic moral worth.

Another of O'Connell's (2014) points concerns Kant's reasons for punishing someone through suffering. Suffering may not have intrinsic moral worth, but is punishing someone through suffering morally justified from the punisher's perspective? If it is, how is it also considered a duty, as claimed by Kant? O'Connell explains that the act of punishing through suffering is closely related to Kant's moral concept of the highest good (O'Connell, 2014, p. 485). Kant's concept of the highest good consists not in maximizing one's happiness but in achieving happiness proportionally to one's moral worthiness (Kant, 2002, pp. 141–142 [110]). Therefore, O'Connell argues, punishment (through retribution) is morally justified by acting as the means to achieve this equilibrium and ensuring no undue suffering can occur by guaranteeing that "virtue is rewarded with proportional happiness and vice is punished with proportional

unhappiness” (O’Connell, 2014, p. 488). In other words, punishing someone through suffering is a way to make them “happy” in proportion to the moral worth of their act. From a Kantian perspective, it justifies the moral imperative of punishing through suffering.

Finally, O’Connell’s third major argument regards the duty to punish. Answering a claim that aiming for the highest good means being virtuous but does not necessarily imply punishing others for their lack of it, O’Connell explains that the act of punishing can be justified in two ways. First, it can be deduced from Kant’s moral theory, where some of his examples clearly insist on punishment and suggest that it is a perfect duty and must be done categorically (2014, p. 489). Second, because one can make moral judgements that approximate proportionality, one can also set proportional punishment. Because acting towards the highest good is always encouraged, punishing through retribution then becomes justified as it works towards the highest good (2014, p. 490).

O’Connell’s (2014) article explains well the moral imperative to punish vice and the role of suffering in Kant’s views on punishment. However, his article was designed to address specific critiques (Byrd, 1989, Hill, 1999, Wood, 2008) and mainly focused on suffering. It gives a credible account in favour of Kant’s retributivism but leaves certain topics unexplored that could undermine the link between Kant’s views on punishment and his ethics. For instance, many examples Kant gives in *The Metaphysics of Morals* appear to contradict his principle of proportionality, such as his insistence that a mother committing infanticide or a soldier duelling for his honour should not be punished by death for their crime (Kant, 2007, p.118 [6:336]). How can there be exceptions to moral principles that are meant to be universal and categorical? Without exploring why Kantian examples sometimes contradict his own proportionality thesis, one cannot fully assert that his views on punishment are loyal to his moral theory.

Another supporter of Kant's retributivism is Fleischacker (1988). Where O'Connell (2014) focuses largely on the notion of suffering, Fleischacker's attention is placed on the tension between the coercive aspect of punishment and the need to regard people as free and autonomous beings. More specifically, Fleischacker explores how one can justify, from a Kantian moral perspective, applying force on someone else for an act that has already been committed and how as a criminal, one can will as a universal law a system that will punish him for his crimes. In other words, argues Fleischacker, "we need a moral ground ... not only for our right to punish, but also for our right to consent to be punished" (1988, p. 435).

With regards to the right to punish, Fleischacker (1988) explains that an objection often brought up by critics is that, although it is easy to justify the use of force to stop a crime in progress (because the crime is hindering someone else's freedom), it is more difficult to justify the use of force on someone to punish them for a past offence when no hindering of freedom is currently occurring (1988, p. 435). Fleischacker argues that Kant sees retribution as serving two purposes. First, from the perspective of positive law, it acts as a deterrent, as all punishments do, to prevent future hindrances to freedom (1988, p. 435). Second, it also has a moral function: "making the criminal live under the law he implicitly sets up in his criminal act" (Fleischacker, 1988, p. 442). That is to say that by using retribution, society applies to the criminal his own criminal maxim, as if he had willed it to be universal, showing him the irrationality of his act (Fleischacker, 1988, p. 442). Therefore, coercion is morally justified through retributive punishment even if the crime is not currently happening. The second issue is how a criminal could morally wish for a system that exerts coercion on him (through fear of punishment) to prevent him from doing crimes he wishes to do (Fleischacker, 1988, p.435). Fleischacker explains that, for Kant, the deterrent aspect of punishment is not as crucial as its moral worth. A

punishment must be just, meaning it must be deserved and proportional to the harm committed (Fleischacker, 1988, p.435). According to Fleischacker, a criminal does not will the act of being punished but can morally will punishment in itself, as long as it fulfills the requirement of justice. If the punishment is deserved and proportional to the crime, Kant says the criminal will “have to admit that justice had been done and that his reward perfectly fitted his behavior” (Kant as cited in Fleischacker, 1988, p. 435; Kant, 2002, p. 55 [37]).

Overall, Fleischacker (1988) provided an account of how Kant sees the deterrent effect of punishment as not necessarily undermining its moral worth. His interpretation of retributivism by which the criminal is served “a rational interpretation of his irrational act” (1988, p. 442) can provide us with a moral ground to punish and be punished. However, as with O’Connell (2014), textual tensions remain between some of Kant’s examples of punishment and his moral theory. If one of the moral aims of retribution is to show the criminal the irrationality of his act (1988, p. 442), as Fleischacker claimed, why does Kant believe we should be more lenient with punishment in certain circumstances? Once again, Kant’s examples of punishment in *The Metaphysics of Morals* appear to get in the way of establishing Kant’s views on punishment as theoretically aligned with his moral work.

Fleischacker’s (1988) interpretation of Kant’s retributivism is similar to the one written by Pincoffs (1966) a few years earlier. In his book, Pincoffs established Kant’s views on punishment as retributive on the ground that the reason to punish is that a crime was committed, that the punishment must be proportional, and, most importantly, that it is deserved (1966, p. 4). Pincoffs also reflects on the deterrent effect of punishment and wonders how one can justify and explain that a punishment must not act as a deterrent for future crime but strictly as a punishment for the committed crime. He proposes an explanation for the problem using Kant’s categorical

imperative formulas. He explains that by committing mischief, a criminal is accepting its universalization. To punish him is to show him the consequences of what he willed, as a free rational agent, to be a universal law (Pincoffs, 1966, p. 8). Therefore, the punishment does not primarily aim at preventing other crimes but shows the criminal the consequences of what he has universalized. By acting on him as he acts on others, society is showing him the logical consequence of his actions (Pincoffs, 1966, p. 9). From this perspective, although the punishment might deter the criminal from doing it again, it does not result from the deterrent aspect of the punishment but rather from the criminal's realization that universalizing his action would be wrong. It makes retributive punishment a deterrent without explicitly being one and would be consistent with Kant's reasoning that punishment must serve no other purpose and only be executed because the person has committed a crime (2017, p. 114 [6:331]).

Pincoffs's (1966) interpretation is not without critics. Wood states that attempting to merge and justify Kant's retributivism through the formula of universal law is wrong, as the former belongs to the doctrine of right and the latter to Kant's doctrine of virtue (2008, pp. 217–218). He also claims that, although it is true that a criminal's maxim cannot be universalized and that for Kant, a breach of juridical law is also a breach of the moral law, there is no proper justification in Kant's writing that would allow punishing the criminal by following his own non-universalizable act (2008, p.217). As well, if treating a criminal according to his criminal maxim was consistent with the universal law, it should also apply to maxims that breach ethical duty (2008, p. 218). It would make it morally acceptable to lie to a liar or to not be charitable to a selfish person in need. As Wood points out, acting as such requires the punisher to act on a maxim that cannot be willed as a universal law, which does not work with Kantian ethics (2008, p. 218).

Overall, Pincoffs's (1966) views on the matter reflect an understanding that Kant's views on punishment are compatible with his ethics. However, Wood (2008) also brings excellent counterarguments against punishing someone using his own criminal maxim against him. Also, for the punishment to show the criminal the consequence of what he has will to be a universal law, the punishment would always have to be precisely the same as the crime, which is not always the case.

On the other side are scholars that see Kant's views on punishment as not entirely retributive and conflicting with his moral theory. As mentioned above, Wood (2008) is one of these scholars. Wood acknowledges that some aspects of Kant's theory of punishment are unequivocally retributivist, such as Kant's claim that criminals must be punished only because they have committed crimes and that punishment must be immediate, necessary, and equal in severity to the blameworthiness of the crimes (Kant, 2017, p. 114 [6:331-32]; Wood, 2008, pp. 210–211). However, he also argues that Kant's views on punishment do not receive much support from his practical philosophy and are even in "serious tensions with some of his most fundamental moral doctrines" (Wood, 2008, p. 206). First, Wood claims that Kant does not provide any convincing account of how the law of punishment fits within his concept of external rights, most specifically why it has to be immediate and necessary that a criminal be harmed equally to his crime (2008, p. 214). He claims that as a law that is also a categorical imperative, the necessity, means, and urgency of punishment should come as self-evident, which is not the case, and only rests on the fact that Kant declared it to be (2008, p. 214).

