

University of Texas at Arlington

MavMatrix

2021 Fall Honors Capstone Projects

Honors College

12-1-2021

A PRACTICAL MORAL THEORY THE FAILED STATE OF NORMATIVE ETHICS

Gage Martinez

Follow this and additional works at: https://mavmatrix.uta.edu/honors_fall2021

Recommended Citation

Martinez, Gage, "A PRACTICAL MORAL THEORY THE FAILED STATE OF NORMATIVE ETHICS" (2021).
2021 Fall Honors Capstone Projects. 18.
https://mavmatrix.uta.edu/honors_fall2021/18

This Honors Thesis is brought to you for free and open access by the Honors College at MavMatrix. It has been accepted for inclusion in 2021 Fall Honors Capstone Projects by an authorized administrator of MavMatrix. For more information, please contact leah.mccurdy@uta.edu, erica.rousseau@uta.edu, vanessa.garrett@uta.edu.

Copyright © by Gage Martinez 2021

All Rights Reserved

A PRACTICAL MORAL THEORY
THE FAILED STATE OF
NORMATIVE ETHICS

by

GAGE MARTINEZ

Presented to the Faculty of the Honors College of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

HONORS BACHELOR OF ARTS IN PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2021

ACKNOWLEDGMENTS

This project would not have been possible without the guidance of my advisor, Dr. Williford, and all others in the Philosophy department that have taught me over the years. I would also like to extend my deepest gratitude to all the people who created the work that I reference and engage with in this thesis, both living and dead. And of course, to all those that supported me in the process of the creation of this thesis, emotionally or otherwise, have my sincerest thanks.

November 18, 2021

ABSTRACT

A PRACTICAL MORAL THEORY THE FAILED STATE OF NORMATIVE ETHICS

Gage Martinez, PHIL

The University of Texas at Arlington, 2021

Faculty Mentor: Kenneth Williford

Normative ethics can be described as the investigation into how one ought to act, morally speaking. The state of contemporary normative ethics is, essentially, a split between consequentialist (the idea that things are morally right or wrong based on their consequence) and non-consequentialist thought. Yet in practice, no one seems to be able to live by these moral theories, and I argue that it would be bizarre to do so as well. I examine reasons to believe that neither theory provides a complete account of real moral life, and how they fail at being prescriptive in practice. I provide a naturalistic account for the place consequentialist and non-consequentialist reasoning has in moral life, and analyze the differing meta-ethical groundings for them. This leads to insights about the exact nature of decision making, and ultimately having to deciding what role normative ethics plays in life.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
ABSTRACT.....	iv
Chapter	
1. INTRODUCTION TO THE PROBLEMS OF NORMATIVE ETHICS.....	1
2. CONSEQUENTIALISM AND NON-CONSEQUENTIALISM.....	6
3. GROUNDING A PRESCRIPTIVE THEORY.....	14
4. STRATEGY AND TACTICS	25
5. A PRACTICAL THEORY	33
6. CONCLUSION.....	49
REFERENCES	50
BIOGRAPHICAL INFORMATION.....	53

CHAPTER 1

INTRODUCTION

What motivation do we have for studying ethics? One of the arguments philosophers often use to justify the purpose of ethics is that by examining morality, we are increasing moral understanding and bettering people overall. We can develop better moral reasoning and rational capacities, discover and ground objective moral truths, eliminate biases, and come to a greater overall harmony. There is only one problem: none of this is true. Even professional ethicists, the best-case scenario, by and large behave indistinguishably from the general populace (Schwitzgebel & Cushman, 2012). The people who should be the “moral experts” among us have, even in the hypothetical, the same moral intuitions making virtually the same moral judgements. This is not even getting to the practical, where ethicists performed the same, if not worse, as the general public in putting their moral beliefs into practice (Schwitzgebel & Rust, 2013). This study shows that philosophy can lead to a more stringent view of moral obligation, as that was what really rose across the board no matter the particular normative view, but almost no change in practical action is produced. Basically, ethics seems to assert that we really do have certain moral obligations, but that understanding the basis behind it is not enough motivation for any real change in behavior. For all our sophisticated reasoning and widely differing theoretical beliefs, it does not seem to amount to much in practice. Is the conclusion to be made from this that, simply, humans do not behave rationally? That is probably one of the safest takeaways. But what might this say for ethics overall?

If all this really is the case, then the question to ask is why even bother with ethics at all? Normative ethics, with the goal of prescribing how moral agents ought to act, might be in jeopardy here. This is not even taking into consideration how little difference the various normative theories themselves produce. The contemporary split in normative ethics can be viewed as a conflict between consequentialism and non-consequentialism (primarily deontology, but others such as virtue ethics as well). Does the key in deciding what we should do lie in the consequences of an action, or is it irrespective of them? Well, it would seem that the upholders of both ultimately judge actions and make decisions independently of the theories. The picture is not looking great.

And as for leading theories themselves, it is not hard to understand why people generally do not adhere to them in reality. A famous example of a philosopher seemingly failing to ‘pull the trigger’ when it goes from abstract to concrete is the anecdote from Jean-Paul Sartre regarding a student in need of advice (Sartre, 2007). The student was trying to decide between joining the French resistance against the German occupation or staying and protecting his mother who had no one else left to stay with her. Sartre’s response was to, effectively, say that it was up to the student to decide. Many took this as a weakness in Sartre’s ethical views, but in truth it is doubtful that other philosophers would fare much better in their advice. Both consequentialism and non-consequentialism prescribe actions that almost no real individual would, and sometimes can, ever undertake, or are simply so far from reality as to be unacceptable on principle. And that is only in practice. Even theoretically, there are plenty of issues that can render a theory moot. There are problems that intersect with metaphysics, epistemology, psychology, and more that endanger the viability of normative theories even just within themselves.

With this in mind, there might be a temptation to downplay normative ethics and accept its status as an interesting, but ultimately worthless pursuit in practice. Certainly, it would be good for us to leave behind the convenient lie that a solid normative theory assists in making people behave more morally. After all, we do not take criminals and those in need of reformation and teach them normative ethical theories. There might still be justification for normative ethics however, but a reexamination, and perhaps a redefinition, seems to be in order.

The first problem is answering why exactly it is that differing normative theories do not lead to differences in practice. This requires to first understand the theories themselves. The primary theoretical distinction that will be made in this discussion is between that of consequentialism and non-consequentialism. This is because the dual process theory of moral judgement that will be addressed later gives us good reason to believe that this is really the fundamental split in morality (Greene, 2007, Dubljević et al., 2018). But for a more basic formulation of that split, we can turn to the famous trolley problem to get a sense for the tension. Assuming everyone knows the basics, typical consequentialist theories will prescribe pulling the lever (and the equivalent decision in other cases) while nonconsequentialist theories can prescribe not pulling the lever for one reason or another. Is what we should choose to do decided by the consequences of an action, or the action itself (or somewhere else)? Consequentialism is in the former, non-consequentialism (but particularly deontology) strays to the later.

Now, the immediate rejection to this is that this is not true of all consequentialist and non-consequentialist theories. It is true that this is only the fundamental characteristic of each. A consequentialist theory can prescribe not pulling the lever while a deontological

or virtue-based account can lead to pulling the lever. And this is, essentially, what happens in practice. Sidgwick's utilitarianism, for instance, is argued to encompass common sense morality, in other words one could make a consequentialist argument for a great deal of what intuition suggests (such as pulling the lever to be acceptable while pushing the man is not) (Sidgwick, 2007). Non-consequentialists also make arguments to the same conclusion, such as Judith Jarvis Thompson, a deontologist, relying on a rights conception to explain why the lever is ok, but pushing the man over the footbridge is not (Jarvis Thompson, 1985). This itself, however, is another reflection of the core problem. If it is possible to construe any theory to get any result (with enough ingenuity), what is the purpose of having these different schools at all? There is the fear that what we are doing is not really objective reasoning that challenges our intuitions, but rather just crafting theories that comport with them. But, maybe that is exactly what we should be doing. This is one potential conception that needs to be examined in the redefinition of normative ethics.

