ANTI-CODIFIABILITY IN NORMATIVE ETHICS

A SENIOR HONORS THESIS

SUBMITTED TO

THE MORRISSEY COLLEGE OF ARTS AND SCIENCES

DEPARTMENT OF PHILOSOPHY

BY

THOMAS CARRABES

ADVISED BY PROFESSOR MICAH E. LOTT

MAY 2024

Acknowledgements

To my family, I am forever grateful for your unwavering love and support throughout my time at Boston College, which has culminated in this project. With all my love, thank you.

To my advisor, Professor Micah Lott, thank you for your generosity and guidance throughout this endeavor. This thesis would not have been possible without your expertise. It was an honor to pursue this project under your counsel.

To Professor Marius Stan, thank you for your investment in the success of the members of the honors cohort. Your kindness and support made this process incredibly rewarding.

To Professor Cherie McGill, thank you for igniting my passion for normative ethics, which inspired the concentration of this thesis. The four consecutive semesters I spent in your classroom transformed my perspective as a student of philosophy and person.

To my fellow cohort members and classmates at Boston College, it was a privilege to have been surrounded by such amazing peers. Thank you for being a constant and continued source of inspiration.

Contents

Introduction

> "All people possess within them a moral sense that cannot bear the suffering of others. The former kings had such a moral sense and thus they devised means of government that would not allow people to suffer. If a ruler were to employ the moral sense that makes human suffering unendurable in order to implement such humane government, he would find bringing the entire world into order to be simple, as though he were turning the world in his hand. Why do I say that all people possess within them a moral sense that cannot bear the suffering of others? Well, imagine now a person who, all of a sudden, sees a small child on the verge of falling down into a well. Any such person would experience a sudden sense of fright and dismay. This feeling would not be something he summoned up in order to establish good relations with the child's parents. He would not purposefully feel this way in order to win the praise of their friends and neighbors. Nor would he feel this way because the screams of the child would be unpleasant. By imagining this situation we can see that one who lacked a sense of dismayed commiseration in such a case simply could not be a person. Moreover, anyone who lacks the sense of shame cannot be a person; anyone who lacks a sense of deference cannot not be a person; anyone who lacks a sense of right and wrong cannot not be a person. The sense of commiseration is the seed of humanity, the sense of shame is the seed of righteousness, the sense of deference is the seed of ritual, and the sense of right and wrong is the seed of wisdom. Everyone possesses these four moral senses just as they possess their four limbs" (Mencius 2A6).

Ancient Chinese philosopher Mencius regarded moral intuition as the defining feature of humanity through which all social and political customs are to be realized. This belief that the inherently altruistic hearts and minds of humanity offer sufficient direction in the development of an ethical society is liberating; if humans shared a universal sense of morality and aversion to the suffering of others, ethical action would manifest naturally.

Unfortunately, history has disproven the universal benevolence of humanity, and suggests moral intuition is alone incapable of informing moral behavior. Upon this recognition, modern ethicists reject the notion that morality is informed by human instinct, and instead appeal to

robust frameworks through which ethics can be quantified. John Stuart Mill's assessment of the nature of ethics is as follows:

> "The difficulty is not avoided by having recourse to the popular theory of a natural faculty, a sense or instinct, informing us of right and wrong. For—besides that the existence of such a moral instinct is itself one of the matters in dispute—those believers in it who have any pretensions to philosophy, have been obliged to abandon the idea that it discerns what is right or wrong in the particular case in hand, as our other senses discern the sight or sound actually present. Our moral faculty, according to all those of its interpreters who are entitled to the name of thinkers, supplies us only with the general principles of moral judgments; it is a branch of our reason, not of our sensitive faculty; and must be looked to for the abstract doctrines of morality, not for perception of it in the concrete. The intuitive, no less than what may be termed the inductive, school of ethics, insists on the necessity of general laws. They both agree that the morality of an individual action is not a question of direct perception, but of the application of a law to an individual case. They recognise also, to a great extent, the same moral laws; but differ as to their evidence, and the source from which they derive their authority. According to the one opinion, the principles of morals are evident à priori, requiring nothing to command assent, except that the meaning of the terms be understood. According to the other doctrine, right and wrong, as well as truth and falsehood, are questions of observation and experience. But both hold equally that morality must be deduced from principles; and the intuitive school affirm as strongly as the inductive, that there is a science of morals. Yet they seldom attempt to make out a list of the à priori principles which are to serve as the premises of the science; still more rarely do they make any effort to reduce those various principles to one first principle, or common ground of obligation. They either assume the ordinary precepts of morals as of à priori authority, or they lay down as the common groundwork of those maxims, some generality much less obviously authoritative than the maxims themselves, and which has never succeeded in gaining popular acceptance. Yet to support their pretensions there ought either to be some one fundamental principle or law, at the root of all morality, or if there be several, there should be a determinate order of precedence among them; and the one principle, or the rule for deciding between the various principles when they conflict, ought to be self-evident" (Mill 6-8).

By the nineteenth century, ethicists widely recognized the distinct advantages of systematizing normative ethics. Transcending the apparent shortcomings of moral instinct allowed thinkers such as Mill to develop universalized frameworks through which ethical behavior can be determined systematically without the troublesome variable of human instinct.