Perhaps more importantly, Wood (2008) describes Kant's retributivist views as conflicting with his ethics for two reasons. The first is the duty toward the happiness of others (2008, p. 219). Kant states that the happiness of others is one of the two duties that are also ends

in themselves (Kant, 2017, p. 160 [6:386]). When one puts the happiness of others as one of one's own ends, one reduces adversity, pain and other conditions that could tempt people to act immorally (Kant, 2017, p. 162 [6:388]). However, as pointed out by Wood, because a requirement of retributive justice is to make the offender suffer in accordance with the moral blameworthiness of the crime, retribution appears to be in contradiction with the requirement to put the happiness of others as one's own end (2008, p. 219). In other words, it creates rather than mitigates the conditions inducing people to act immorally. Second, Wood sees an issue with the coercive aspect of retribution. He claims that if Kant sees retribution as permissible to enforce juridical law, it cannot act as a way to coerce people into doing their ethical duties, as these need to be conducted from our own rational free will to classify as categorical imperatives (2008, p. 220). It is problematic for Wood as he claims that Kant implies in many instances that the goal of punishment is to attack the "inner wickedness" of a person, which goes against that principle (2008, p. 220). Finally, Wood asserts that the principle of the highest good is primarily a way to assess one's own inner worth and does not entail that one should hurt themselves or others to achieve this goal (2008, p. 222). Wood supports his claim by underlining how Kant condemned religious penance and mostly linked the highest good to retribution when talking about God and how He is the only one entitled to inflict punishment (2008, p. 221).

The following conclusions are made from Wood's (2008) analysis of Kant's retributivism. First, morality in Kant is all-encompassing. For punishment to align with Kant's moral theory, one needs to understand and interpret the examples of punishment from the doctrine of right without contradicting the core of his doctrine of virtue. Second, the highest good can be interpreted as a principle strictly used to assess one's own inner worthiness and not as a right to punish others, which Wood and Hill say is a matter that belongs to God only (Hill,

1999, p. 409; Wood, 2008, p. 221). The punishment of vice as a heavenly matter is worth investigating because religion plays a role in Kant's work that is sometimes close to a pragmatic application of Kantian morality. Allowing for elements of pragmatic morality in Kant, although controversial, could explain the discrepancies with some of Kant's examples of punishment.

Similarly, Byrd (1989) describes the irregularities between Kant's retributivism and ethics and attempts to reconcile the two. She argues that Kant discusses not one but two different elements of punishment. The first element is coercive and sees punishment as a threat to deter crime and enforce each others' rights (1989, p.151). The second element of punishment regards its execution. It is primarily retributive and serves no other purpose than to punish because justice requires respect for individuals as ends and not merely as a means to some other ends (e.g., to deter crimes; 1989, p. 151). This way, retribution becomes a limitation on the state's right to punish rather than its goal (Byrd, 1989, p. 153). Byrd's argument reflects Kant's distinction between positive law, which regulate external behaviour and is concerned with hindrance to freedom, and moral law, which regards internal motivation and is concerned with the maxims of actions. Byrd argues that the deterrent aspect of punishment is a necessary feature of positive law to ensure everyone's right to freedom (1989, p. 171). She says it is something acknowledged by Kant that "if a certain exercise of freedom is itself a hindrance of freedom under universal laws (i.e., wrong) then the force, which is employed against it, as a limitation on a hindrance of freedom, is harmonious with freedom under universal laws, i.e., right" (Kant as cited in Byrd, 1989, p. 170). However, she argues that because using deterrent punishment to preserve freedom must act preemptively to be effective, a strictly retributive theory of punishment cannot work because retribution must occur only after the fact (Byrd, 1989, p. 183). Because Kant clearly sees retribution as the only morally just punishment, Byrd argues that the

best way to explain it is by making a distinction between punishment as a threat (within the context of legislating) and the execution of punishment (through the judicial and penal system; Byrd, 1989, p. 183).

Wood praises Byrd's (1989) thesis as "the most defensible theory of the state's right to punish that is available to Kantian Ethics" (2008, p. 212), even though he suggests it to be untenable. From Wood's perspective, Kant's retributivism cannot be strictly about the legal execution of punishment. As mentioned earlier, Wood states that as a philosopher, Kant's views on punishment cannot be taken strictly from a legal practice perspective and must be viewed more broadly, from a general perspective, including both his philosophical theory of right and ethics (Wood, 2008, p. 213). He asserts that attempts (such as Byrd's [1989]) to ground the justification of punishment primarily in deterrence rather than retribution neglects the importance put by Kant on retributivism in some of his texts, where retribution is cited in moral and even religious contexts (Wood, 2008, pp. 213, 216). Although Byrd adequately describes the coercive aspect of the law with respect to positive law in Kant's doctrine of right, it does not account for Kant's straightforward declaration that the law of punishment must serve no other purpose than to punish and must be inflicted only because a crime was committed (Kant, 2017, p. 114 [6:331]). This statement alone makes relegating the moral aspect of punishment strictly to its administration debatable. Although Byrd provides one of the most plausible interpretations of Kant's views on punishment, Wood's response is also convincing. If we are to allow punishment to deter in any capacity, we must find a way to also make it work with Kantian pure ethics. As we argue, one way to do this is by challenging the primary assumption that Kantian ethics is strictly grounded in metaphysics with no consideration for the pragmatic reality of life.

Finally, a similar approach is envisioned by Brooks (2003). Brooks argues that Kant's

views on punishment are retributivist regarding moral law but open to consequentialism regarding positive law (p. 206). Brooks highlights examples from Kant's *Metaphysics of Morals*, where the philosopher's views reveal some consequentialist thoughts. For example, Brooks mentions how preserving the community in Kant outweighed the need for retribution if "the number of criminals be so great a number that the state would be left without any subjects if they were all executed" (2003, p. 213). Consequentialist consideration also appears in his example of the right of necessity, where he states that the murder of a castaway man when it is for one's own survival should not necessarily be punished with death (Brooks, 2003, p. 214).

Overall, Brooks' (2003) argument about the consequentialist nature of some of Kant's examples of punishment also seems to indicate a tendency by Kant for pragmatic moral consideration. However, as mentioned earlier, it is also a problem as it puts Kant's views on punishment at odds with his moral theory grounded in pure reason. As Wood mentioned (2008), it appears unlikely that Kant's ethics can be separated from his doctrine of right because it needs to be all-encompassing for the former to work. The question as to why some of Kant's examples on punishment demonstrate consequentialist thoughts when he also states that punishment is a categorical imperative that "can never be inflicted merely as a means to promote some other good for the criminal himself or for civil society [but rather] only because he has committed a crime" (Kant, 2017, p. 114 [6:331]) remains unsolved. Consequently, Brooks' (2003) article also confirms the dichotomy between Kant's views on punishment and his moral theory that we will attempt to reconcile.

Overall, this brief literature review touched on an array of interpretations of Kantian retributivism, as well as some sense of its most common problems. To see if Kant's views on punishment are truly compatible with his moral theory and understand why some of his examples

seem to point otherwise, we must first reconcile the need to punish strictly based on desert while also accounting for the deterrent aspect of punishment concerning positive laws. We must also explain why the proportionality thesis, which is so important to Kant's retributive claim and the highest good, is not always considered in some of his examples of punishment. In doing so, we must not discount Kant's other works or rely only on the examples and concepts from *The Metaphysics of Morals*. Specifically, the notions of God and divine justice could make a closer examination of Kant's critique of religion a productive endeavour, because *Religion within the Boundaries of Mere Reason* was published only 4 years earlier than *The Metaphysics of Morals* and Kant's idea of grounding religion within the scope of his moral theory could have influenced his views on punishment.

In the next part, we will first go over some of the core notions of Kant's pure moral theory. Because the aim is to establish Kant's view on punishment as supported by his moral theory, we focus strictly on the concepts most often cited in the literature on punishment. It is not an extensive review and assumes some familiarity with Kant's moral theory. After that, we will explore the work of Louden (2000) on Kant's impure ethics. Louden's work suggests Kant was well aware of the pervasive hold of some of our inclinations and, therefore, allowed room for temporary pragmatic moral practices without necessarily contradicting his whole moral system based on pure reason. Louden's work serves as our theoretical framework to analyze the ostensible subjectivity in some of Kant's most puzzling examples of punishment.

Kant's Moral Theory

Kant's moral theory is complex, grounds morality in the metaphysical world, and is established through the use of pure reason. Two prominent doctrines can be found at the heart of it: the doctrine of right and of virtue. The doctrine of right concerns our external conduct as

human beings interacting with each other, and it sees actions as being right when they do not impede others' freedom (Kant, 2017, p. 27 [6:230]). On the other side, the doctrine of virtue is inwardly focused and regards what is morally good such as our ends and the maxims of our actions (Kant, 2017, p.156 [6:381]). Both doctrines are connected to the concept of duty because external and internal laws must also be duties to be morally good. It means that they are binding and must be done categorically because they are good in themselves and not because they can further some other goals (Kant, 2017, p.18 [6:223]).

The concept of duty is essential to understand Kant's views on punishment because "what essentially distinguishes a duty of virtue from a duty of right is that external constraint to the latter kind of duty is morally possible, whereas the former is based only on free self-constraint" (2017, p. 158 [6:383]). In other words, a duty of right can be forced upon another through coercion as an incentive to act rightly while still being considered moral, but a duty of virtue cannot. It must be done without outside pressure and out of duty alone. Punishment is an external and coercive act imposed on others. Therefore, it should naturally belong to the doctrine of right. This situation causes tension because to envision punishment as a duty of virtue, one must envision it as an act of free self-constraint.

Other concepts of virtue found in *Grounding for the Metaphysics of Morals* are also important to understanding Kant's theory of punishment. The first ones are the concepts of categorical and hypothetical imperatives. In Kant's moral theory, in order to be morally good, acts must stand as morally good in themselves (they cannot be good because they accomplish some other purposes), which Kant calls categorical imperatives (Kant, 1993, p. 25 [414]). On the other side are hypothetical imperatives, which are completed to achieve some other ends, and that, although not necessarily bad, cannot be considered "morally" good in themselves (1993, p.