Another consideration is in the functional role of normative ethics. Is it a potential guide for decision making, or merely an evaluative meter for determining what is right and wrong? After all, if it is not to act as a guide in the first place, the practical failings of it are of no great concern. Traditionally, there have been efforts to do away with the former and focus on the later so as to preserve the integrity of the theory. For instance, Henry Sidgwick defended utilitarianism by insisting that it was not to be used as a conscious way to make decisions, with others positing similar views that have held throughout the tradition (Sidgwick, 2007, Mill, 2000). It can be argued, however, that normative ethics primary application should actually be as a prescriptive tool rather than an evaluative one. This

thesis will, in fact, be devoted to articulating and defending this view. But first, a review of consequentialism and non-consequentialism is in order to get things started.

CHAPTER 2

CONSEQUENTIALISM AND NON-CONSEQUENTIALISM

Consequentialism is the family of theories which hold that “whether an act is morally right depends only on *consequences*” (Kagan, 1990). The progenitor to all modern consequentialist theories is utilitarianism, though many scarcely resemble its ancestor anymore. However, in the interest of simplifying the explanation of the problems that consequentialism faces, utilitarianism will be used as a base, from which each problem creates an offshoot to some more particular form of consequentialism (such as preference utilitarianism following issues with utility). Utilitarianism, simply, is the position that the right action is “that which will produce the greatest amount of happiness on the whole” (Sidgwick, 2007, p. 411). One of the main problems to be taken with utilitarianism, and many consequentialist theories, is that the means justify the ends. Utilitarianism commands one to pull the lever that kills one to save five. That is not so controversial, in fact most people agree with it. It also, however, commands one to push the man over the footbridge, killing one to save five. This is already hard to swallow for many, but it can be made much more unacceptable just by changing the numbers a bit. Imagine there are two boats at maximum capacity sailing together: boat A carrying 100 people and boat B carrying 101. A hole opens up at the bottom of both boats, and only one can be repaired for one reason or another. Fatal in-fighting between the boats will be inevitable unless action is taken. You are also, in this situation, the sole possessor of a lethal weapon, let’s assume an assault rifle, on the boat. What should be done? Utilitarianism might readily suggest that not only

is the right course of action is to gun down all 100 people on boat B, but that you are *obligated* to do so. In this scenario, at the minimum 100 people must necessarily die, and it seems that killing all the people on boat B is the best way to ensure that the ‘best case scenario’ is realized.

Of course, few people’s intuitions would accept gunning down 100 people to save 101 as the right thing to do, even in this situation where the consequences are more or less clear. Hopefully, no one’s intuitions would *obligate* them to gun down the 100, but so long as it is this action that maximizes overall happiness (and there is reason to believe it does), utilitarianism necessitates it. But this, of course, leads to another, much greater problem that plagues both utilitarianism and consequentialism as a whole: how do we know that this is what leads to maximizing overall happiness? After all, it might be the case that the trauma inflicted by the murdering of the 100 weighs down overall happiness to the point where it would be better to allow for the in-fighting that would claim more lives, but ultimately maximize happiness. This problem holds no matter what unit is swapped for happiness, so long as it is some abstract quantifier. The natural and fair rejection to this is to point out that this is an epistemological issue, not a moral one. It does, however, become a moral issue if there is no reason to believe that even in all-things-considered cases, there is no way to calculate the happiness of one outcome as compared to another. What this means in practice is a lot of *post-hoc* reasoning. If the conclusion suggested by the theory is unsavory (such as with the murder of the 100), then no problem; just change the calculus such that whatever decision is desired is that which has the best consequence under some specific criteria.

Moving to an issue that concerns utilitarianism in particular is the centering of utility (though particularly the presence of pleasure and the absence of suffering) as the ultimate good. If morality is all about maximizing pleasure and minimizing suffering, then consider this situation. Science has been able to step up a system where a rat's brain would receive a direct stimulation to its pleasure center upon doing something like pushing a button (Wise, 2002). This would lead to the rat compulsively pushing the button thousands of times, foregoing all other stimulus in favor of the button, leading to the rat's starvation. Clearly, the rat gained more pleasure from this artificial system than anything else, to the point where eating was a trifling distraction, and starvation was irrelevant. If we are to maximize pleasure and minimize suffering, the ideal life of the rat would then seem to be simply pushing the button as many times as possible. To make this more pertinent, let's switch from rats to humans. Say this same system is developed for a human, and has, predictably, the same results (if one believes the results would differ, we need only look to drug addiction as a counter example). Furthermore, let us stipulate that life support systems are provided so that an individual can press the button as much as they want without any concerns of survival to stop them. Is this a superb life? On top of that, if this was possible, would we be obligated to provide this sort of life? Intuitions certainly suggest that not only is this not the ideal life, but a somewhat terrifying and meaningless life. Nevertheless, if what is right is maximizing pleasure and minimizing suffering, the creation of pleasure batteries, as they will be called, would be a great thing. Not only would they be great, but if they generate enough pleasure, the creation of them may very well be our highest moral obligation. In fact, the most ethical thing we can do to animals would be to stimulate the pleasure centers of their brain for their entire lives. No, this is clearly not the case. Also, it

does not matter if pleasure is swapped with happiness, or some other sensation or psychological state that is purported to be the ultimate good, as then we simply develop a way to cause *that* via brain manipulation. Hedonism as the base of utilitarianism is the motivation for many splits from it, such as negative consequentialist views and, in an indirect way, preference utilitarianism.

Another potentially serious problem for consequentialism is in the demandingness objection. The objection resulted from a famous argument made by Peter Singer about our obligation to distant, unknown strangers (Singer, 1972). In it, Singer argues that we have the same obligation to save a child drowning in front of us as we do, for instance, to donate to a charity that will save a child's life in some distant country. The point of the argument, in that we should be more charitable, is certainly agreeable. The philosophical implications are much more suspicious. As demonstrated by the argument, consequentialism leads to the view that supererogatory acts (actions beyond the 'call of duty') do not exist. That is a problem itself, but to go further, all supererogatory actions would then be morally obligated, such as donating a great deal to charity. This seems to put such a high demand on morality as to render virtually anyone making even basic personal allowances to be moral criminals equal to murderers (that is in Singer's own words). This alone will lead many to either abandon the theory or make significant alterations (Scheffler, 1986). In any event however, it certainly gives one difficulty in pairing it with intuitive morality.

And then there is the matter of intention. In many forms of consequentialism, it is the actual consequence of an action and not the foreseen or intended consequence that is morally relevant. The former can result in predictably absurd things, but this is one case where the offshoot (subjective consequentialism rather than objective consequentialism)

seems rather safe, though some will wish for the strength of actual consequentialism, nevertheless.