Jeremy Bentham's Felicific Calculus is the most thoroughly developed attempt at creating a codified formulation of morality. Bentham posits that the ultimate utility, or "U", of an action can be determined using the provided mathematical algorithm (Bentham).

$$U = \sum_{i=1}^{P_{total}} \int_{t=0}^{\infty} I_i(t)\ dt$$

This equation is intended to calculate the units of pleasure and pain of any given action, which are respectively denoted as "posend" (*positive end result*) and "negend" (*negative end result*). The algorithm operates according to the following seven variables of pleasures and pains to calculate the ultimate utility of an action (Sinn).

1. *Intensity*: How strong is the pleasure?
2. *Duration*: How long will the pleasure last?
3. *Certainty/Uncertainty*: How likely or unlikely is it that the pleasure will occur?
4. *Propinquity*: How soon will the pleasure happen?
5. *Fecundity*: The probability that the action will produce other pleasures.
6. *Purity*: The likelihood that the action will not cause pain.
7. *Extent*: How many people will be affected?

The net posend/negend of an action can be determined by the process below (Sinn).

1. Consider a conscious being most immediately to be affected by an action. Rate the following for this being on a scale of 1 to 10:
    - Each different pleasure that appears to be produced by it in the first instance.
    - Each pain that seems to be produced by it in the first instance.
    - Each pleasure that appears to be produced after the first. The sum constitutes the fecundity of the first pleasure and the impurity of the first pain.

- Each pain produced by it after the first. The sum constitutes the fecundity of the first pain and the impurity of the first pleasure.

2. Repeat the process for each conscious being impacted.
3. Sum the posends for everyone and subtract the negends for everyone.
4. The act is a net good for the community if the total exceeds 0 (i.e., posends exceed negends). It's a net evil if the result is below 0.

I admire Bentham's efforts to provide an algorithmic procedure through which moral decisions can be made with absolute certainty. However, as Peter Tsu Shiu-Hwa recognizes:

> "Although the unification project has achieved considerable success in science, it does not seem to do well in the field of ethics. It is a notorious fact that no moral theories have ever gained the same kind of popularity and respect as scientific theories" (Tsu Shiu-Hwa 148).

I believe Bentham's Felicific Calculus fails to meet the intended standard of scientific soundness and validity. This model depends on the subjective evaluation of pleasures and pains, which effectively undermines the universal reliability of its calculations. As there is not a codified procedure for identifying pleasures and pains, and where they objectively belong on the scale of 1 to 10, the foundation of this algorithm is flawed in its dependence on unscientific assessments of positive and negative phenomena. Thus, despite its robust framework, I believe Bentham's Felicific Calculus fails to offer a universally objective and valid formula for ethical action.

Thesis Statement

I believe this apparent failure is to be attributed to the nature of ethics, rather than the shortcomings of Bentham. As I previously stated, I appreciate and understand the motivation

behind Bentham's commitment to the codification of morality, and am impressed by the depth of the framework he developed. However, I believe the complexity of morality will inevitably outmaneuver rigid models of action guidance dependent on universalized principles, evoking the need to rethink the dominant epistemological vision of morality and revise the accepted standard of action guidance in normative ethics.

I argue that the insensitivity of universalized moral laws and rigid principles of action guidance to the contextual nuances of moral dilemmas suggests ethics cannot be codified. Without appeal to a non-principle entity, the strict frameworks of consequentialist and deontological theories are prone to necessitate immoral action, therefore, normative ethics should embrace the imperative roles of virtuous character, contextual literacy, and practical discernment in the determination of moral action.

Structural Overview

In advance of my argument, I will address two red herrings surrounding the theory of anti-codifiability. I intend to preemptively dismiss potential concerns surrounding universal moral truths and moral relativism that are commonly, but inappropriately, leveraged against claims of anti-codifiability.

The first measure to substantiate my theory is to expound the structure of the anti-codifiability argument. I will provide a comprehensive overview of the theory with reference to Peter Tsu's *Can Morality be Codified?*. Upon articulating how imposing a single, rigid ethical code fails to account for non-principle entities and the necessary practical discernment required in moral considerations, I will introduce Rosalind Hursthouse's Neo-Aristotelian model of Virtue

Ethics as an estimable demonstration of the strengths of approaching normative ethics with an appreciation for anti-codifiability.

Subsequently, I aim to augment the theoretical model of anti-codifiability by demonstrating the flaws of codification in practice through criticisms of John Stuart Mill's Utilitarianism and Immanuel Kant's Categorical Imperative. I will challenge the doctrinaire of the theories to illustrate the dangers of relying on universalized laws and ultimate principles of ethics to dictate moral action. Beginning with Utilitarianism, I will use examples of moral dilemmas to demonstrate how action guidance according to the principle of utility does not adequately account for the contextual complexities of ethically challenging situations. In response to Kantianism, I will criticize the universalization of moral laws in Kant's Categorical Imperative. I believe there are examples which clearly reveal how universally codified laws dictate highly immoral action. My refutation of both theories will include evidence to suggest that there are circumstances in which each model prescribes action, according to their universalized principles, that is evidently immoral.