25 [414]). Therefore, in its most simplified form, a virtuous act is based on the categorical imperative's laws. It is an essential moral notion to our punishment analysis since Kant clearly stated that to punish is a categorical imperative (2017, p. 114 [6:331]). From the categorical imperative, three formulas are essential in understanding and interpreting Kant's views on punishment. The first is the formula of universal law, and the second is the formula of autonomy of the will. The third is the formula of the end in itself. After these formulas, the highest good and the two ends that are also duties will also be examined.

The formula of universal law states that one should "act only according to that maxim whereby you can at the same time will that it should become a universal law" (Kant, 1993, p. 30 [421]). Regarding punishment, this means that to consider retributivism as moral, one must be able to rationally will a world where all punishments would be retributive. As seen earlier, some of Kant's examples do not always appear to support purely retributive punishment, conflicting with this formula of the categorical imperative. Consequently, to claim that Kant's views on punishment align with his moral theory, we must investigate and explain Kant's examples of punishment that go against the formula of universal law.

Another important formula of the categorical imperative is the formula of autonomy of the will. This formula implies that morality, from a wilful perspective, is not simply a matter of agreeing to follow laws but thoroughly willing them as if one were the legislator of those laws (Kant, 1993, p. 38 [431]). That inner will makes one obey a law out of duty. For example, the law prohibiting murder is not ethically good in itself. It is right, as it protects the collective freedom necessary for happiness, but it is only ethically good because people genuinely believe that murder is wrong. What makes it a law of virtue is that we autonomously (without constraint) and rationally will it to be a universal law of nature. The formula of autonomy is essential to

determine moral practice and is crucial in determining how punishment can be both right and virtuous.

The third formula is that of the end in itself. This formula of the categorical imperative states to always “act in such a way that you treat humanity, whether in your own person or in the person of another, always at the same time as an end and never simply as a means” (Kant, 1993, p. 36 [429]). This formula means that when interacting with others, one should always respect them as humans and never only as means to achieve something else. For example, one should feed a hungry person because it is the right thing to do (treating him humanely) and not because of hoping to gain something from him. It is relevant to punishment as Kant states that it should be done strictly because the criminal deserves it and never to achieve some other purposes such as deterrence for future crimes or any other things that could benefit society in the long term (Kant, 2017, p.114 [6:331]). Yet, Kant allows for deterrence in his doctrine of right, and many of his examples of punishment also express some consideration for it. Therefore, it is a part of Kant’s moral theory that needs reconciliation with his views on punishment.

Another concept that needs to be defined is the highest good. In Kant, morality and happiness are two different things. Unlike utilitarianism, Kant’s highest good in life is not achieved through maximization of the amount of happiness but, instead, when happiness is distributed in equal proportion with one’s moral worth (Kant, 2002, p. 142 [110]). The moral end of the virtuous man is not to be happy but to be worthy of happiness (Kant, 2002, p. 141 [110]). It is an important moral concept because it allows us to interpret the notion of proportionality that is often mentioned by Kant in his views on punishment when determining the required severity of the punishment.

Finally, another important concept (reflected by Wood [2008] in the literature review) is

the notion of ends that are also duties. According to Kant, there are only two types of ends that are also duties: the duty to work toward one's own perfection and the duty of the happiness of others (Kant, 2017, pp. 160–161 [6:386-87]). Making the happiness of others a duty works towards the overall preservation of moral integrity because it helps remove obstacles such as fear, hunger, and disease (among other adversities) that make people unhappy and tempt them to follow their inclinations (Kant, 2017, p. 162 [6:338]). As mentioned earlier, punishment as the infliction of physical harm contradicts this concept by promoting vice rather than virtue and deserves further investigation.

Louden and Kant's Impure Ethics

The theoretical framework for this research is based on Louden's work on the presence of pragmatic morality in Kant, which he calls impure ethics (2000). Louden's work helps structure some of Kant's controversial examples as nonrepresentative of his pure moral theory without necessarily invalidating his claim that to punish is a moral duty. Any serious Kantian readers should have noticed how Kant's examples are notorious for being confusing and sometimes hard to put into context within his moral theory. His examples on punishment are no different and often suggest exceptions to what should otherwise be a moral duty (as we will describe later). Louden also points out rightfully that using practical examples to explain metaphysical notions is peculiar and appears paradoxical as one cannot prove pure theoretical concepts using real-life experience, something Kant would have been unquestionably aware of (Louden, 2011, p. 99). Any attempt at claiming compatibility between Kant's views on punishment and ethics must, therefore, go through an investigation of the significance and meaning of examples in Kant's works.

While acknowledging that Kant is best known for his pure philosophy, Louden argues

that a key element of Kant's philosophical work was the importance of "judgment sharpened by experience" (Louden, 2011, p. 72) in the application of the moral law, which he deemed essential to make pure laws effective in real-life applications. Louden first points out a few textual and contextual circumstances to support his argument. First, Kant was a university professor and no stranger to practical anthropology, having taught and written on subjects such as geography, education, and religion. In his years at the University of Königsberg, Kant spent a considerable time teaching empirical disciplines where he was said to be unsatisfied with giving only theoretical knowledge but also wanted to be involved in finding pragmatic applications of it (Stuckenberg as cited in Louden, 2000, p. 6). However, contrary to his pure ethics, Kant never really finished or developed a pragmatic moral system that could stand on its own, as most of his thoughts are scattered across his legacy (Louden, 2000, p. 6).

Louden (2011) also points out that Kant's interest in finding a pragmatic application to his moral law is reflected in his works and appears more consistent as his moral theory matures over the years. For example, morality in *Groundwork for the Metaphysics of Morals* (1993) has few pragmatic notions. Kant asserts in it that moral philosophy does "have its empirical part" (Kant, 1993, p. 1 [387]) and that if morality "does not borrow the least thing from acquaintance with the human being in justifying its basic principles...it does take *experience* [emphasis added] to distinguish in what cases they have their application" (Kant as cited in Louden, 2011, p. 73). However, although the morality explained in the groundwork shows some traces of empirical consideration, it is mostly considered a work of pure, nonempirical ethics (Louden, 2011, p. 66).

In *The Metaphysics of Morals* (2017), the original version of which was published 12 years after *Groundwork for the Metaphysics of Morals*, Kant acknowledges a more significant role of experience in shaping human behaviour. He includes both references to "the particular

nature of human beings, which is cognized only by experience” (Kant as cited in Louden, 2011, p. 66) and to a counterpart to the metaphysics of morals, which concerns “the subjective conditions in human nature that hinder people or help them in the carrying out of the laws of a metaphysics of morals” (Kant as cited in Louden, 2011, p. 55). In other words, Louden argues that older Kant recognized that some of our passions and inclinations make it hard for us to adhere to pure moral laws simply out of duty and that some moral anthropology is needed to guide us toward a metaphysics of morals (Louden, 2011, p. 55).

Louden also mentioned that, although Kant never established an official pragmatic moral system, he did recognize that reducing ethics to strictly pure rational principles would be unwise (Louden as cited in Muchnik, 2013, p. 463). His main concern was to keep pure rational principles separate from pragmatic moral laws so that no one “would then run the risk of bringing forth false or at least indulgent moral laws, which would misrepresent as unattainable what has only not been attained just because the law has not been seen and presented in its purity or because spurious or impure incentives were used for what is itself in conformity with duty and good” (Kant as cited in Muchnik, 2013, p. 463). The possibility of Kant acknowledging our human inability to always act accordingly to moral laws without guidance allows room to interpret some of Kant’s examples of punishment from a different angle. Instead of being representations of the pure moral law in action, some of them could be pragmatic guidance which is undoubtedly imperfect but leads towards a metaphysics of morals.

Kantian Pure Morality and Punishment

In the literature review, we described two prevalent tensions between Kant’s views on punishment and his moral theory. The first problem was the idea of suffering and, more precisely, how one can make someone suffer through punishment without going against the

fulfillment of the highest good or our duty toward the happiness of others. The second problem of importance was the difficulty of fulfilling the categorical imperative requirement while considering the undeniably coercive and deterrent aspects of punishment. This next part will exclude Kant's use of examples to focus on the question strictly from a theoretical perspective. It will demonstrate that although these issues appear to pose some difficulties, they are theoretically compatible with a Kantian morality grounded in metaphysics.

The first critique that making someone suffer goes against the fulfillment of the highest good and our duty towards the happiness of others can be explained through an interpretation of the two concepts. Kant's concept of the highest good does not aim to provide the highest amount of happiness achievable nor to provide it to as many persons as possible. Doing so would simply transform Kant into a utilitarian. Instead, Kant's end is for everyone to be deserving of happiness, and that each receives happiness according to his desert (Kant, 2002, pp. 141–142 [110]). Kant's retributivism serves this function. Punishing a criminal equates to giving out a negative amount of happiness proportional to his desert, which reflects the highest good. Similarly, the happiness of others, which Kant calls an end that is also a duty, is also fulfilled by following the same reasoning. It is fulfilled when, from all our might, we attempt to serve the criminal with a negative amount of happiness reflective of his desert (Kant, 2017, p. 161 [6:387]). It is not following our duty to punish, through exemption or by allowing a punishment larger or smaller than what the criminal deserves, that would be contrary to the highest good and the happiness of others.