To move on to non-consequentialism, this is, as one might infer, the wide range of theories that do not regard the consequences of actions to be the sole determiner of what is right. Though this includes many different theories, the most popular family of theories, and the one that will be used as the primary non-consequentialist example, is deontology. This is because while consequentialism holds that the results of an action are what matter, deontology, in contrast, claims they are irrelevant; it is the actions themselves that are either permissible or not (Alexander, 2021). And similar to utilitarianism, modern deontological theories are derived from the moral philosophy of Immanuel Kant, but depart in many different directions. At the heart of Kantian ethics is the idea that the only intrinsically good thing is a good will (Kant, 2012). From this follows the categorical imperative, which is, supposedly, a perfectly objective and rational principle which commands agents to act regardless of their own desires, or of the consequences of the action. It is from this that all of the more simple moral laws are derived, such as the prohibition of lying. The most general principle by which a duty can be derived is in its universalizability. It is a true duty only if it can be willed upon all agents as a universal law without violating rationality or autonomy. The character of the ethics that results from this is strongly duty based, hence the name deontology.

As a first observation, one notable weakness of deontology is its tendency towards being grounded in non-natural, non-physicalist metaphysical worldviews, though this is not necessarily so. This is because of a general problem with deontology in the precise grounding of misty concepts that are central to it, particularly those of rights and duties. If

one takes an approach such as divine command theory, it falls upon a deity to ensure these rights and duties are real (and, in some cases, rewarded). For one seeking a purely naturalistic ethics, this is obviously not acceptable.

The most famous problem with deontology is one that Kant himself raised: the murderer at the door. Imagine a terrified individual is running from a murderer and wants to hide in your home. You, of course, let them in. Then next appears the murderer, axe in hand, and asks whether their target is in the house or not. Kant, bafflingly, insists that in this case one is obligated to tell the truth to the murderer (Kant, 1949). Whether or not the murderer kills the person you are hiding as a result is irrelevant, as you followed your duty to not lie. Virtually no one, not even Kant's contemporaries, were persuaded by this argument. Modern Kantians insist that Kant misinterpreted his own theory here, and that a more perfect conception of moral duties resolves the issue (Cholbi, 2009). But it is just that that is the problem; how to determine what these moral duties are, how they interact, and how they are justified. In this sense, Occam's razor certainly favors consequentialism.

To illustrate this in the example of the murderer at the door, there seems to be multiple duties at work here. The obligation not to lie is at work here, but what about other duties such as protecting others? Kant held firm that, in truth, moral obligations can never be contradictory. Yet there seems to be numerous cases in which they are, such as this one. If that is the case, just how is one to determine which moral duty overrides the other? Ultimately, in that example we want to be able to factor in the apparent horrible consequences of telling the truth into whether we should do it or not. Deontology simply does not allow for that direct consideration. Everything must follow from the categorical imperative.

A critique of Kantianism that will gain relevance in the following discussion comes from Friedrich Nietzsche over the role of reason in morality. For Kant, reason is the force that overrides our base desires in order for one to follow the commands of duty. This is because, allegedly, reason can come to a decision independent of aims or desires. Nietzsche points out that this is not the case, as reason is just not capable of being entirely separate from the self's desires and motivations (Janaway & Robertson, 2012). In reality, when it seems that reason makes decisions against our desires, what is really happening is that one stronger desire overtakes another and nothing more. As will be discussed later, this critique is relevant not only against deontology but any rationalist normative theory in general.

To give some attention to non-consequentialist theories that are not deontological, virtue ethics is seen as the next leading competitor in normative ethics against consequentialist and deontological views. It, as its name suggests, is centered in the importance of moral virtue. What distinguishes it from consequentialism and deontology is the unique importance that it puts on the evaluation of agents' character, rather than actions themselves or the consequences of them. It is important to note that our intuitions seem to put a similar weight on character as we do actions and their consequences when making moral evaluations (Dubljević et al., 2018). However, it is somewhat mistaken to place virtue ethics in opposition to the other theories, as it is commonly viewed to be consistent with both consequentialist and deontological theories (Nussbaum, 1999).

So, to boil down all of this, the main takeaway is that consequentialism says that consequences are the only thing that matter, while non-consequentialism says that consequences do not matter at all, and that it is the actions themselves which are morally relevant (such as with deontology). But does it have to be one or the other? Why is it that

there cannot be consideration for multiple components in moral judgement? In other words, each are on a binary opposition regarding the moral value of consequences, when this might not be necessary at all. The main problem in reconciliation, however, is how exactly we should balance (and even evaluate) good consequences and good actions. Perhaps the real question to be investigated is not consequentialism or non-consequentialism, but *when* consequentialism and *when* non-consequentialism.

CHAPTER 3

GROUNDING A PRESCRIPTIVE THEORY

Thus far there has been a description of consequentialism and non-consequentialism in theory, but there may be a hidden story between the two on the psychological level. This means a foray into moral psychology. There is a general resistance towards doing this in fear of violating the is-ought distinction. This is, however, largely unnecessary, and an understanding of the nature of our moral intuitions can in fact have normative implications.

The work of Joshua Greene, a moral psychologist, gives a unique insight into what is happening on the level of intuitions in regards to the conflict between consequentialism and non-consequentialism. He gives a plethora of empirical evidence to suggest that there is a dual-process in moral judgements (Greene, 2007). This refers to two distinct modes of thinking that are correlated with making consequentialist and non-consequentialist judgements. It is shown that in consequentialist judgments, the centers of the brain that are associated with calculation, reasoning, arithmetic and so on are primarily what is engaged. Meanwhile with non-consequentialist judgements, there is a strong association with the engaging of emotional intuitions. Greene calls them “cognition” and “emotion” respectively, though the name of the label can differ. On the one hand there is a calculative, reasoning, cognitive, manual mode of thought, while on the other there is an emotional, intuitional, automatic mode of thought. This distinction is used to explain why there tends to be different results in “trolleyology”, i.e., why the switch scenario gets different results

than the footbridge (Greene, 2014). This is because in the switch scenario, the emotional distance between the agent and the person on the track is such that the cold, calculative mode of thought can take over, and point out that five is greater than one. When direct violence is needed to achieve the same end however, such as with the footbridge, an emotional reaction is produced against it. This is thought to be because our emotional intuitions evolved to be averse to physical violence but does not hold the same strength in distant forms of killing. This also goes for our revulsion towards things like incest, even with contraceptive preventions. What makes these cases interesting is that the emotional reaction evolved for a function that, due to some circumstantial factors, are not being fulfilled in these scenarios. For instance, our emotional reaction against killing is founded in cooperating with others and ensuring their survival. For the trolley cases, however, killing is necessary to save the greater number of people, and so the function of our intuition is not being satisfied.

Though Greene does not point this out, it is the social emotions in particular that spur the unique moral quality of the intuition. We humans, like all animals, are equipped with a wide range of emotional intuitions, but not all of them necessarily serve a moral purpose. For instance, while the desire for food and sex are some of our strongest intuitions, they do not seem as intrinsically moral as, say, loyalty or compassion appear. Compare the moral quality of physical fitness with the quality of keeping one's word, for instance. The latter, being a social intuition, appears moral in a way that the former simply does not, despite the former having more direct evolutionary weight.

Greene uses the dual-process theory to ultimately argue for consequentialism, claiming that it involves true moral reasoning while deontology is instead moral

rationalization of emotional intuitions. The argument that deontology is moral rationalization is rather persuasive, while proving consequentialism to be different is not as simple. Regardless, the takeaway from dual-process theory in the interest of this project does not lie in the state of moral reasoning or moral rationalization (though that is relevant). The idea to keep in mind here is the state of consequentialism and non-consequentialism as a theoretical formalization of two distinct modes of thought. Henceforth, any mention of ‘moral modes’ will refer to the differing modes of cognition as described by the dual-process theory.