If this critique succeeds, I will then appeal to the indispensable role of non-principle entities in the supposedly formulaic processes of identifying moral action, which further supports the anti-codifiability argument. I will acknowledge and respond to criticisms against Virtue Ethics and anti-codifiability. I intend to dialogically engage the action-guiding objection presented by Joshua Duclos in order to demonstrate the robustness of Hursthouse's Neo-Aristotelian model of Virtue Ethics and its framework of indirect action guidance which I believe should serve as the paradigm of action guidance in modern normative ethics.

To close my argument, I will argue the structure of Virtue Ethics is better suited to guide moral action than its codified counterparts due to its grounding in the established virtues and cultivation of necessary moral wisdom and contextual discernment.

Preliminary Acknowledgments

Prior to presenting my defense, I will establish a few preliminary qualifications to elucidate the scope of my argument and avert misguided criticisms of my intended model of anti-codifiability. First, the term "anti-codifiability" itself is somewhat esoteric, and comparatively used in relevant philosophical literature. Nonetheless, I believe the term anti-codifiability best encompasses the essence of my argument, which is not quite congruent with related theories such as moral particularism, and therefore, what I will use as the comprehensive denomination of my argued theory going forward.

Additionally, criticisms of the anti-codifiability argument are often directed more specifically towards Virtue Ethics. Therefore, I will use the theoretical structure of Rosalind Hursthouse's Neo-Aristotelian Virtue Ethics to respond to criticisms of anti-codifiability and demonstrate the soundness of theories that do not depend on universal codes of action.

Finally, before proceeding any further, I want to note that throughout this paper, I use examples of ethically challenging situations to explicate my argument, some of which involve unpleasant details. I sincerely hope these references are evidently requisite given the subject matter, and in no way offend anyone who does me the honor of reading my thesis.

Section I: Two Red Herrings

Universal Truths

In defense of anti-codifiability, I appeal to a lack of universal moral principles and rules that can tactfully guide action. It is important to clarify that I recognize the possibility that certain specific actions are absolutely immoral, and that my rejection of the legitimacy of universal moral absolutes pertains to ethics on a normative level of action guidance. This acknowledgement should not subvert nor detract from my presented argument, rather, understanding the nature of observed moral absolutes contextualizes and I believe supports the anti-codifiability thesis.

I argue moral absolutes emerge only as the scope of ethics narrows beyond the normative level, and therefore, these truths cannot be universalized to guide action beyond their specific contexts. I credit the existence of absolute or universal moral truths, but reject the legitimacy of truly universal moral rules or principles. Working through an example is helpful in illustrating this distinction; take for instance the very common moral principle of not inflicting harm upon another human being. As a generalization, I accept the value of this statement without demur. However, while the broad scope of this claim certainly qualifies it as a potentially action guiding principle, self-defense and other similarly justified inflictions of harm render this rule far from exceptionless. Narrowing the scope of this claim to eliminate exceptions demonstrates how the essence of principle must yield to contextual specificity in order to arrive at a truly universal moral truth. Suppose one were to amend the original statement, declaring, “do not inflict harm upon others through torture”. Once again, this statement is sufficiently broad but prone to

exception; even principlist theories such as Utilitarianism would recognize the circumstantial permissibility of enhanced interrogation. In order for a moral assertion to be truly exceptionless, its absolute utility will be limited to exceedingly specific contexts, rendering it a no longer broadly actionable principle. A realistically exceptionless derivation of the aforementioned principle may be, for example, "do not inflict harm upon others through torture solely for the sake of personal enjoyment". Such a statement can only guide action under the given circumstances, which are defined by the sadistic execution of harm, not the essence of the act itself.

Therefore, when I reject the existence of universal moral absolutes, I contend moral assertions must be contextualized to the extent of compromising their ability to guide action beyond exceptionally narrow or isolated circumstances, and that absolute moral verdicts can exist only as unique truths, not universal principles.

I do want to acknowledge that a rule of thumb, such as not harming others, can be a valuable starting point for determining the morality of an action. As demonstrated by the previous depiction of how a general rule can gradually contract and inform absolute moral verdicts, I do believe that rules of thumb are valuable tools for moral philosophers, if used correctly. I will discuss the indispensable role of non-principle entities and the shortcomings of rules of thumb in a later section.

Moral Relativism

The final preliminary measure is to dismiss potential criticisms of anti-codifiability that appeal to moral relativism as the second of the two red herrings. Ethical theories that do not adhere to a strict code of laws or principles are often expected to justify their rejection of

universal principles, which is a central tenet of moral relativism. I recognize this consistency between anti-codifiability and moral relativism, and do not believe it to be problematic; I trust that my previous statement regarding universal moral principles substantiates this dismissal of such protest.