A second critique relevant to sufferance brought up by Hill (1999) was the ability to judge a crime's blameworthiness. Hill answered that Kant's views on punishment refer only to an intrinsic desert to suffer rather than a duty to make others suffer for their crimes and that our

inability to judge the inner moral blameworthiness of an offence (because of its subjectivity) forces us to separate moral law from judicial law (Hill, 1999, pp. 413, 429). From a Kantian perspective, one could answer this conundrum by arguing that the inability to guarantee an adequate judgement of the moral blameworthiness of an offence is not a justification for inaction. First, labelling punishment as a categorical imperative makes it a duty that commands regardless of experience. As stated in *Grounding for the Metaphysics of Morals*, reason “commands actions of which...feasibility might well be doubted by one who bases everything upon experience” (Kant, 1993, p. 20 [408]). For example, Kant states that although we cannot ascertain that we ever had sincere friendships in our lives, it should not be a rejection of the a priori value of sincere friendship, and it is something that every person is required to attempt out of duty (Kant, 1993, p. 20 [408]). Through the same logic, our inability to guarantee an equal amount of retribution for the moral blameworthiness of a crime shall not preclude us from seeking it out of duty, even though we can never guarantee an appropriate punishment.

Following on the right to make others suffer for their crimes, a third perspective can be taken by looking at the doctrine of right. According to it, a punisher cannot be held responsible for any injustice arising from the execution of the law. Someone who punishes within the boundaries of the law (excluding revenge) is but the executor applying the laws enacted by the lawmakers. He does not punish the criminal from a personal perspective but on behalf of society, in accordance with duty and out of respect for the law. Because of this, he cannot be blamed nor punished if the punishment outmatched the moral blameworthiness of the crime, as he is only applying “coercion in conformity with the law” (Kant, 2017, p. 102 [6:317]). Therefore, punishing someone else for his crime through suffering is not in friction with Kant’s moral theory as long as it follows the law’s requirement, is retributive in its application and is executed

out of duty.

The second problem of importance was the difficulty of fulfilling the requirement of the categorical imperative while taking into account the undeniably coercive and deterrent aspects of punishment. In the scholarly literature, this is explained predominantly in two ways. Some argue that punishment as a categorical imperative should only apply to breach of the moral law, which would allow using deterrence for a breach of positive law only. Others argued that punishment must be all-encompassing and apply to both doctrines to be truly grounded in Kantian morality (to be a categorical imperative). For example, Pincoffs makes punishment compatible with both doctrines because he sees the deterrence aspect as only a derivative effect, and its primary purpose to make the criminal understand that his act cannot be universalized (1966, p. 8).

Alternatively, Wood (2008) asserts that the requirement of the categorical imperative should apply to both doctrines but cannot see in Kant's writings any compelling evidence that it is the case. He says the inherent coercive aspect of punishment makes it incompatible with Kantian morality because no one can coerce someone into virtue, and, according to Kant's moral theory, one can only become virtuous through his own free will (Wood, 2008, p. 220).

Like Wood, this paper argues that punishment requires us to look at both moral doctrines simultaneously, although it does not see its coercive and deterrent effects as contradicting Kant's claim that punishment is a categorical imperative. It is important to remember that when following Kant's pure ethics, it is not the act but the maxim of an action that needs to be universalized and follow the formulas of the categorical imperative. An act is always accompanied by its maxim. They must be viewed as complementary and analyzed similarly. The act of punishment must always be in accordance with the doctrine of right, and the maxim behind its legislation and execution must follow the moral law. The act of punishment can coerce

and deter as long as its maxim is grounded in virtue and its execution is based on justice and proportionality. Deterrence and retribution are not competing aims of punishment; they require each other (Ripstein, 2009, p. 301).

Kant's views on punishment are not at odds with his moral theory because, in civil society, every act of legal punishment inevitably creates deterrence and cannot serve as a benchmark of its morality. Only the maxim of the act of punishment can convey its morality by being just in its enactment (because it is deserved) and retributive in its execution (to ensure justice through proportionality). Confronting a criminal with the irrationality of his actions (Pincoffs's [1966] claim) may eventually lead to subsequent actions that are morally grounded in pure reason. However, this is a hopeful consequence of the act of punishment (the deterrent aspect), not its moral intent. It cannot relate to its morality because it would interfere with the criminal's ability to become virtuous of his own free will (Wood's [2008] claim). The moral intent of punishment for Kant is to aim for the highest good by fulfilling the requirement of justice through desert and proportionality.

In any way, whether an offender really becomes deterred due to the coercive effect of the act of punishment or because he has ultimately realized the wrongness of universalizing the maxim of his action from his own free will, is something that no one but God can assert. Therefore, similar to why the inability to judge the blameworthiness of a crime should not prevent us from seeking justice, the inability to establish why an offender ceased to act against the law should not prevent us from enforcing it. It is why we assert that Kant's views on punishment accord with his moral theory, at least theoretically. The next step is to determine why Kant appears to contradict himself in some of his examples of punishment.

The Role of Examples in Kant's Work

As previously mentioned, there is difficulty in giving examples of pure ethics. Kant was cognizant of that and highlighted in the *Groundings for the Metaphysics of Morals* that “worse service cannot be rendered morality than that an attempt be made to derive it from examples” (Kant, 1993, p. 20 [408]). He also states that if, sometimes, examples may be appraised for their validity, “in no way can [they] authoritatively furnish the concept of morality” (Kant, 1993, pp. 20–21 [408]). There are two main reasons why this presents itself as problematic. First, moral examples are derived from a conception of rightness and good acquired empirically based on a preexisting standard (Louden, 2011, p. 92). Second, the value Kant gives to the autonomy of the will emphasizes that we should judge a situation ourselves, following universal rules and principles, which is the opposite of what one does when simply following examples (Louden, 2011, p. 92). That said, Kant does not shy away from examples in his moral theory or to explain his views on punishment. It is essential to understand why, as it allows us to forge a little place for pragmatic morality within Kant's work.

Louden says that examples in Kant reflect his acknowledgment of the manifest human connection with empirical knowledge and are designed to guide us when human nature makes it challenging for us to act morally (Louden, 2011, p. 94). We can view Kant's use of examples from two perspectives. First, they help demonstrate that morality is humanly possible and guide those whose ability to reason autonomously is challenged or underdeveloped (such as children and people with intellectual disabilities) by providing something to emulate, a model to strive for (Louden, 2011, pp. 93–100). They are pathways for social transformations toward an ethical society. The second perspective presents examples that share similarities with real life but are different enough that they do not allow for a direct association through experience. They are not

there to tell us what we should do but to either symbolize abstract moral concepts that are not sensible to us or to expose human conditions that prevent us from achieving them.

Two examples from *The Metaphysics of Morals* will serve as our main cases in evaluating how Kant's views on punishment can be regarded as manifestations of impure ethics rather than scant embodiments of his pure morality. The first case explores examples in which honour conflicts with punishment. In Kant's views on punishment, there are various times when honour is brought up as a mitigating factor for determining the proportionality of a punishment. It is not immediately clear how such a subjective and temporal feeling can be mentioned together with fair punishment or pure Kantian morality. More specifically, we look at the example of the Scottish rebellion, where Kant states that the rebels who acted out of duty for their master were less deserving of punishment than those who rebelled for personal gains (Kant, 2017, p. 116 [6:334]). We also look at the examples of military duels and infanticides where, in all cases, Kant states that although both deserve the death penalty, none should be enacted because they are crimes driven by honour (Kant, 2017, p. 118 [6:336]). As we can see, Kant's consideration for honour appears at best incoherent with the core of his moral theory and deserves to be explored in further detail.

The second case looks at Kant's statement that if a society were to disband, it would be primordial to execute all murderers before it happens so that "each has done to him what his deeds deserve and blood guilt does not cling to the people for not having insisted upon this punishment" (Kant, 2017, p. 116 [6:333]). On one side, this statement appears to put punishment as a categorical imperative as it reflects the requirement of duty, regardless of the existence of civil society. However, it is not clear what Kant refers to as "bloodguilt" in this segment. How does not punishing a murderer for his crime equate to transferring bloodguilt to the people? It is

also worth noting that the term “bloodguilt” is used later in the book with reference to divine punitive justice, making it intriguing and deserving of further exploration.

Honour and Punishment

The importance of honour in determining the fairness of punishment in Kant’s work is interesting and seemingly controversial. Honour is a subjective concept shaped by experience. Its meaning and implication can evolve considerably over time and from one place to another. It is surprising to see Kant use it in conjunction with punishment because, as just discussed, punishment supposedly rests on a priori principles. It is even more surprising considering that Kant is generally known as someone who “[rejects] socially defined measures of human worth” (Bayefsky, 2013, p. 811) and who, in his early life at least, saw honour as undermining morality precisely because of the sway public opinion holds on it (Kant, 1997, p. 20 [27:45]; LaVaquer-Manty, 2006, p. 721). Two questions come to mind. First, what exactly was the meaning of honour for Kant, and could it have evolved between his early lectures on ethics and his late metaphysic of morals? Second, is there a way to conceptualize honour and punishment without contradicting Kant’s moral theory? Using Louden’s (2000, 2011) work on impure ethics as a reference, it will be demonstrated that the concept of honour reflects a pragmatic morality that was increasingly present in Kant’s mature works, although never put into a formal system. Kant’s examples of honour do not invalidate the pure moral claim on which punishment is based but show an imperfect roadmap to it rather than an embodiment of it. Two examples of honour are particularly predominant in *The Metaphysics of Morals* and deserve a closer examination.