Important to the coming discussion is the establishment of a metaethical framework. In short, a form of moral anti-realism is to be adopted, though it is functionally not so far from moral realism. Moral realism holds that the world contains objective moral facts. Moral non-objectivism challenges this by asserting that moral facts are not, in fact, mind independent (Joyce, 2021). Morality is a result of certain mental states, but within that mental world, morality is objective so long as the certain features that constitute morality are shared. This will be further articulated in the coming arguments.

In order to gain a better understanding of what normative theories really strive to be conceptually, there is an analogy that can reveal a surprising amount about what it is we are doing when we build normative theories. We can compare it to game theories (not in the mathematical sense, but rather in the sense of the analysis of specific games); chess theory in particular will be used to illustrate this.

When we analyze a game like chess, we are looking at how we should play the game in general, and on the micro level, what moves in particular we should make in a given position. With normative theory as well, we are analyzing what we should do in

general, and on the micro level, what decisions we should make in turn. Let's imagine for a moment that we are looking at a position in chess and have to decide which move to play, so we consult our much more skilled and knowledgeable friend and ask "What move should I play in this position?" We will probably receive some response along the lines of "You should make X move because, for all these different reasons, it will lead to the best position out of your available options." This is a fair, appropriate response, but then it raises the question, "Why is the position this move leads to the best?" Assuming we are asking in the philosophical sense here (in other words, Why is the best position 'the best?'), the correct response would be something like "Because that position gives you the greatest chances of winning as opposed to all other options." But we can ask the question once more: "I understand that we should play this move in order to win... but why should we win?" In other words, is there anything in chess theory (defined as a body of theory concerned with how the game of chess should be played) that suggests we should play to win rather than lose or draw? Or even for some other reason, such as achieving a particular position, or gaining as much material as possible, and so on? The answer is no. There is nothing in any position itself that makes it intrinsically valuable, not even a checkmate because, after all, the only thing entailed by a checkmate is that the game is decided as either a win or loss. Winning and losing itself, however, only has value outside of the chess board. And yet, chess theory does still prescribe moves. It does not, however, determine the value of a win, draw, or loss. It is prescriptive only under the assumption that the player *intends* to win (or in the case that this is not possible, to draw rather than lose).

Accordingly, when consulting your friend on what move you should play, we are not asking whether we should win, or draw, or achieve some other aim. The desire to win

is implicitly assumed, as if it is the case that if you are asking whether you should win or lose, then you are no longer asking about the game itself. How is one to give a response then? We should win because... winning is just good? Though there might be some debate on it, 'good' is not analytically found within 'winning.' If one played with the intention of losing for some reason or another, then, in that case, winning would be bad. In practice, when we speak of best positions, good moves, optimal play and so forth, we are doing so under the assumption that we are playing to win, or at least to not lose. There is no chess theory to determine what our objective ought to be when playing the game. That is left to only the intention of the player, but the overwhelming majority intend to win, as that is what they want; and so that is our assumption in the creation of chess theory.

So how does this relate to normative theories? Essentially, whatever our ultimate value is—whether that be maximizing utility, or abiding by the categorical imperative, or anything else that's considered to be an intrinsic good—that itself is not justified within the theory, just as chess theory doesn't justify winning. It is reasonable to think that what is really going on here when we speak of what we ought to do is similar to what's happening in chess theory. We all assume some common intention, an ultimate moral desire, then evaluate situations and decisions based on how well they advance this aim. Normative ethics is not concerned with deciding what this intrinsic good should be, just as chess theory doesn't concern itself with why we should win. It is, in truth, about the *how*.

To be clear, this doesn't mean that the identity of this assumed common intention isn't relevant; it is critical to a clear ethical theory (you need to know a player's intention to prescribe a chess move, after all.) The thing that is being highlighted here is that prescriptive theories don't actually concern themselves with what has intrinsic value; it's

already assumed. For instance, do we consult chess theory when trying to determine if we should win, lose, or draw? No. We consult chess theory to learn how to play better chess, already under the assumption that we play to win. Accordingly, the debates within chess theory reflect this, being matters of what principles work better, what strategies are characteristically sound, and so on. Essentially, a tool to learn how to play good chess. Normative theory should follow suit and be a tool to learn how to live a good life. Furthermore, what is being thought of as a conflict of value between normative theories is, in fact, really a debate on practical effectiveness. The question of moral ought is about what helps us achieve assumed intrinsic value x better. Is it making decisions based on reasoning out the consequences of actions, or relying on our emotional intuitions regarding the actions themselves? This is the real, practical conflict between consequentialist and non-consequentialist theories.

An immediate objection to this is that this is actually mischaracterizing normative ethics and what 'ought' means. What normative ethics is doing is asking what we ought to do in the same sense of asking what game state ought to be reached in chess; be that a win, loss, or other (or so the objection goes). This, however, is a misapplication of 'ought'. There is no intrinsic value within the game of chess unless there is already an intended game state by the player. To ask about the intended game state is to transcend the game. It is no longer a question about the game but is instead about the world beyond the game (the player's psychological state) from which value is generated. In short, it is a question of *is*, not *ought*. The real-life moral equivalent is an attempt to transcend the phenomenal world into the noumenal. There is no ought in the noumenal realm. Consider these questions in the noumenal sense: Ought massive objects attract each other? Ought the stars be arranged

some particular way in the sky? Ought fire burn rather than sputter out? And probably most interestingly, ought anything exist rather than not exist? Remember that these are asked in the noumenal sense.

It appears dubious that these questions could even apply to matters of pure existence. In the objective, noumenal world, there are only two states; is and is not. Something either is the case, or it is not. Therefore, there can be no noumenal ought without a phenomenal intent, due to the is-ought gap. If one is a moral realist, then yes, but mere states of affairs appear to receive ought claims differently from agents. It is merely the case that massive objects attract each other, that the stars have their particular arrangement, that fire burns or does not, and that anything exists or does not. The question of ought is, fundamentally, one derived from the will, our very intention itself, and that is exclusive to the phenomenal world, the 'world of the game' to return to the chess analogy.

Here is where we need some Humeanism both to enforce this claim and to avoid glaring pitfalls. There has been much talk about assumed intentions, desires, the will and so forth as being what values are derived from and being immune to questions of ought. There is a saying being parroted in popular culture from time to time along the lines of "Feelings aren't right or wrong, they just are." This is spoken in the sense of claiming that having a certain sentiment (what Hume refers to as 'passion') is non-moral, but what acts result from it is a different matter. A sentiment is a matter of *is*, not *ought*. When combined with Hume's claim that "reason alone can never be a motive to any action of the will" and "can never oppose passion in the direction of the will," it can be concluded that the will (one's intention, in other words) is in itself non-moral (Hume, 2003, p. 413).

The first problem to address here is actually accepting both the premises. The first claim, as it is merely a common saying, is likely to be accepted on mere intuition. If intuition disagrees in fact, think of the situation this way. If one person says, “I’m angry” and someone responds with “Well you shouldn’t be. You ought to be happy,” it’s clear that we are not talking about the moral value of being angry or happy. Those claims only make sense in the sense of how justified they are relative to their causes. To outline the exact reasoning behind it, one has no more direct control over what sentiments arise within one than, for instance, a physical reflex (we can condition ourselves as Aristotle pointed out with virtue, but it will regardless always be outside of our conscious control.) When we ask if a sentiment is justified or not, we are directing our attention not at the moral quality of the sentiment itself, but rather whether or not the sentiment is serving some appropriate aim given the stimulus. If one says, for instance, that they feel scared, when we evaluate that sentiment we look to see if the stimulus justifies it (i.e., whether or not being scared is serving its functional role), not at the moral value of being scared itself. If one takes sentiments themselves to have moral qualities, it could lead to the conclusion that merely having some emotional disposition to things like sadness (such as with depression) or anger is directly moral. For instance, a theoretical inverse of depression where one feels inappropriate happiness would be a morally good condition. Of course, some theories like Utilitarianism take emotional states like happiness to be, in fact, the ultimate moral quality. But this was cast into doubt with the pleasure battery discussion in chapter 1, and the reasoning here justifies why. One’s brain being directly stimulated is not what we consider to be an appropriate justification for happiness; in the same way that we don’t consider social contact to be an appropriate justification for fear in someone with agoraphobia for

instance. The sentiment is still felt, but it is non-moral to the extent it is irrelevant in serving our implicit moral desire (to be clear, it often is relevant; the purpose here however is to illustrate that this is not necessarily so.)