The only potentially valid criticism leveled against anti-codifiability on account of moral relativism is the argument that the theoretical model does not allow for objective moral truths, which I also believe can be swiftly discredited. Critics of Virtue Ethics, such as Joshua Duclos, argue that, because the theory does not adhere to a strict set of action guiding principles, it collapses into cultural relativism. To illustrate this claim, Duclos appeals to how the virtue of courage for the Crow, a Native American Tribe, conflicts with modern American values. Traditionally, the Crow considered actions such as "planting a coup-stick, stealing an enemy's horse, and taking an enemy's weapon while he was still alive courageous" (Duclos 11). American colonization has disturbed the traditional Crow lifestyle, effectively antiquating such acts, and disrupting their cultural notion of courage. In response to the role culture plays in determining what constitutes virtue, Duclos calls for an Archimedean point from which to assess competing cultural conceptions of morality (Duclos 13). I understand this aspiration, however, I do not believe identifying a universal standard for each virtue is a necessary step to avoid cultural relativism. Virtues, such as courage, do not need to be defined by some universal quality, as their manifestations can be determined to be objectively moral or immoral based on the context of the situation. The cultural and contextual sensitivity of anti-codifiability and Virtue Ethics does not infringe upon the ability to determine the objective moral value of a specific situation or action. For example, the perverted sense of courage behind suicide bombings and other forms of terroristic violence inflicted upon innocent civilians by extremist organizations

can be unequivocally condemned under Virtue Ethics, regardless of the cultural justifications for the act. Therefore, the absence of an established universal standard of courage, or any virtue, does not relegate anti-codifiability or Virtue Ethics to a form of moral relativism.

Section II: Anti-Codifiability

The "Trolley Problem" is perhaps the most widely recognized and referenced moral thought experiment that effectively illustrates the operations of leading codified ethical theories. The original model supposes five people are restrained on the tracks ahead of a trolley which will kill them, unless a bystander, the subject of the thought experiment, pulls a lever, diverting the trolley from its original course, killing one person to save the lives of the other five (Foot). Since popularized by Philippa Foot in 1967, this moral exercise has facilitated significant discourse among philosophers and inspired introspection within participants regarding their individual ethical outlooks. When facing dilemmas of this nature, it is easy to become overwhelmed by the gravity of the situation, and label it an impossible decision. However, codified theories such as the Categorical Imperative and Utilitarianism can provide clear and decisive verdicts to this notoriously challenging dilemma. According to the Categorical Imperative, pulling the lever would be categorically immoral as it violates the two central principles of Kant's model of deontological ethics; the act of killing a person cannot be universally justified, and doing so treats the individual as a means to an end, rather than an end themself. Thus, the Categorical Imperative would dictate that one is not to pull the lever, allowing the original five workers to perish. In contrast, the Utilitarian model would require the bystander to pull the lever, sacrificing one life in order to save five, reasonably maximizing the happiness of the most people.

The formulaic nature of consequentialist and deontological models of ethics allows these theories to systematically and definitively determine the moral verdict of an action according to a universal principle or set of maxims. The codification of ethics has the ability to frame dilemmas such as the "Trolley Problem" as objective equations that can be solved through frameworks that

simplify moral conundrums, and provide a clear course of action according to the principles of the given theory. Dismissing the role of human intuition in the determination of moral action makes the notoriously elusive and volatile scope of ethics tractable. Although they may yield conflicting verdicts on dilemmas such as the "Trolley Problem", consequentialist and deontological theories are united in their commitment to producing a strict code of action guidance by systematizing ethics.

While I understand the appeal of this approach, I ultimately believe the project of universally codifying ethics to be a futile endeavor. Competently evaluating the integrity of action guiding ethical theories requires an incisive understanding of the ethos of morality. It is possible that the evolution of moral philosophy in accordance with post-enlightenment epistemological standards has corrupted our understanding of the essence of ethics. Consequentialist and deontological ethical theorists such as Bentham, Mill, and Kant attempt to formalize ethics into a comprehensive and universally applicable model, as has become customary in the natural sciences and mathematics.

However, I believe ethics is qualitatively distinct from scientifically tractable disciplines, such as mathematics, which can be wholly defined by systems of universally valid principles and algorithms. Thus, theoretical models of morality should not imitate the epistemological rigor of these systems. Rather, ethicists should embrace the theory of anti-codifiability in order to develop frameworks that tactfully guide moral behavior in accordance with the established accessibility of moral truth. Peter Tsu Shiu-Hwa supports this departure from a rigidly scientific approach to ethics by arguing:

> "...there is a need to revise our common principled conception of morality and change our current practices in normative ethics and applied ethics. Instead of trying to formulate basic moral principles to distinguish between right and wrong in general, normative

> ethics should then concern itself with how the particular configuration of the action's contextual features come to determine its moral status" (Tsu Shiu-Hwa 147).

The theoretical definition of anti-codifiability is reasonably simple: general principles of morality fail to capture the complexity and uniqueness of particular circumstances, therefore, ethics cannot be codified according to a universally valid and coherent system of moral principles (Tsu Shiu-Hwa 145). As established in my introductory remarks, to espouse anti-codifiability is not to absolutely dismiss the value of ethical generalizations as rules of thumb. In the final section of this text, I will demonstrate how rules of thumb can coexist with, and complement, ethical frameworks that reject codifiability. However, as I previously suggested, in order for principles to be exceptionless, their scope must narrow beyond normative utility. Therefore, the imminence of exception to sufficiently broad generalizations renders them rules of thumb rather than universal truths. Naturally, this relegation of such principles from moral absolutes to rules of thumb establishes the need for an alternative determinant and source of action guidance, to which Tsu Shiu-Hwa astutely responds:

> "...some non-principle entity such as virtue, intuition or perception will have some indispensable roles to play in our moral reasoning process as the failure of principlism implies that there are bound to be some cases moral principles fail to cover…In cases where principle runs out, the notion of virtue kicks in" (Tsu Shiu-Hwa 152).