The first case consists of a failed Scottish rebellion. Kant explains the problem as follows:

Suppose that some (such as Balmerino and others) who took part in the recent Scottish rebellion believed that by their uprising they were only performing a duty they owed to House of Stuart, while others on the contrary were out for their private interests; and suppose that the judgment pronounced by the highest court had been that each is free to make the choice between death and convict labor. I say that in this case the man of honor would choose death, and the scoundrel convict labor. This comes along with the nature of the human mind; for the man of honor is acquainted with something that he values even more highly than life, namely honor, while the scoundrel considers it better to live in shame than not to live at all. Since *the man of honor is undeniably less deserving of punishment* [emphasis added] than the other, both would be punished quite proportionately if all alike were sentenced to death; the man of honor would be punished mildly in terms of his sensibilities and the scoundrel severely in terms of his. (2017, p. 116)

This case reflects very well the “desert thesis” mentioned earlier by Wood (2008) that Kant believed criminals should be punished because they deserved it. It also demonstrates that finding a proper punishment is not strictly related to the moral blameworthiness of the crime itself but also considers the criminal’s own moral worth because Kant states that those with honour are less deserving of punishment even if they did commit the same crime as the “scoundrels” (2017, p. 116). At its source, this is consistent with Kant’s moral theory through the idea of achieving the highest good, as everyone is punished based on his moral worthiness. However, this example demonstrates two other things. First, Kant presumes that for a moral person accused of a crime, preserving one’s honour has primacy over punishment, even if it means losing his life. It is something reflected in Kant’s writing as far back as his lectures on ethics, where he stated that “it

is better to sacrifice life than to forfeit morality. It is not necessary to live, but it is necessary that, so long as we live, we do so honourably” (Kant, 1997, p. 147 [27:373]). Second, it also implies that somehow, a person of honour would be less deserving of punishment than others. Kant’s belief that honour influences desert and is worth preserving at all costs makes investigating further the various meanings of honour in Kant a worthy endeavour.

Kant distinguishes between two broad kinds of honour. One is the corrupting form of honour, the one associated with social standing, which is hierarchical and is a “socially defined view of human status” (Bayefsky, 2013, p. 812). The other is what he called true honour, which implies conducting oneself in such a way as to preserve one’s dignity, which includes not letting others hold oneself in contempt (2013, p. 824-25). True honour is a way to “secure rightful honor in this life” (Held, 2010, p. 77) by ensuring that each person gets his due and no injustice is left unchecked. A crime is an infringement on the dignity of the victim, and an unjust punishment is an infringement on that of the criminal. Therefore, honour is a sentiment that, although very subjective, allows society to feel that justice has been served and dignity maintained, and the proportionality so important to retribution allows for this to happen. Kant asserts that to punish disproportionately or for any reason other than to preserve justice would lead to a degradation of honour, hence dignity, in favour of following one’s inclinations (Held, 2010, p. 91). It reflects that for Kant, “it is the shamefulness of vice, not its harmfulness, that must be emphasized above all, and of course, the dignity of virtue” (Kant, 2017, p. 243 [6:483]; Welsh, 2008, p. 145). It explains why in the example, Kant says the men of honour would likely choose death over labour, as it is the most proportional punishment to the crime committed (Kant did state that rebellion against the State was punishable by nothing less than death [2017, p. 105 [6:320]), and allows them to restore their dignity, while the scoundrels with no honour would follow their

inclination and go for the lesser punishment.

Arguably, honour is subjective, derived from experience, and not as popular in today's world as in Kant's. However, the point Kant is trying to demonstrate is the importance of true honour as a concept not grounded in pure morality but one that can lead us closer to it. Although this might explain the importance of valuing honour over the severity of the punishment, however, it does not explain why Kant explicitly stated that "the man of honour is undeniably less deserving of punishment" (Kant, 2017, p. 116 [6:334]). For this, it is useful to look at two other prominent examples of murder: the military duel and the case of infanticide (Kant, 2017, pp. 118–119 [6:336]).

In the second case found in a later section of *The Metaphysics of Morals*, Kant talks about the inadequacy of imposing the death penalty for crimes of military duels and infanticide. In both cases, Kant says that although the murderer deserves death, it is doubtful that a legislated death penalty would be appropriate because the feeling of honour is what led to both murders (2017, p. 118 [6:336]). The idea reflected here is not that one is justified to do what it takes to rectify a perceived affront to his honour. People are still required to comply with this world's positive and moral laws, even if they feel offended. However, what is insinuated is that in some cases, breaking a positive law is not always a manifestation of free will (which is necessary for guilt), so these people should not be punished as severely. It is part of the reason people feel that a beaten woman who killed her abuser is less deserving of punishment than a contracted killer or that the poor man who steals to provide basic commodities for his family is less deserving of punishment than the rich one who steals out of greed. It is not out of free will but of necessity that both crimes were committed. They are still blameworthy but feel less deserving of punishment because they were committed to restoring a fundamental right to dignity and

preserving their honour.

It is certain that what consists of an infringement of dignity can change over time and that arguing that the crime of infanticide and military duel should not be punished as regular murders is certainly more controversial today than in Kant's time. Most people today would disagree that murdering an illegitimate child is different from killing any other child and that being called a coward is a justification for duelling to the death. Kant states his reasoning for keeping these two cases outside of the realm of legislative punishment:

Legislation cannot remove the disgrace of an illegitimate birth any more than it can wipe away the stain of suspicion of cowardice from a subordinate officer who fails to respond to a humiliating affront with a force of his own rising above fear of death. So, it seems that in these two cases human beings find themselves in the state of nature, and that these acts of killing, which would then not even have to be called murder, are certainly punishable but cannot be punished with death by the supreme power. (Kant, 2017, p. 118 [6:336])

Kant established clearly that people seek to restore their honour because society would hold them in contempt, and they would be permanently disgraced from it for not doing so (Bayefsky, 2013, p. 828). According to the moral standards of the time, not reacting to a military duel challenge or disposing of a child from an illegitimate birth would damn the person concerned to a life as an outcast, and they would not be able to live with dignity for the rest of their lives (LaVaquer-Manty, 2006, p. 729; Livingston & Soroko, 2007, p. 496). Therefore, the social conditions and positive laws of the time did not provide the external conditions of freedom required to make moral choices.

As Livingston and Soroko argue, Kant was very well aware that 18th century Prussia did

not have the perfect external condition for ethics and that the “social evolution of moral evaluation” (2007, p. 494) had to change. For Kant, traditions such as duelling or the social rebuke of mothers of illegitimate children were practices and signs of a still very “underdeveloped and immature public” (Livingston & Soroko, 2007, p. 495). It is also part of the reason why Kant makes an exception on capital punishment for military duel and infanticide crimes. It is a recognition by Kant that public opinion, although a bad thing for ethics, nevertheless has a more substantial impact on people than positive law can sometimes achieve. To be clear, Kant sees infanticide and duelling as a regression to a state of nature (Livingston & Soroko, 2007, p. 495). It is not condoned as ethical in any way. It is rather seen as a temporary regression to a state of nature, and honour, as an early manifestation of people’s rational potential by which we develop “increasingly complex and abstract moral self-understanding” (Livingston & Soroko, 2007, p. 497). In other words, honour is a temporary pragmatic moral application until humans mature enough to take in pure moral concepts out of their own free will.

Although this new perspective on honour makes both examples more comprehensible, one difficulty remains with the case of infanticide. Where a duel requires consent from both parties and could be considered a fair trial based on honour (as discussed above), there is reason to question where justice lies from the perspective of the child murdered by his mother. It is one thing to choose honour over life regarding one’s own life. It is another to take someone else’s life to preserve one’s own honour. Kant’s explanation is rather bleak. Because the child was born outside of marriage, he was also born outside of the law and its protection. Therefore, his murder is not a murder but a killing within a state of nature, and his annihilation can be ignored (Kant, 2017, p. 118 [6:336]). Kant’s explanation is surprising and is naturally in tension with our modern acknowledgment of fundamental human rights. Therefore, the meaning of the state of

nature and its implication on the mother's punishment deserves further inquiry.

Kant's state of nature is similar to other modern philosophical views. Briefly, it is a state of complete freedom where no higher authority can make and enforce laws (Kant, 1983, p. 115 [354]). When humans choose to live in society, they give up on some of that freedom (such as the right to kill) in exchange for a legal system that gives them essential protection and social order (Uleman, 2000, p. 185). However, if the system of laws is ineffective at protecting them, people can sometimes live in a situation that emulates the state of nature. Timmermann (2022) argues that it is likely what Kant is referring to in this example. Because the unwed mother and the illegitimate child have reverted to a partial state of nature (because the positive laws cannot adequately protect them from social ostracism and cruel treatments), "it would be within the mother's rights to defend herself against the threat posed by the child, which would not be entitled to the law's protection at all" (Timmermann, 2022, p. 14). Sussman also argues that this partial state of nature is why Kant argues for a lesser punishment (than death) and not a complete absolution for her crime (2008, p. 308). Otherwise, in a complete state of nature, the crime of murder would have been outright unpunishable through juridical laws (Sussman, 2008, p. 308). As for the child, by saying that his annihilation can be ignored, Kant only highlights the problem of the child who finds himself in this partial state of nature. It is not to be interpreted as an endorsement of the mother's action. The crime of murder must still be punished, but the partial state of nature in which the unwed mother finds herself favours lenience over proportional retribution, and it does so for two reasons.