To defend Hume's claim that reason alone can never be the motivation to act, the social theory of reason can give evidence to believe it. The relevant aspect in this discussion is how we do not in fact go from reasons to conclusions as is often idealized, but in fact reasoning is evolutionarily designed to produce reasons for conclusions we've already decided upon (Mercier & Sperber, 2017). This manifests itself in the form of many flaws in reasoning that are so pervasive, they seem less like mistakes are more like features. Examples are the ease with which we can perform post-hoc reasoning for decisions that were never actually made, the prevalence of cognitive dissonance and its resulting confirmation bias and rationalizations, and so on. When making decisions, it is really intuitions that make the decision, and then reason gives justification by attempting to demonstrate that this action will actually achieve an intended end. But that intended end is beyond reason. When reason seems to be guiding our decision making, what is really going on is that reason is making inferences that give us a greater understanding of the consequences of a given choice. How we actually evaluate the consequences that are revealed by reason is by relating them to some end that is determined by intuition, not reason. It is closer to revealing the route to a destination, rather than deciding the destination itself. In short, one way to think of it is that reason reveals where a potential path will lead, but cannot by itself choose one path over the other. That is left to how well the path leads to a 'passion.' This is enough to give credibility to the claim that in moral matters, the "direction of the will," contrary to common thought, is actually non-moral.

A reasonable concern with the claim that the will/intentions are non-moral and beyond reason is something like this: It might be concluded that if an individual feels they should bring about mass murder, then they ought to kill a bunch of people. After all, what ought really is is just acting according to a value generated by a pre-existing non-moral intention, as a way to realize a given end. There are multiple clarifications to be made that resolve this concern. The first is to recognize a correct part of the concern; this desire to bring about mass murder is, in fact, non-moral as all sentiments are. Realizing it, however, obviously isn't.

The objection to the statement about the mass-murderer-in-the-making lies in what 'ought' means. Here is where things might get a bit speculative. The function of the general ought would indeed prescribe that if one has some aim, they ought to do whatever it may be to realize said aim (that's the principle of practical rationality). Indeed, just asking "I am hungry, ought I to eat?" would in one sense result in answering in the affirmative. But that isn't the moral ought we are talking about. When speaking of the moral ought, in the same way that one assumes the intent or will to win in chess, we too assume some particular intent in morality, some end that is what makes an agent a moral one rather than non-moral. Whatever the identity of this assumption is what grounds our moral *ought* in what actually *is*. This assumption will henceforth be referred to as the 'moral aim.' Different schools of thought each of course suggest an identity for this moral aim; the maximization of utility, the categorical imperative, and so on. These are critical to any normative theory, but the suggestion being given is that they are not in fact claims of what our moral aim *ought* to be, but rather a matter of what already *is*.

It should be noted that even at this level, there is already a level of consequentialism present. This is said in the sense that, under this framework, prescriptive morality is concerning itself with what actions will be effective in achieving some particular desired end (that is already assumed).

CHAPTER 4

STRATEGY AND TACTICS

And so, the question looms large: what exactly is the identity of the moral aim? I will suggest that there are in fact two different, though very much related, moral aims that will ultimately decide whether we should use consequentialist or non-consequentialist thinking in a given situation. Our highest-level moral aim, the one from which the other is derived, is, essentially, an idealization of desire itself; the intent to persist (biologically). It can be formulated as something like “We, collectively, aim to (biologically) persist, therefore we ought to do that which furthers this aim.” This does not run afoul of the naturalistic fallacy due to what has been previously discussed. The general function of ‘ought’ itself is in determining the most effective means to some end. The difference between the moral ‘ought’ and the general ‘ought’ lies solely in the identity of the end in question. To give a clearer definition of persistence, this is said in the sense of a sort of ontological disposition that life has. That which is disposed to existing continues to exist, that which does not, fails to continue existence. Organisms are disposed to extending their causal chain in a particular way (as opposed to, for instance, a rock which needs not do anything a biological creature does to continue existing), and this is what persistence really is. This assuages the misconception that biological success is strictly the proliferation of one’s genes. This is not the case with persistence. For instance, an organism (such as a human) could, on the causal level, be a great benefit to persistence without ever having

proliferated their genes. Put simply, one does not have to reproduce to further the aim of persistence (though that is often the simplest and most necessary way to further the cause).

This desire to persist is inescapably ingrained in our psychology, in fact it is the end of desire itself. The other moral aim is set in our particularly human way of living: the intent to sociality. That can be formulated as “We, collectively, aim to behave prosocially, therefore we ought to do that which furthers this aim.” To define prosociality, there are many potentially accurate formulations, but the most well-known encapsulation is the Golden Rule: Do unto others what you would have done unto you. This is not an infallible conception, but the core prosocial ethos is contained within it. It is also probably no coincidence that it somewhat resembles the categorical imperative (Blackburn, 2003, p, 103). We ultimately want to persist, but the particular way we do so is via our hypersocial nature. It is, in reality, the means we employ to the greater end of persistence, but it is the means that we are cognitively locked into due to what our intuitions actually do. If we were biologically structured in some different way, it might be to the greater interest of persistence to be anti-social rather than the highly social creatures we actually are (to give an example of a mammal that is this way, wolverines are highly anti-social, almost never contacting each other). In other words, prosociality might not be the objectively best strategy for any organism to persist. This is relevant in limited (but extremely important) circumstances that will later be highlighted. It doesn’t take much to see that due simply to human nature, prosociality helps us persist more than some other tactic does. When the assumed moral aim is persistence, consequentialism shines, but when the assumed moral aim is prosociality, this picture becomes murkier.

Of course, the prosocial moral aim is just a derivative of the greater aim of persistence. It is the tactic to the greater strategy. In that case, why even make this distinction? Bluntly, it is for the same reason we make the strategy-tactic distinction. There are instances in which focusing on the strategy is best, as well as cases where sticking to a tactic is best. Strategy encompasses both the intent/aim and what needs to be done to realize it. Tactics are, essentially, the micro-strategies employed to further the strategy. Let's take a mixed martial arts bout as an analogy to put it concretely. Your strategy would, of course, encompass winning, but winning in whatever way is suitable to the circumstance, ideally in accordance with the strengths and weaknesses of you and your opponent. This corresponds to the conception of persistence that is the aim of our desires. Note that it's not mere ontological persistence, but instead is the particular way life persists, adapting via reproduction (otherwise something like achieving immortality would be our ultimate goal). The tactic is taking whatever is suggested by the strategy, say winning via leg kicks for example, and employing effective means to reach it. In short, it would be the actual kicking of the legs and whatever needs to be done to reach it. This is in correspondence with a prosocial aim, as hypersociality is the most effective way for humans to persist. This analogy will be returned to in later discussions on ways to decide between strategy and tactics.