The embrace and implementation of non-principle entities as critical determinants in the frameworks of theories that observe anti-codifiability allow such models to ascertain elusive moral verdicts under complex conditions. As Tsu Shiu-Hwa indicates, virtue is a particularly efficacious non-principle entity that has informed moral action since Aristotle's original development of Virtue Ethics. I believe Rosalind Hursthouse's Neo-Aristotelian interpretation of Virtue Ethics, which will be the focus of my fourth section, demonstrates the indispensable role of non-principle entities in the determination of moral action. Hursthouse's framework, which I

will later use to defend the sufficiency and advantage of a revised standard of rigidity for action guidance, illustrates how the development and cultivation of virtue equips individuals with the contextual literacy and practical discernment necessary to successfully navigate the moral dilemmas that destabilize consequentialist and deontological frameworks.

Section III: Critique of Codified Theories

Utilitarianism

The failure of consequentialist and deontological theories to adequately account for contextual determinants acts as an obstruction to morality. Established applications of Kant's Categorical Imperative and Mill's Utilitarianism to moral thought experiments demonstrate how the operating formulas of each theory overlook contextual information that is absolutely essential in the determination of moral action.

The role of context in the pursuit of ethical action is well demonstrated by the impact of adjusting the details of the "Trolley Problem" without changing the ultimate result. According to Marc Hauser's *Moral Sense Test*, the morality of the sacrifice of one life for the salvation of five varies significantly depending on the contextual features. When presented with the original scenario, roughly 90% of participants believed it would be ethically permissible to intervene by pulling the lever (Hauser 139). However, adjusting the details of the experiment changes this figure significantly. An alternative form of this dilemma is to assume that the only way to save the five people on the track would be to push a large man off of a bridge, onto the tracks, in order to stop the train from killing the others. Under the Utilitarian theory, the implications of this scenario are no different from the original; the subject has the agency to sacrifice one innocent life in order to save five. Nonetheless, only 10% of respondents believed it would be morally permissible to push the man off the bridge (Hauser 139). This discrepancy suggests there is more to ethical decisions than the strict application of a code of action that solely serves to maximize utility.

Furthermore, Philippa Foot demonstrates the nuance of moral reasoning beyond the adherence to the universal principle of utility by adapting the "Trolley Problem" to the context of medical ethics, supposing:

> "A surgeon has five patients who are waiting for organ transplants. The patients will die if they do not receive the organs, but the organs are not available at the time. The whole prospect changes, however, when a young traveler comes to town and goes in for a routine checkup. The doctor is performing the checkup when he realizes that the traveler's organs are healthy and incidentally compatible with his dying patients. The young man is the perfect donor, and no one would associate him with the surgeon if he were to disappear. The dilemma here is, should the doctor remove the organs from the healthy man in order to distribute them to the dying patients?" (Andrade).

A faithful application the codified framework Mill provides would oblige the sacrifice of the traveler in order to save the five patients. According to the principle of utility, the surgeon would be not only justified, but obligated to distribute the traveler's healthy organs to his moribund patients. However, in surveys posing the same question, the overwhelming majority of respondents disapproved of the sacrifice (Edmonds). Foot herself accepted the moral impermissibility of authorizing such a transplant regardless of the consequential maximization of utility. The observed dissonance in the moral statuses of the original Trolley Problem and Foot's subsequent examples despite their equivalent production of utility indicates that the Utilitarian code of action fails to detect contextual determinants that distinguish the different cases.

Foot rationalizes this shift in moral consensus according to the principle of non-maleficence, which suggests the morality of an action is to be determined according to the precedence of negative over positive duties. According to Foot, the surgeon's negative duty of not harming the traveler is greater than the positive obligation of saving the five patients, while the original driver of the trolley is implicated as a source of harm regardless, and thus must minimize said harm by diverting the trolley (Andrade). In defense of this claim, Foot specifies

the original driver of the trolley as the agent rather than a completely innocent bystander. As this individual is responsible for setting the trolley in motion, Foot argues he is ultimately responsible for the consequences of its movement, while the doctor is not inherently responsible for the death of his patients, as he himself has not put the ill patients in peril. Foot uses this specific configuration of the thought experiment to support the primacy of non-maleficence in medical ethics, which I respect. However, for the sake of evaluating the integrity of the Utilitarian attempt at codifying ethics, I believe the more logical and effective strategy is to assume the agents in each respective scenario are operating under the same conditions of responsibility.

Although Foot's analysis does undermine Mill's Principle of Utility as the ultimate determinant of moral action, I believe she unnecessarily invokes the principle of non-maleficence to address exceptions to this code of action. Foot's designation of the driver as the original mover allows the primacy of negative over positive duties to function as a principle of action guidance that picks up on subtle differences in the premises of each solution, which alter the moral verdict when Utilitarianism stalls. Thus, Foot's argument suggests there are meaningful exceptions to the supposedly universal Principle of Utility which require the principle of non-maleficence to determine moral action.