First, social norms and laws were partly to blame for the mother's actions (Timmermann, 2022, p. 20). At the time, a woman's dignity and life's prospect was mostly acquired through marriage. Marriage meant respectability, legitimacy, and viable life chances for the mother and

child. They had the legal rights to the fathers' protection and resources and the social benefits of family life (Uleman, 2000, p. 189). Without it, the unwed mother had no economic support or legal protection from social ostracism. As well, 18th-century Prussia was a time when honour was very present in people's lives and pressured upon them early (Welsh, 2008, p. 139). Even without committing any crime, being an unwed mother was considered dishonourable, and humiliation and harsh treatment were prevalent. For example, unwed mothers were excommunicated until they repented and often flogged by parents or employers (Uleman, 2000, p. 178). Local authorities also subjected them to public shame and ridicule, as well as their family, whose means of subsistence could become threatened (Timmermann, 2022, p. 22).

Second, for the mother committing infanticide, the death penalty was also disproportionately cruel in contrast with that of other murders and, therefore, unjust towards the mother (Timmermann, 2022, p. 22). It did not respect the mother's humanity; she was not respected as an end in herself. Historically, punishment for infanticide by the mother has always been harsh and barbaric. Confession was often extracted through torture (Timmermann, 2022, p. 22). Forms of punishment for women killing a family member could go from being buried alive to some ruthless form of drowning (Timmermann, 2022, p. 22). Although these particular examples were already abandoned by Kant's time, other deterrent measures remained, such as displaying the head of the executed woman or her dead and broken body on a wheel. They were also usually not buried in the consecrated ground so that they could be shamed even after death (Timmermann, 2022, p. 23). Overall, Kant's lenience with infanticide does not rest on a lack of desert or legal standing to punish the mother, but on the lack of proportional retribution in punishment and on the archaic social norms and laws of the time that made chastity "*the point d'honneur* of unmarried women" (Timmermann, 2022, p. 21). Going back to the initial inquiry

of where the justice lies for the child whose been murdered, the answer is that justice is simply not served, and for his murder, the mother deserves death. Our main argument remains despite Kant's reference to the state of nature or the importance of honour in punishment. Kant's examples do not always exemplify Kant's pure morality at work and sometimes, as in this case, reflect attempts at a pragmatic morality that accounts for the reality of the time.

It is also worth noting that these are not isolated examples of pragmatic morality in Kant's work. Allowing a ruler to punish by other means than capital punishment should the criminals be so many that it would deprive the state of its population and threaten its existence is also an example of earthly consideration (Kant, 2017, p. 117 [6:334]). On the right of necessity, Kant also shows his concerns over our inability to act against our inclinations. He alleges that had a castaway killed another one by pushing him off a shipwreck to save his own life, he would undoubtedly be morally blameworthy but could not be punished for it, for the positive law (against murder) is not enough to deter someone facing otherwise a certain death (Kant, 2017, p. 31 [6:236]). In a perfect moral world, where people follow the law out of duty, no one would value their own life over another. However, Kant knows that this is a lot to ask, and the right of necessity is an acknowledgment of it. It does not refute the idea of Kantian morality but reflects a pragmatic moral thought that is often denied to Kant.

Punishment and the Fulfillment of Justice

The second case examines how Kant sometimes uses examples to symbolize moral concepts or expose how our inherent human conditions get in the way of morality. To help us in this endeavour, let us recall one of Kant's famous examples of the necessity to impose capital punishment for the crime of murder. In this example, Kant proclaims that

Even if a civil society were to be dissolved by the consent of all its members (e.g., if a people inhabiting an island decided to separate and disperse throughout the world), the last murderer remaining in prison would first have to be executed, so that each has done to him what his deeds deserve and blood guilt does not cling to the people for not having insisted upon this punishment; for otherwise the people can be regarded as collaborators in this public violation of justice. (2017, p. 116 [6:333])

This case leaves many questions. It appears to put punishment as a categorical imperative as it reflects the requirement of duty, something that must be done regardless of the existence of civil society. However, if a society were to disband, its inherent condition of right would be ended, and there would no longer be a requirement for deterrence. What would then be the point of executing the last murderers? Furthermore, why does Kant say the failure to execute them constitutes a moral lapse for “the people” (O’Connell as cited in Thomason, 2021, p. 80)? If this example appears confusing, it is because we focus on judging its value from a utilitarian perspective. We wonder, for example, why the state should kill the murderer before the collapse of society because we are concerned that the result of our action serves certain ends or prove the purity of our intention (footnote in Kant, 2018, p. 42). Instead, we must look at this example as Kant’s attempt to make an abstract moral concept sensible to us. Through it, Kant wants to symbolize the demandingness of justice (Thomason, 2021, p. 81).

According to Kant, justice, holiness, and goodness form “the moral concept of God” (Kant, 2018, p. 21 footnote). Kant observed that if we do not cry injustice when the good we believe we deserve does not come our way, we rejoice and call it justice when some ill befall the wicked (Kant, 2018, p. 21 footnote). It is a truthful reality that crimes often remain unpunished, and Kant says that when nature gets back at the wicked, we tend to see this as a sign of God’s

work and reconcile ourselves with heaven (Kant, 2018, p. 24 footnote). Justice is a deed to be done on evil. Therefore, justice is a feeling that must be perceived as superseding mere institutions of legal punishment to encompass the punishment of vice. The imperative of killing the murderers comes from the need for justice to punish vice. Ending society without killing the murderers would not remove the feeling of injustice. It would “cling to the people” (Kant, 2017, p. 116 [6:333]) and make them guilty, not of collaboration to the crime of the murderers, but of having let vice succeed over good. Therefore, by symbolizing the demandingness of justice, Kant emphasizes that the penal system should not contradict our idea of justice and punish vice unconditionally, as it otherwise only leads to further wickedness (Thomason, 2021, p. 91).

Another possibility is that Kant’s example serves as a “psychological [description] of how human moral agents think about justice” (Thomason, 2021, p. 83). We see justice as intemporal and unconditional. It is why society pursues war criminals decades after war and why mistakes from past centuries are still perceived as deserving justice today. Thomason argues that Kant’s use of the term bloodguilt refers to the intensity of the feeling and that it is why people sometimes feel willing to “spill the blood of people who are associated with the author of the original wrong even if they themselves are innocent” (2021, p. 83). It is “how people think about justice” (2021, p. 86). From this perspective, rather than symbolizing a moral concept, Kant’s example exposes a human condition that can be detrimental to morality, similar to how the child born out of wedlock was perceived as being in a semi-state of nature. Kant’s use of the word “bloodguilt” is also worth noting because it is used in other parts of his work and has a specific religious connotation. Kant’s *Religion within the Boundaries of Mere Reason* was published 4 years before *The Metaphysics of Morals*. Therefore, the former may have influenced his views on punishment in the latter. Consequently, the last part of this paper focuses on exploring the

relation between religion and punishment.

Religion and Punishment

Religion, along with a perfect civil constitution and a good educational system, are deemed essential by Kant to the establishment of a worldwide ethical society because he believes religion can achieve through internal constraint what external laws cannot (Kant, 2006, p. 238 [333 footnotes]; Louden, 2011, pp. 59–60). In other words, it allows for a morality based on voluntary self-restraint rather than external coercion and, therefore, displays similarities with Kant’s pure morality. Consequently, and in light of Louden’s (2000) argument on impure ethics, we argue that Kant’s work on religion is akin to a pragmatic moral project which directly influenced his views on punishment in *The Metaphysics of Morals*.

First, it is necessary to have some contextual awareness of the impact of religion on Kant’s life. In reaction to the enlightenment, Kant’s time was marked by a strong revival of contemporary religions (Wood, 2020, p. 9). As a child, he received a religious education and studied at the *Collegium Fredericianum*, where he became acquainted with Pietism (Wood, 2020, p. 9). He was also the son of devout Pietist parents and lived in Prussia, where the court of Friedrich Wilhelm I was also highly influenced by the clergy (Wood, 2020, p. 9). It was a time when morality was characterized mainly by “obedience owed by created beings to their God, and the unfortunate inability of most people to behave without guidance, or indeed threats of punishment and offers of reward” (Welsh, 2008, p. 150).

Being a philosopher of the enlightenment era with a moral theory that strongly emphasized the use of reason, Kant had an early interest in finding a relation between empiricism and the metaphysical conception of God (Wood, 2020, p. 10). It would be easy to picture Kant as an atheist; however, although he criticizes ecclesiastic faith and many of the church’s religious

practices that based morality on pleasing a deity and avoiding punishment rather than for the sake of being good, nowhere in his critique does he make a firm affirmation of secular beliefs. In fact, contrary to our first assumption, he even claims that “the needs of morality demand and justify a sort of faith in the existence of God” (Adams as cited in Kant, 2018, p. viii). Religion allows us to believe that the highest good is possible, and its revelations serve as channels through which God reveals Kant’s ethics under the name of the pure religion of reason (Hare, 1997, p. 40). Kant appears particularly sympathetic to Christianity and some of its moral doctrines that he views as benefiting human society (Adams as cited in Kant, 2018, p. x). Although this does not amount to an admission of religious beliefs, it certainly shows a receptive position toward religion. Some early interpretations argued that this receptivity towards God in Kant’s system was “for the sake of Prussian censor or to appease his faithful manservant, and that he did not really mean it” (Hare, 2017, p. 70). However, this view changed over time and a more prominent place for God in Kant’s religion is now seen, even if only conceptually (Hare, 2017, p. 70).