One motivation for the seemingly arbitrary distinction in moral aims is purely practical; it does well to explain cases where we simultaneously have differing moral judgments that change with slight adjustments to the moral sense under which we're viewing it. Another motivation, also practical but in a different sense, is that each aim corresponds to a different moral mode under the dual process model. A practical

prescriptive theory is really one that will be able to direct us when to think under one aim or the other. But yes, prosociality's own value is derived from its relation to the greater aim of persistence. This also means that, at the most fundamental level, morality is consequentialist. However, this does not entail that we should always use consequentialism to prescribe actions.

To give a real example before delving into the nitty-gritty, there is an example that can perhaps illustrate morality on the level of persistence versus the level of prosociality. Noam Chomsky has claimed that former president Donald Trump is “the worst criminal in human history,” surpassing the likes of even figures like Adolf Hitler (Buncombe, 2020). It is prefaced by admitting that it is “an outrageous claim... [but] it's not false.” His logic in making this claim is straightforward: “Adolf Hitler was pretty hideous – [but] he wasn't trying to destroy organised human society on earth.” This is in reference to Trump's policy on climate change and the catastrophic effects it would wreak upon humanity, or really, all of life. It is commonly known that unchecked climate change and the factors it is likely to aggravate, such as probability of nuclear war, pose an existential threat to modern society. Critically, it is known that the Trump administration knew this when making its decisions on climate policy. If we examine this, assuming the moral aim to be at the level of persistence, the strategic level, it is indeed undeniable that this constitutes a worse crime than anything Hitler did. Our best scientific understanding suggests that Trump's climate policy would in fact lead to greater doom on the level of biological persistence than anything Hitler ever did intended to do, far more so than even the Holocaust. Yet of course, even the harshest critics of Trump are likely to be hesitant in accepting the claim that what Trump did was even worse than what Hitler did. This is because the “moral aim” that we

most often assume, when judging character for instance (or more specifically, actions made by moral agents), is not concerned directly with persistence. We still can, of course, enter that frame of mind by considering the consequences accordingly. It is under a different assumption that Chomsky's claim is outrageous: the prosocial intent.

To reiterate points made in the discussions on the dual-process model and the social theory of reason, our prosociality is really grounded in moral emotions (namely things like compassion, love, and so forth) that constitute the impetus for our prosocial behavior. We do not reason out that these prosocial behaviors are beneficial for our persistence, though they largely are indeed so, of course. Due to prosociality almost always lending itself to persistence, we aren't often aware of the distinction it entails in moral judgments. The situation with Chomsky's claim, however, makes it clear. On the social level, which is primarily how we think in regard to judging both character and types of actions themselves, Hitler is far and away the worse criminal. This is because, essentially, the psychological compulsion (or moral intuition, rather) to not cause harm to others is directly violated in the case of perpetrating genocide. The genocide is the intended consequence. In Trump's case, abject devastation via climate change is just the irrelevant byproduct of keeping the highly profitable fossil fuel industry churning. Both are clearly atrocious, but, in Schopenhauerian terms, the malice necessary to allow for one to choose genocide is a greater psychological retardant than the egoism necessary to put personal gain above all else. Neither are admirable in any way though. In the social sense, we judge based on the emotions that the choosing of an action inspires (as we make decisions based on pre-existing intent, never pure reason) rather than basing them in any rational account of consequences.

Another example is our general tendency to regard social crimes as ‘moral’ irrespective of how insignificant it might be in the greater scale of overall persistence. Let’s take the adulterer for instance. If one has an affair, even when there is no consequence to the greater good nor any intended consequence for such, there is still a revulsion as it violates our prosociality. Let’s look at a case where actions offend our prosociality, but have a markedly beneficial effect on persistence. Thomas Edison, or at least the increasingly common depiction of him, stole the ideas of other inventors and took undue credit for their work (regardless of how true this actually is, let’s presume it to be so). He is also, however, responsible for many inventions that changed the world for the better, such as electric generation, motion pictures, and so on. Accordingly, there is a split in our evaluation of him dependent on what moral aim we assume; great in one sense, less than savory in the other.

Another clear difference between the persistence aim and the prosocial aim is in the existence (or not existence) of the supererogatory. To drive home the point, let’s take the real story of Hisashi Nishizawa for example. To make a long and touching story short, Nishizawa’s fiancée developed a rare form of brain inflammation that caused her to have such memory loss so as to forget the identity of her friends and family, Nishizawa included, before falling into a coma. It was almost entirely unknown if she would ever recover, but, in any case, she was unlikely to regain her memory. Nishizawa tended to her for eight years, most of those with her being in a vegetative state, while she slowly regained functionality and, eventually, her memories, before being married. At one point even the fiancée’s parents urged Nishizawa to abandon her and find someone else, but he refused. If this is not supererogatory moral behavior, who is to say what is? But on the persistence level of

moral thinking (and under the consequentialists theories that best represent it), this is either mostly irrelevant or even condemnable. The prosociality in the form of love and loyalty displayed by Nishizawa's actions is, however, abnormally superb, and so the evaluation given reflects this. Should he have done what he did? On the strategic level of moral aim, the answer is either that it doesn't matter much, or that he should not have. On the tactical level however, this was not obligatory at all, yet is evaluated more highly than the alternative of finding someone else. This also highlights how ought claims function differently on the tactical level as opposed to the strategic level. This exemplary conduct is probably what we *should* do in the prosocial aim, but is far from what is *required*. There is the possibility for morally permissible moral mistakes, for instance when charging a poor person who needs the money for a given service (Harman, 2015). Nishizawa leaving his comatose fiancée is a similarly morally permissible moral mistake. There is no such thing in the case of the strategic aim, only that which works better and does not (ignoring cares where there's no bearing on persistence, of course).

And so, which are we to evaluate more positively: the actions of Nishizawa, or the actions of Edison? On the level of character, it is clearly Nishizawa of course, but on the level of action it depends on what aim we assume. This leads us to the central question: what aim should we be thinking under, in what circumstances, and why? The answer is, as expected, not as flashy as many would prefer but probably the most accurate: it depends on the circumstance. The first step in explaining this is to point out that this question is only interesting in the cases where assuming one aim would produce a different result than the other. This is not as common in real life as contemporary theory might suggest; this

only occurs when prosocial acts don't coincide with that which is beneficial for persistence.

In most decisions we make, what is conducive to one is often conducive to the other.

CHAPTER 5

A PRACTICAL THEORY

To address a potential concern, the exact identity of the moral strategy is, in truth, not incredibly important for the discussion of this chapter. Reason has been given to believe that it is in fact persistence, but so long as there is a similar structure in how it relates to tactics, most of what will be argued should follow. So long as prosocial behavior is the tactic to some greater strategy, methods to balance them in practice should remain similar. Furthermore, as consequentialism and non-consequentialism have been shifted to occupy a position in normative ethics more resembling the two moral modes rather than their explicit theories, reference to them in this chapter will be in the sense of moral modes. And so, to reveal more about how we should balance moral strategy and moral tactics, we can return to the chess analogy. In fact, this is why chess in particular was chosen to be the example in the first place.

When analyzing chess in the modern era, there are two primary tools available to evaluate a position with: the computer and the opening database. Chess computers have long since surpassed any human's player's skill at the game and are only rapidly improving (though chess still seems to be far from a solved game). They function as a tool to evaluate positions by assigning a number value (called centipawn value) based on things like material, positional advantages, and so on. This is done by an algorithm of human design; in other words, we input our theoretical principles of what value pieces and their positions have into the computer. Then, using the calculative powers of machinery that humans could

only wish to reach, it evaluates candidate moves and the value of the positions they lead to, say, 20 moves down the line for instance. They then show the move that leads to the highest value position assuming optimal play (that part is critical). When we humans are analyzing the game, we typically consider this the ‘objectively best’ move as, relative to human skill, it might as well be so. Then there is the opening database. It is a database of hundreds of millions of games played, and when viewing a position, it displays the number of times each move was played in the position as well as the game’s result. It is, essentially, a way of seeing what the most common move in a position is, as well as how effective it is in practice.