Having established the susceptibility of Utilitarianism to exceptions through Foot's analysis, let us consider a variant of the Trolley Problem offered by Judith Jarvis Thompson to understand how exceptions to Mill's Utilitarian framework are not readily explained according to Foot's principle of non-maleficence. Thompson revises Foot's original scenario so that the agent is now an innocent bystander rather than the individual who set the trolley in motion (Tannsjo). As I stated, I believe this configuration in which neither agent, bystander nor surgeon, is directly

responsible for the five deaths should they fail to act is a more logical framework for evaluating the Utilitarian attempt at codification.

Referring to the results of the previous study, I am skeptical of Foot's attribution of the moral dissonance that occurs in response to her two scenarios to the perception of non-maleficence among the general public. It seems likely that the majority of respondents in Hauser's survey lacked Foot's abstruse understanding of non-maleficence, which she uses to determine:

> "This first variant of the Trolley Problem supports the primacy of non-maleficence in medical ethics. The five patients may die as a result of the transplant not taking place, but the surgeon is not ethically at fault since he has done no harm, and that is a doctor's most important duty. In order to save the five, he would have had to kill the one person. The surgeon wisely refuses to engage in such a procedure in deference to non-maleficence. The driver of the trolley, by contrast, does have the moral obligation to kill one in order to save five, because those five will die as a result of his own initial action. As opposed to the doctor, the driver is not in a position to claim that his duty is to first do no harm. This is because the driver already has done some harm by setting the trolley on course to kill five people. His moral duty is to take additional action to minimize his initial harm. Killing one is not better than letting five die, but killing one is indeed better than killing five" (Andrade).

While this analysis is certainly intelligible to those familiar with relevant philosophical literature, I struggle to believe the shift in sentiment among respondents was primarily informed by deference to the principle of non-maleficence.

Research conducted using Thompson's variant of the Trolley Problem indicates that the overwhelming majority of respondents still believed the innocent bystander is morally obligated to pull the lever and divert the trolley despite not being culpable for harm of the other five (Tannsjo). This reality confirms my suspicion that the principle of non-maleficence cannot be the only additional determinant that accounts for the moral discrepancy between the trolley and medical cases.

The lack of concord in the moral verdicts regarding Thompson's variation of the Trolley Problem and Foot's medical scenario demonstrates that the Utilitarian code of action fails to adequately account for the contextual features and that the qualitative differences between the two sacrifices that cannot be explained through Mill's principle of Utility, nor Foot's appeal to non-maleficence, but rather some non-principle entity.

Categorical Imperative

The supposed universal validity of the moral laws behind deontological codes of ethics also deteriorates under the pressures of moral dilemmas. I believe the dependence of Kant's Categorical Imperative on strictly exceptionless maxims to inform moral action ultimately compromises the integrity of the theory. The dialogue between Kant and Benjamin Constant initiated in 1797 demonstrates the shortcomings of Kant's efforts at developing a universal code of ethics. Constant argued:

> "The moral principle, 'it is a duty to tell the truth' would, if taken unconditionally and singly, make any society impossible. We have proof of this in the very direct consequences drawn from this principle by a German philosopher [Kant], who goes so far as to maintain that it would be a crime to lie to a murderer who asked us whether a friend of ours whom he is pursuing has taken refuge in our house. . . . It is a duty to tell the truth. The concept of duty is inseparable from the concept of right. A duty is that on the part of one being which corresponds to the rights of another. Where there are no rights, there are no duties. To tell the truth is therefore a duty, but only to one who has a right to the truth. But no one has a right to a truth that harms others" (Kant 1996: 425).

Constant challenges the universal proscription of lying, maintaining that the moral status of a lie is dependent on circumstantial features, such as to whom one is lying, and on what grounds. According to the contextual features of this scenario, Constant soundly determines the

murderer forfeits their right to the truth, and hence, there is no corresponding duty of truthfulness (Varden). In an attempt to defend the righteousness of his moral code of action, Kant responded:

> ". . . if you have by a lie prevented someone just now bent on murder from committing the deed, then you are legally accountable for all the consequences that might arise from it. But if you have kept strictly to the truth, then public justice can hold nothing against you, whatever the unforeseen consequences might be. It is still possible that, after you have honestly answered 'yes' to the murderer's question as to whether his enemy is at home, the latter has nevertheless gone out unnoticed, so that he would not meet the murderer and the deed would not be done; but if you had lied and said that he is not at home, and he has actually gone out (though you are not aware of it), so that the murderer encounters him while going away and perpetrates his deed on him, then you can by right be prosecuted as the author of his death. For if you had told the truth to the best of your knowledge, then neighbors might have come and apprehended the murderer while he was searching the house for his enemy and the deed would have been prevented. Thus one who tells a lie, however well disposed he may be, must be responsible for its consequences even before a civil court and must pay the penalty for them, however unforeseen they may have been; for truthfulness is a duty that must be regarded as the basis of all duties to be grounded on contract, the laws of which is made uncertain and useless if even the least exception to it is admitted. To be truthful (honest) in all declarations is therefore a sacred command of reason prescribing unconditionally, one not to be restricted by any conveniences" (Kant 1996: 427).