Before examining how Kant’s work on religion relates to some of his views and examples of punishment, let us briefly look at how it generally exemplifies his concern for pragmatic morality. First, as pointed out by Loudon (2000), there are textual clues in Kant’s work suggesting a concern for pragmatic morality. For example, Loudon argues that Kant’s repeated use of the words “human being” and “human nature” in the main title and all four subsection titles of Part 1 of *Religion within the Boundaries of Mere Reason* reflects that his primary concern in it is for “the moral psychology of human beings and not that of any other kind of rational being” (Louden, 2000, p. 132).

Second, Kant says that human understanding is a “discursive understanding that has need

of images” (Kant, 2007, p. 236 [408]) to make abstract concepts sensible to people. Therefore, religion appears naturally well-suited for the task, as it can provide the images needed and find “ways to make the message of pure ethics graspable to human beings” (Louden, 2000, p. 127). To do this, however, Kant has to find a way to adapt religion within the limits of reason alone, without denying revealed religion itself, so that it becomes acceptable for the predominantly religious population and the prevailing political condition in Prussia at the time (Tönissen, 2022, p. 162).

Third, Kant says that the church as a social structure is adequate for teaching virtue (Kant, 2018, pp. 132–133 [6:101,102]). Virtue requires us to act of our own free will, so people need to be compelled by inner motivation. The church is appropriate because it inculcates voluntary compliance, enabling the teaching of virtue. In contrast, the state is inappropriate because its laws are inherently coercive, and no one can coerce people into being virtuous. Kant also finds some concepts of Christianity interesting, such as the idea of salvation, which shares similarities to Kant’s highest good and uses a similar language, emphasizing one’s worthiness rather than its actual realization (Kant, 2018, p. 160 [6:133]).

Finally, religion exemplifies Kant’s interest for a pragmatic morality through the concept of faith. People can progress as religious, moral beings only through faith in God’s existence and our possible salvation (Kant, 2018, pp. 100–101 [6:68]). Similarly, people must have faith that the highest good is possible to progress as moral beings. Faith is “rooted in our obligation to will the highest good: a systematic unity of virtue and happiness... [that is only possible] if we accept the existence of God because only He makes the highest good a possible object of our will” (Tönissen, 2022, p. 163). Therefore, religious faith provides us with the necessary mindset to develop a morality based on free will and pure reason. Now that religion has been established as

a potential pragmatic moral project, it is hoped that two central concepts impacting punishment will reinforce this claim: the origin of radical evil in humans and the concept of bloodguilt.

The inherent moral predisposition of humans in a theoretical state of nature is not a new topic in moral philosophy. Whether one believes people to be inherently good or bad has implications for a philosopher's conception of society and humanity's future. For Kant, humans have an inherent moral predisposition towards evil without being completely evil (Kant, 2018, p. 72 [6:40]). His justification is that if they were born perfectly good, from birth, they would follow the moral law out of duty and would never be prone to inclinations, which is not the case (Kant, 2018, p. 74 [6:42]). He also says humans are predisposed to morally corrupt each other, to the point that even if we were to create a moral society, it would always be in danger of relapsing into immorality (Kant, 2018, p. 126 [6:94]). However, all hope is not lost. According to Kant, people are not unalterably corrupt and have the propensity to be good. It just falls on us to work towards it or not (Kant, 2018, pp. 76–77 [6:47]). Kant says that ideally, good action should be based on objective moral grounds, something that is possible only if humans are first “habituated to morality [and] the moral feeling first made active” (Kant, 1997, p. 80 [27:287]). Until that moral feeling is active, Kant does not object to using a more subjective morality, such as punishment and rewards (Kant, 1997, p. 80 [27:287]). When punishment happens, it serves as means of “moral training” because it deters people from doing immoral acts through repetition and consistency. Kant states that through punishment, people can eventually learn to act morally out of their own free will and not out of fear of punishment anymore (Kant, 1997, p. 80 [27:287]). Therefore, the law of punishment works toward accomplishing the perfect ethico-civil society. It trains people towards pure morality and away from a predisposition towards evil. The final victory of good over evil on earth can only happen in a society strictly ruled by the moral

law, a society in “which reason makes it a task and a duty of the entire human race to establish in its full scope” (Kant, 2018, p. 126 [6:94]).

The second concept reinforcing religion as a pragmatic moral project is the concept of bloodguilt. References to bloodguilt appear in two specific passages from *The Metaphysics of Morals* and shorter passages from its drafts (Thomason, 2021, p. 82). Its uses appear unusual in a work of morals based on reason. Some authors (Murphy 1987, Cottingham 1979) point out that bloodguilt has a biblical connotation which does not seem to belong there (Thomason, 2021, p. 80). Cottingham describes one of the passages as “heavily tied up with Old Testament notions of sacrifice and placation” (Cottingham as cited in Thomason, 2021, p. 79). This section aims to demonstrate how the use of bloodguilt in *The Metaphysics of Morals* was likely not a question of style nor coincidence but to link Kant’s work on morals (including his views on punishments) with his earlier work on religion.

Bloodguilt first appears in the previously analyzed example, where the faith of imprisoned murderers on the eve of a society’s dissolution was discussed. Thomason argues that this unusual example served either to symbolize the demandingness of justice or to exemplify our human perception of it as a transcendental concept (Thomason, 2021, p. 83). Regardless of the argument, the use of the term bloodguilt is unusual. If Kant wanted to say that justice is a sentiment that transcends our mere institutions of justice and that murder deserves punishment, surely there was a better way to do it than by stating that bloodguilt would “cling to the people”(Kant, 2017, p. 116 [6:333]). The second passage of bloodguilt, appearing at the end of *The Metaphysics of Morals*, is even more intriguing and intense in its wording. It reads as follows:

Blood innocently shed cries out for vengeance. – Crime cannot remain unavenged; if punishment does not strike the criminal, then his descendants must suffer it, or if it does not befall him during his lifetime, then it must take place in a life after death, which is accepted and readily believed in expressly so that the claim of eternal justice may be settled. – I will not allow blood-guilt to come upon my land by granting pardon to an evil, murdering duelist for whom you intercede, a wise ruler once said. – Guilt for sins must be expiated, even if a complete innocent should have to offer himself to atone for it. (Kant, 2017, p. 250 [6:490])

This passage, like in the first example, refers to how we think about justice, albeit in a much more explicit manner, with the exception that in the latter, bloodguilt is also expiated through the blood of innocents, something of which Kant clearly disapproves. Otherwise, the themes of bloodguilt reflected in Kant’s two passages share similarities. First, bloodguilt is acquired through the spilling of innocent blood (murder) and, if not expiated, can be transferred and “cling” to others (Thomason, 2021, p. 83). Second, bloodguilt is a feeling that is strong enough that we feel punishing people vaguely associated with the crime seems justified, even if they have committed no crime themselves (Thomason, 2021, p. 83). These characteristics of bloodguilt in Kant’s two passages do not appear coincidental and are strikingly similar to those found in its biblical definition, although some differences exist.

According to the *Oxford Encyclopedia of the Bible and Law*, bloodguilt is “an arena in which legal and social institutions, on the one hand, and moral and religious beliefs, on the other hand, are closely related and intertwined” (Bartor, 2015, para. 11). Bloodguilt entails the killing of an innocent (shedding innocent blood) and is considered a crime against God which can be atoned only through the death of the murderer (shedding the guilty blood) and under the

principle of retribution (Bartor, 2015, para. 11). Furthermore, blood in biblical Hebrew has a few meanings. First, aside from the substance, blood refers to an improper state of being spilled (Kedar-Kopfstein as cited in Bartor, 2015, para. 4). Second, blood is seen both as a contaminant and a cleanser. Killing an innocent contaminates the earth with innocent blood while spilling the blood of the guilty helps in its cleansing (Bartor, 2015, para. 15). Finally, blood reflects a person's moral character and has an existence of its own, which can move from one person to another (Abusch as cited in Bartor, 2015, para. 5). It can also cling to an object such as a house or land or contaminate a community, which is why in some communities, ritual slaughtering of a heifer was often conducted when a criminal was unknown, in an attempt to purify or relocate the contaminated blood elsewhere (Bartor, 2015, para. 17).

The biblical understanding of the term bloodguilt leaves little doubt that it is what Kant was alluding to in *The Metaphysics of Morals*. Bloodguilt assumes a murder has been committed and its perpetrator bears responsibility. It also assumes that it can be transferred from one person to the people if the crime remains unpunished, reflecting that murder is not only a judicial crime but also a moral one as it transcends the institutions of justice. However, the idea that innocents should suffer if the criminal cannot is different. Kant clearly established that before any punishment can occur, the accused must have been found guilty (Kant, 2017, p. 114 [6:331]). If the second passage of bloodguilt describes punishing innocents, it is, Thomason argues, because Kant attributes this view to that of the theologians (2021, p. 82). Thomason supports his claim with a passage from *The Metaphysics of Morals* drafts in which Kant states that

the theologians have found it so inappropriate that crimes should go unpunished that they prefer to assume that an innocent could assume them (for others) simply in order to

satisfy justice or that children must atone for their parents' guilt. (Kant as cited in Thomason, 2021, p. 82)

In other words, Kant's passage from *The Metaphysics of Morals* about punishing innocents is not his own thoughts but a recapitulation of the theologian's perspective, which he condemns. Although bloodguilt relates to religion, Kant remodels its interpretation to fit his views on punishment and moral theory. Bloodguilt becomes an intangible feeling of justice, but one that must confine its punishment within the boundaries of our institutions of justice because the theological interpretation would make it impossible to punish vice without doing others wrong (Kant, 2017, p. 250 [6:490]). The demand that bloodguilt puts on us helps portray retribution as an imperative and religion as a pragmatic moral application of Kant's moral theory.