We can see how these two tools share similarities with consequentialism and non-consequentialism, the former being the computer and the latter being the database. This is, of course, assuming that Greene’s claims about non-consequentialism being grounded in emotional responses is accurate. Keep in mind that these emotional responses themselves are evolutionary. In the same way that the computer evaluates moves by assigning value to positions based on a certain calculative algorithm, consequentialism urges us to evaluate decisions by assigning some value (such as total utility) to the consequence of the decision based on some mistily defined method. Reasoning/calculation is critical to both. For the consequentialist, reason is what reveals the expected outcome, and then it is also expected (in some theories) to assign a value to that outcome. Of course, reason only does not actually generate any value as that is not within its capacity, but so goes the myth. The computer, similarly, gives its evaluation not via raw calculating all the way to a checkmate (when it does, it displays how many moves until checkmate, not a centipawn value), but instead using the algorithm already programmed in it. The opening database resembles our

emotional responses in that it is the collection of all games previously played. Our emotions are constructed from, essentially, what worked in the past and is carried on by our genealogy. The database displays what is essentially the most natural instinct in a given position via the most commonly played move. Using the database for suggesting moves involves weighing various factors such as the total number of times a given position has been reached, how many times the move has been played, and what game result it leads to. What really leads to their similarity is their shared basis in learning from trial and error.

Despite the computer giving us the so-called ‘objectively best’ move, how we use these tools (as that is, in fact, all they are) in practice is quite different. When deciding what move to play in a singular position, following the computer's suggestion is surprisingly unadvisable. For instance, it is not so uncommon to find yourself in a position where the most intuitive move is a quiet, but principally solid one. Even if it does not generate a decisive advantage, it is a safe development. Then the computer will suggest either a seemingly pointless move, or even a seriously principally unsound one, such as advancing your king into danger. If a human player were to play this move, it would surely lead to disaster, as even though it is defensible with the perfect play of a computer, the human mind simply will not be able to find and play the following moves that justify the first. This is particularly true if it leads to an unfamiliar position. Put shortly, the computer can recommend one to turn an otherwise calm position into a very sharp one. Sharp positions are those where there are only a handful of good moves (or maybe even just one good move) with other moves leading to a decisive disadvantage. The computer, of course, has no problem navigating sharp lines, as it has already calculated all potential responses and knows that with optimal play, it is safe. Human players, however, do not play like this.

Even the highest-level grandmasters play primarily using intuitions and principles, only using calculation in certain positions. Because of this, unintuitive moves that create sharp positions are only played when knowing exactly what line will lead to an advantage.

Consequentialist thinking shares this same weakness of turning otherwise safe situations into likely disasters. Sidgwick entertained the idea that it would perhaps be better for utilitarianism to only be used by the elite that have the correct calculative abilities to navigate the equivalent of ‘sharp lines’ in chess (Sidgwick, 2007). This is an extremely suspicious proposition. The famous 19th century Russian novel *Crime and Punishment* by Fyodor Dostoevsky was written basically with its entire purpose being to illustrate why this is a horrendous idea. The plot centers around Raskolnikov, the main character, and his axe-murder of a wealthy, but rotten, inconsequential, and overall disliked pawnbroker (Dostoevsky, 2017). His justification is, essentially, a consequentialist one. His logic was centered around the idea of a ‘Napoleon,’ in the same vein of the Nietzschean *Übermensch*; one that can transcend common morality in order to do greater things. Raskolnikov, who was living in poverty, reasoned that if he killed the pawnbroker and took her money, he would be able to go on and realize his great potential as a ‘Napoleon.’ The death of the pawnbroker would cause no great harm, and indeed would be used as a means for much greater achievements. The result is, of course, much different from the theory. Raskolnikov fails even to secure anything of significant value after the murder, and is so haunted by his actions as to hide away the little he did gain. The psychological pain and guilt that is brought about by the murder ultimately renders him totally unable to carry out his consequentialist calculations (which were pathetically self-centered in the first place), and results in him repenting and facing a prison sentence for it.

It is in this that the essential danger of consequentialism is exposed. Raskolnikov's error was, bluntly, a complete lack of self-awareness which consequentialism can enable to incredibly dangerous ends. In the perfect world, in the world of computer calculation, perhaps Raskolnikov's murder would end up being justified. This is, however, neglecting the fact that this plot would really only have worked if Raskolnikov was a total psychopath. And indeed, in reality low-anxiety psychopaths make for very consistent consequentialists, due to their absence of prosocial emotions (Koenigs et al., 2012). But, as shown in the novel, Raskolnikov is not a psychopath and, for all his flaws, is in possession of a very deep conscience as demonstrated by the depth to which he suffers psychologically following the murder. And herein lies Dostoevsky's compelling implication: this depth of feeling that led to Raskolnikov's 'failure' is not a failing at all. On the contrary, it is his redeeming quality. As stated before, this 'greater good' scenario is only possible if the one enacting it has a serious antisocial disposition (in the sense of ASPD, also called psychopathy). But it is not in the interest of the greater good to be a psychopath, even worse so in the case of Sidgwick's 'Napoleon'-esque utilitarian cabal (just ask yourself if you want a group of psychopaths performing utilitarian calculus in your interest). In fact, calling someone a psychopath is sometimes used to imply that the person in question is inherently evil, or at least morally suspect. This is, of course, not truly a justified implication, but the connection is not totally baseless. After all, being a psychopath would potentially mean a lack of the appropriate aim to be a moral agent in the tactical sense; the prosocial aim. Returning to Raskolnikov's thought process in killing the pawnbroker, we are hinting towards a crucially neglected moral feature by consequentialist theories. It is

the asking of this question: “Can I carry this out *psychologically*?” We will return to this later.

The larger implication from this is to demonstrate that what looks like a strategically sound move (killing the pawnbroker) can be, and often is, subverted by our own tactical strength (not being antisocial). Furthermore, the solution is *not* to promote the discarding of prosocial emotions in this situation, as that is often a foolishly naive and self-centered endeavor that neglects to factor in one’s own overall strengths. At least, *Crime and Punishment* paints a picture of this. To go back to the strategy-tactic discussion in a mixed martial arts bout, it is something of the equivalent to jumping on a guillotine when you are not skilled in them. To translate that, it is a bad idea because if you fail at successfully performing the guillotine (a submission maneuver), your opponent will end up in a dominant position on top of you. It appears a good strategic decision as it has the potential to end the fight right there, but is a terrible tactical decision considering you are not suited to it, and failure puts you in a losing position when you would otherwise be at an advantage. So overall, attempting the submission is a bad decision in the same way killing the pawnbroker for her money is a bad decision.

To give another similar situation, let us consider the following. Think of a friend that has an overall trusting nature. This friend has invested their time into developing a relationship with an individual who, unfortunately for them, will betray and abandon them. In the aftermath, this friend confides that they are now doubting whether it was right to trust or not in the first place, that maybe it would be better to keep others at a safe distance. What are we to recommend them to do? On the tactical level, their trust was no mistake. It is highly conducive to, if not outright crucial, the interest of prosociality to be trusting and

trustworthy in turn. Strategically however, in this situation it was a mistake; this was simply a bad decision as demonstrated by the bad results (assuming these bad results were a foreseeable possibility, as is often the case). That does not, however, change the principle one should operate under on the tactical level.