Kant's deontological model undermines the essence of morality by obliging immoral action for the sake of maintaining the supposed universality of his maxims. Regardless of his hedging toward a concession to the agent's discretion, Kant's codified framework ultimately prescribes truth telling in this situation. I believe Kant's defense of telling the truth to the murderer according to its legal defensibility demonstrates his prioritization of the political coherence of the Categorical Imperative over its facilitation of truly moral behavior. The overwhelming majority of individuals believe the ethical procedure would be to lie to the inquiring murderer in order to protect the potential victim. Many supporters of the Categorical Imperative recognize the validity of this response while still appealing to Kant's argument that, whether or not it is justifiable, the act of lying is inherently immoral (The Inquiring Murderer). I

believe the attempts of Kantians to reconcile the universal condemnation of lying in the context of the inquiring murderer through concerns of culpability demonstrate how fixation on absolute adherence to the Categorical Imperative impedes moral behavior; as long as the duty of veracity is regarded as a universal maxim, the framework requires one to reveal the victim's location to the murder. The necessitation of palpably immoral action under the Categorical Imperative supports the argument that ethics cannot be successfully codified, as universal maxims of action guidance are prone to detrimental exceptions.

Section IV: Indirect Action Guidance

The failure of codified theories to adequately account for contextual determinants in ethical dilemmas evokes the need for a revised standard of action guidance in normative ethics that embraces the complexity of moral decision-making through the integration of non-principle entities such as virtue into frameworks of ethical reasoning. In response to the failure of the Categorical Imperative to account for critical contextual determinants in the evaluation of the inquiring murderer scenario, Peter Tsu Shiu-Hwa argues:

> "The moral status of an action of lying is thus not determined by a set of principles that specify features as invariantly wrong-making or right-making, because such principles do not hold in certain cases. Rather, the moral status of an action of lying is determined by how its features are configured in a particular context" (Tsu Shiu-Hwa 152).

Tsu Shiu-Hwa emphasizes the contextual nature of moral evaluations, maintaining that the moral status of an action, such as lying, cannot be determined solely according to rigid principles that do not account for the specific context in which it occurs. The anti-codifiability argument posits that normative ethical frameworks should guide action indirectly through the careful consideration of the nuances and complexities of each situation.

Rosalind Hurtshouse's Neo-Aristotelian model of Virtue Ethics provides a sufficiently robust framework of indirect action guidance that accounts for critical contextual features that are overlooked by consequentialist and deontological codes of action. Hursthouse's understanding of the *telos*, or ultimate end, of virtue is rooted in the Aristotelian concept of *eudaimonia*, which refers to human happiness, flourishing, or well-being (Hursthouse 9). Through her Neo-Aristotelian framework, Hursthouse posits that the ultimate purpose of virtuous action is to promote *eudaimonia*, arguing the development of virtue guides individuals toward an ethical and fulfilling existence (Hursthouse 167). By grounding virtue in the pursuit of

eudaimonia, Hursthouse provides a comprehensive framework for ethical decision-making that emphasizes the development of virtuous character as central to human flourishing.

However, critics of Hursthouse continue to argue her framework of indirect action guidance fails to offer agents adequate direction in determining ethical behavior. Joshua Duclos provides a detailed evaluation of Hursthouse's model of action guidance through Virtue Ethics, ultimately arguing the framework she offers does not provide meaningful action guidance. Duclos summarizes Hursthouse's response in *On Virtue Ethics* to modern criticisms of virtue-based action guidance in the form of the following criteria for right action:

> P1- An action is right if and only if it is what a virtuous agent would characteristically do in the circumstances.
> P2- A virtuous agent is one who acts virtuously, that is, one who has and exercises the virtues.
> P3- A virtue is a character trait or a disposition to act in a way that contributes to the flourishing or well-being of the agent.
> P4- Courage is the disposition to persist in a valued project despite the risk of injury or death.
> P5- Courage contributes to an agent's flourishing by helping her pursue and succeed in valued projects.
> ---
> Conclusion - Courage is a virtue (Duclos 3-4).

Through this standard of moral action rooted in the necessary properties of virtue, Hursthouse provides what appears to be a sufficiently robust framework of indirect action guidance. However, Duclos argues that Hursthouse's model of Virtue Ethics cannot provide useful action guidance unless it is grounded in cultural consensus regarding the meaning and *telos* of the given virtue (Duclos 13).

At this point, I ask you to recall my refutation and dismissal of Duclos' relativism objection as a red herring. Hursthouse supports my expulsion of this criticism by maintaining that virtues are grounded in human nature and contribute to flourishing across humanity, thus providing a non-cultural criteria for evaluating virtue and determining right action (Hursthouse

206). Presuming my argument against this objection endures, I believe Duclos' action guidance objection can also be dismissed, as it is predicated upon the false assumption that Hursthouse's model of indirect action guidance cannot meaningfully inform moral behavior because it is relativistic.

Hursthouse's Neo-Aristotelian model of Virtue Ethics provides a nuanced approach to ethical decision-making that addresses the contextual nature of moral judgments while offering actionable guidance for moral conduct. I believe the sufficiently robust and contextually sensitive framework Hursthouse provides for indirect action guidance should replace the rigidity of codified frameworks as the paradigm of action guidance in normative ethics.