Conclusion

When we look at the compatibility between Kant's views on punishment and his moral theory, we arrive at a conclusion that appears very close to Byrd's (1989) assessment of it. Punishment as a law can act as a deterrent but must be retributive in its execution (Byrd, 1989, p.183). However, where Byrd sees this as conflicting with Kant's moral philosophy, it was argued that the two complement each other. In a civic society, every act of legal punishment inevitably creates deterrence. It is the maxim behind the act that must convey its morality by fulfilling the requirement of justice. It must be enacted and executed strictly because a crime has been committed and must be retributive in its form to ensure justice is met through proportionality and respect for the criminals as ends in themselves. Only from this perspective can the all-encompassing requirement of morality be constantly reaffirmed. Nevertheless, a consensus on the compatibility between Kant's views on punishment and his moral theory is unlikely, and his use of apparently dubious or flawed examples of punishment certainly does not

help.

This paper hopes to have demonstrated that although there are apparent shortcomings in Kant's examples, they should not be seen as evidence of a clear incompatibility between retributive punishment and his moral theory. In fact, many questions appear to dissipate with an understanding of the role of examples in Kant's moral philosophy. At the outset, empirical examples are to be taken lightly because they can never perfectly epitomize pure moral principles, as one can only assess their validity from the ground of human experience (Louden, 2011, p. 99). Consequently, they are rarely meant to demonstrate pure moral laws in action and are instead more often used to establish what Kant sees as obstacles to pure morality (such as the human condition) or to symbolize concepts that are hardly sensible to us (Louden, 2011, p. 55; Thomason, 2021, p. 81). More importantly, Kant's examples of punishment reveal a pragmatic moral side to Kant that, although still modest, is increasingly more acknowledged in the scholarly literature. They embody what Louden calls Kant's impure ethics (Louden, 2000). They reflect a sort of imperfect roadmap to pure morality, a pragmatic morality that was increasingly present in Kant's mature works, although never put into a formal system. His examples of honour, such as military duels and infanticide, are a testament to it, and his suggested amendments to the punishment for these crimes reflect Kant's pragmatic considerations. Similarly, his work on religion is also related to Kantian pragmatic morality, and many references to it can be seen through the ideas and vocabulary used in *The Metaphysics of Morals*. For Kant, religion is a channel through which a pure morality can grow, by providing the images, concepts and institutions necessary to its development, so long as it is a morality based on reason and not blind obedience to deities (Welsh, 2008, p. 150). It is a pragmatic moral project which directly influenced his views on punishment in *The Metaphysics of Morals*.

Finally, as Kant's moral philosophy officially rests strictly on pure reason, our goal as autonomous human beings is to aim for it. However, it is often easier said than done, and it was even more likely the case in 18th-century Prussia when the masses were uneducated and unable to think and act out of their own free will. Kant's life as a professor and philosopher leaves no other choice but to believe he was aware of humans' limitations affecting their inclinations or capacity to apply pure moral concepts. Punishment was then seen as means of moral training through repetition and consistency, which, Kant asserts, would eventually make people act morally out of their own free will and not out of fear of punishment (Kant, 1997, p. 80 [27:287]). For all these reasons, Kant's views and examples of punishment are not at odds with his moral theory grounded in pure reason but instead reflect a pragmatic morality that was increasingly present in Kant's mature works, even though it was never put into a formal system.

References

- Bartor, A. (2015). Bloodguilt and blood feud. In B. A. Strawn (Ed.), *The Oxford encyclopedia of the Bible and law*. Oxford University Press.
<https://www.oxfordreference.com/view/10.1093/acref:obso/9780199843305.001.0001/acref-9780199843305-e-10>
- Bayefsky, R. (2013). Dignity, honour, and human rights: Kant's perspective. *Political Theory*, 41(6), 809–837. <https://doi.org/10.1177/0090591713499762>
- Beiner, R. (2013). Le “textualisme”: Une anti-méthodologie [“Textualism”: An anti-methodology] (D. O’Toole, Trans.). In D. Karmis, D. Lamoureux, D. Giroux, & J. Maclure (Eds.), *Ceci n’est pas une idée politique* (pp. 29–34). Presses de l’Université Laval.
- Brooks, T. (2003). Kant's theory of punishment. *Utilitas*, 15(2), 206–224.
<https://doi.org/10.1017/S0953820800003952>
- Byrd, B. S. (1989). Kant's theory of punishment: Deterrence in its threat, retribution in its execution. *Law and Philosophy*, 8(2), 151–200. <https://doi.org/10.1007/BF00160010>
- Fleischacker, S. (1988). Kant's theory of punishment. *Kant-Studien*, 79(1–4), 434–449.
<https://doi.org/10.1515/kant.1988.79.1-4.434>
- Goodwin, G. P., & Gromet, D. M. (2014). Punishment. *WIREs Cognitive Science*, 5(5), 561–572.
<https://doi.org/10.1002/wcs.1301>
- Hare, J. E. (1997). *The moral gap: Kantian ethics, human limits, and god's assistance*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198269571.001.0001>
- Hare, J. E. (2017). Kant and theology. *Toronto Journal of Theology*, 33(1), 69–73.
<https://doi.org/10.3138/tjt.2016-0015>

Held, J. M. (2010). Honor, dignity, and the summum bonum: Kant's retributivism in context.

Vera Lex, 11(1–2), 81–106.

Hill, T. E., Jr. (1999). Kant on wrongdoing, desert, and punishment. *Law and Philosophy*, 18(4),

407–441. <https://doi.org/10.2307/3505232>

Kant, I. (1983). *Perpetual peace and other essays* (T. Humphrey, Trans.). Hackett Publishing Company.

Kant, I. (1993). *Grounding for the metaphysics of morals* (J. W. Ellington, Trans.; 3rd ed.). Hackett Publishing Company.

Kant, I. (1997). *Lectures on ethics* (P. Heath & J. B. Schneewind, Eds.; P. Heath, Trans.). Cambridge University Press.

Kant, I. (2002). *Critique of practical reason* (W. S. Pluhar, Trans.). Hackett Publishing Company.

Kant, I. (2006). *Anthropology from a pragmatic point of view* (R. B. Loudon, Ed.). Cambridge University Press.

Kant, I. (2007). *Critique of judgement* (N. Walker, Ed.). Oxford University Press.

Kant, I. (2017). *The metaphysics of morals* (L. Denis, Ed.; M. Gregor, Trans.). Cambridge University Press.

Kant, I. (2018). Religion within the boundaries of mere reason. In A. Wood & G. di Giovanni (Eds.), *Religion within the boundaries of mere reason*. Cambridge University Press.

LaVaque-Manty, M. (2006). Dueling for equality: Masculine honor and the modern politics of dignity. *Political Theory*, 34(6), 715–740. <https://doi.org/10.1177/0090591706291727>

- Livingston, A., & Soroko, L. (2007). From honor to dignity and back again: Remarks on Lavaque-Manty's "Dueling for equality." *Political Theory*, 35(4), 494–501.
<https://doi.org/10.1177/0090591707302196>
- Louden, R. B. (2000). *Kant's impure ethics*. Oxford University Press.
- Louden, R. B. (2011). *Kant's human being: Essays on his theory of human nature*. Oxford University Press. <https://doi.org/10.1093/acprof:osobl/9780199768714.001.0001>
- Muchnik, P. (2013). Reflections on Robert Louden's Kant's human being: Essays on his theory of human nature. *Kantian Review*, 18(3), 461-471.
<https://doi.org/10.1017/S1369415413000198>
- O'Connell, E. (2014). Kantian moral retributivism: Punishment, suffering, and the highest good: Kantian moral retributivism. *The Southern Journal of Philosophy*, 52(4), 477–495.
<https://doi.org/10.1111/sjp.12090>
- Pincoffs, E. L. (1966). *The rationale of legal punishment*. Humanities Press.
- Ripstein, A. (2009). *Force and freedom: Kant's legal and political philosophy*. Harvard University Press.
- Sussman, D. (2008). Shame and punishment in Kant's "Doctrine of right." *The Philosophical Quarterly*, 58(231), 299–317. <https://doi.org/10.1111/j.1467-9213.2007.530.x>
- Thomason, K. K. (2021). The symbol of justice: Bloodguilt in Kant. *Kantian Review*, 26(1), 79–97. <https://doi.org/10.1017/S1369415420000345>
- Timmermann, J. (2022). The quandary of infanticide in Kant's 'Doctrine of right.' *Archiv für Geschichte der Philosophie*. <https://doi.org/10.1515/agph-2021-0034>

Tönissen, B. (2022). Allen W. Wood: Kant and religion. Cambridge University Press, 2020. 270
Seiten. ISBN 978-1-10-842234-5. *Kant - Studien*, 113(1), 143–173.

<https://doi.org/10.1515/kant-2021-0023>

Uleman, J. K. (2000). On Kant, infanticide, and finding oneself in a state of nature. *Zeitschrift
Für Philosophische Forschung*, 54(2), 173–195.

Welsh, A. (2008). *Kant's engagement with honor*. Yale University Press.

Wood, A. (2008). *Kantian ethics*. Cambridge University Press.

Wood, A. W. (2020). *Kant and religion*. Cambridge University Press.