Conclusion

The value of human nature and moral instinct as a vital determinant in identifying ethical behavior is wrongfully overlooked in consequentialist and deontological theories. Hursthouse's framework of Virtue Ethics accounts for human nature, allowing our natural instincts to interact with other contextual factors to holistically determine moral behavior in ethically challenging situations.

Therefore, not only do I trust that Hursthouse's model of Virtue Ethics offers sufficiently robust action guidance, I argue theories that recognize anti-codifiability and embrace the role of non-principle entities in the determination of ethical behavior are superior to their codified counterparts. Throughout my defense of anti-codifiability, I have rejected universal laws of morality while acknowledging their value as rules of thumb. Hursthouse demonstrates how the *Phronimos*, or one who possesses virtue and practical wisdom, possesses a keen ability to navigate moral challenges with practical discernment and acuity while being naturally in tune with the moral generalizations to which consequentialist and deontological theories appeal, stating:

> "So the *phronimos* has a grasp of the important, the fine, and the necessary superior to that of most of us. And this suggests a way to go on in our consideration of the question, "What does the *phronimos* know?" which immediately proves fruitful. From other bits of the text, we see that he has a superior grasp of other concepts too, such as those of the fine (again), the expedient or useful, the (truly) pleasant and their opposites. He has a superior grasp of the right or correct as it occurs 'to the right extent, towards the right people, for the right reason, etc.' He also has a superior grasp of *eupraxia* - acting well - and *eudaimonia*. And he has a superior grasp of the virtues and vices" (Hursthouse 2022: 43)

By embracing moral education, agents of virtue use their experiences to strengthen their grasp of ethical rules of thumb while nurturing a virtuous character through which they can confidently navigate ethically challenging situations using contextual discernment and practical wisdom to determine moral action and the proper mean between virtue and vice. Hursthouse's

model of indirect action guidance naturally instills in agents virtue and an intrinsic sense of morality, while codified formulas do not.

Ultimately, I believe accepting the anti-codifiability of ethics is a pragmatic and beneficial step towards moral truth. As I demonstrated in Section III, codified theories such as Mill's Utilitarianism and Kant's Categorical Imperative face a double-bind between necessitating immoral action and appealing to a non-principle entity. Thus, people must be prepared to exercise moral discernment and practical wisdom as supposedly universal principles of morality inevitably encounter exceptions in order to determine moral action under the given circumstances. Therefore, embracing anti-codifiability prepares people to navigate the most difficult moral conundrums through indirect action guidance provided by theories, such as Hursthouse's model of Virtue Ethics, that cultivate virtuous character and the contextual literacy required to successfully navigate complex situations, unlike codified frameworks, which suppress these fluencies through futile attempts to reduce morality to a set of universal principles.

Works Cited

Andrade, Gabriel. “Medical Ethics and the Trolley Problem.” *Journal of Medical Ethics and History of Medicine*, U.S. National Library of Medicine, 17 Mar. 2019, www.ncbi.nlm.nih.gov/pmc/articles/PMC6642460/.

Aristotle, and W. D. Ross. *The Nicomachean Ethics*. Oxford University Press, 1959.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. Edited by J.H. Burns and H.L.A. Hart, Oxford University Press, 1996.

Duclos, Joshua. *Virtue Ethics and Action Guidance*, 2013.

Edmonds D. *Would You Kill the Fat Man?* USA: Princeton University Press; 2014. 56 pp.

Foot, Philippa. "The Problem of Abortion and the Doctrine of Double Effect." *Virtues and Vices*. Oxford UP, 1978.

Hauser, Marc. *Moral Minds*. HarperCollins, 2006.

Hursthouse, Rosalind. *On Virtue Ethics*. Oxford University Press, 2010.

Hursthouse, Rosalind. "What Does the Aristotelian Phronimos Know?" In *Virtue and Action: Selected Papers*, edited by Julia Annas and Jeremy Reid, Oxford, 2022. Oxford Academic, 15 Dec. 2022. https://doi.org/10.1093/oso/9780192895844.003.0004.

Kant, Immanuel, and Mary J. Gregor. *Practical Philosophy*. Cambridge University Press, 1996.

Kant, Immanuel, et al. *The Metaphysics of Morals*. Cambridge University Press, 2017.

Mencius, and David Hinton. *Mencius*. Counterpoint, 2015.

Mill, John Stuart. *Utilitarianism*. The Floating Press, 2009.

Sinn. "How to Calculate Morality Using the Utilitarian Calculator." *Think by Numbers*, 2 Apr. 2022, thinkbynumbers.org/utilitarianism/felicific-calculus/.

Tannsjo T. *Taking life: Three Theories on the Ethics of Killing.* UK: Oxford University Press; 2015. p. 58, 63, 75.

"The Inquiring Murderer." *Groundwork of the Metaphysics of Morals*, myweb.ecu.edu/mccartyr/GW/InquiringMurderer.asp. Accessed 12 May 2024.

Tsu Shiu-Hwa, Peter. "Can Morality Be Codified?" *Australian Journal of Professional and Applied Ethics*, vol. 11, Charles Sturt University, 2010, pp. 145–154.

Varden, Helga. "Kant and Lying to the Murderer at the Door... One More Time: Kant's Legal Philosophy and Lies to Murderers and Nazis." *Journal of Social Philosophy*, vol. 41, no. 4, 2010, pp. 403–421.