It is just as possible to perform mistakes in the opposite direction, however. One is in a previously mentioned case: the comparison between Trump and Hitler by Chomsky. That is a case where the strategic sense is more important than the tactical one due in part to its particular features. This is because impersonal existential threats, such as climate change or nuclear war, go far beyond the reach of prosociality. We can abide by prosocial principles and still be driven into climate disaster. In fact, that itself may be the very issue that caused the problem. Existential threats are a threat to prosociality itself; it's hard to behave prosocially when there is no environment to be prosocial in. And so, in the consideration of the prevention of extinction (or in general, harm to our biological flourishing), our moral mode should be firmly consequentialist.

A further, general rule to follow in deciding when to act on the strategic or tactical level is in the matter of scale in the issue. As the consequences of an action grow larger, there tends towards a greater impact on persistence rather than sociality, as seen for example with existential threats. Our general behavior seems to conform with this too, as the influence of consequence is higher in moral judgements with high stakes, high impact decisions (Dubljević et al. 2018). Even here, however, psychological limits need to be taken into consideration, lest we inadvertently justify a doctrine like Stalinism.

Greene suggests that we rely more on consequentialist modes of thinking in novel scenarios which our evolutionarily founded intuitions are not suited for. This is true, but

there are also those cases in which we already have appropriate intuitions, but they are not engaged due to the novelty of the circumstances, despite being functionally identical. For instance, this is prevalent in online social interactions. In the online world, our social intuitions are not engaged in the same manner as in person interactions, as many important emotional triggers are missing. This leads to an increase of very antisocial behavior, perhaps the most notable of which being the phenomenon of ghosting. Ghosting refers to sudden and complete cessation of all communication with someone, and disappearing ‘like a ghost,’ usually at the mere whim of the one who ghosts (simply because they do not want to talk to the other, for instance). In real life situations, this is highly morally impermissible for reasons that can be easily inferred. In those situations, our emotions are more readily engaged, and there is a higher sensitivity to the basic respect afforded to another in social communication. This is absent in an online environment. The most shocking element, however, is that ghosting is (currently) largely codified as morally permissible behavior. The solution does not seem to be to turn on ‘manual mode’ and give a consequentialist account for the impermissibility of ghosting. It simply is not the tool for the job, as social consequences are often nigh impossible to predict when only looking at the results of actions. Rather, it seems more natural and effective to find ways to import our already existing intuitions to the functionally identical circumstances of online social interactions.

On a more general level, for most of the decisions we make in life, it is safe to operate under the tactical level. The fact of the matter is, few of us often find ourselves making decisions that are beyond prosociality. The greatest dilemmas we face in our actual day to day living are firmly confined to social matters. It is the very reason why our emotions are such that we tend to be guided by them in social situations. All of our

ancestors had to constantly behave such that cooperation was ensured. Only a rare (and unfortunate) few ever had to face situations like those of the trolley problem, but moral emotions such as compassion, honesty, loyalty, and so on are ironed out quite strongly in prosocial matters. The saying “No need to reinvent the wheel” is applicable here.

To extend this, it is similarly true in matters of agent evaluation. As people simply are not as likely to be in positions that transcend prosociality to become direct matters of persistence, dispositions are more weighted to their effectiveness in social contexts. This is why we view qualities such as empathy and compassion as sublime virtues, while cold emotional detachment is not one at all, despite serving consequentialist decision making well. All of the concerns raised in the discussion of *Crime and Punishment* also apply to agent evaluation. Despite ‘failing’ in the consequentialist sense, we still have a much more positive evaluation of Raskolnikov’s character than if he succeeded in the total control of his emotions such that he really became a ‘Napoleon.’ Agent evaluation should, therefore, be non-consequentialist, and almost solely concerned with prosocial disposition. And yes, that basically entails that the egoist is the lowest one can get in agent evaluation. Truly antisocial individuals are ‘off the scale,’ so to speak, as they do not even have the necessary moral aim to qualify themselves as susceptible to moral evaluation on the tactical level. In the same way that we would not morally judge a bear poorly for its antisocial behavior, such as killing villagers, so too does it apply to those without any prosocial aims. To put it in the concrete, imagine you are trying to reason with a psychopath, explaining why, for instance, killing might be wrong. It would be impossible to justify it from the tactical level. If you ground the argument in prosocial emotions, for example saying, “Killing is wrong because it’s good to care for and respect others,” and they respond with “But I do not care

about others in the slightest”, that is essentially all there is to it. You could move to arguments from persistence, but they would be immune to arguments from prosociality; reasoning that way will never bind them morally, because they lack the moral qualifications. Egoists, on the other hand, do have the qualification; they are just so incredibly mistaken to the point where they believe that behaving selfishly is in fact the best course of action in the interest of the whole.

To further extend agent evaluation as non-consequentialist, an important entailment is that character development is also non-consequentialist. What this means is that moral development in an individual does not consist in, essentially, a greater allowance for consequentialist modes of thought to override non-consequentialist ones. It is not even a greater precision of consequentialist calculation itself. Rather, it lies in strengthening our prosocial emotions, and increasing our proclivity to act on them in opposition to egoist or antisocial motivations. In layman’s terms, selflessness over selfishness.

And so to tie this all back, what should be the answers to the moral dilemmas brought up throughout this project? Let’s begin with examining the trolley problems. It should be clear by now that the correct decision on the strategic level is to kill the one to save the five, and the correct tactical decision is the opposite. So, how are we to balance them? Let us begin with the footbridge variation, as that is the more interesting and pertinent case with a much clearer conflict. Our prosocial emotions say that it is wrong, the persistence-based calculus says it is right. Here is where there is an original injection: psychological limitations. The principle that ought implies can is attributed to Kant, spoken in the sense that if an agent is morally bound to perform some action, they must be logically capable of performing it (Kant, 1908, p. 473). This can, however, be extended to

encompass a *psychological* can as well, and the demandingness objection was criticized with this very point (Corbett, 1995). Consider this for a moment: is it possible for one to recognize what one ought to do, yet be unable to will it? If this is true, then there is the following question: if they are unable, ought they instead be able to will it? Let's input all of this into the footbridge case. Is it possible for one to recognize they ought to push the man over, yet be unable to do it due to psychological limitations, i.e., an overwhelming compulsion to avoid committing the violence? And if so, ought they then have the overwhelming compulsion, or not have it? The answer to the first is yes, but the second needs a more careful evaluation.

To pull another case from literature, consider this scene from *Adventures of Huckleberry Finn* by Mark Twain. In the moral climax of the novel, Huck's companion, the runaway slave Jim, has been captured and is going to be sent back to his original masters. Huck stops to consider if this is something he should fight against. He reasons that he has done something wrong by, essentially, stealing someone's property, and that he should write to the original masters to alert them to Jim's whereabouts. He writes a letter detailing just that, and it results in this passage: "I was a trembling, because I'd got to decide, forever, betwixt two things, and I knowed it... 'All right, then, I'll go to hell'- and tore [the letter] up" (Twain, 1985, p. 321). Huck, whilst fully and earnestly believing that what he feels is wrong—so wrong that it will land him in hell, in fact—chooses loyalty to Jim over *eternal damnation*. The part to highlight from this, however, is not how Huck was mistaken in believing slaver's rights to be moral. It is that Huck, due purely to his own prosocial conscience, was *unable* to will what he recognized as the right decision.

Recall the discussion in chapter 3 on the moral status of emotions and the will itself. They are themselves non-moral. Nonetheless, they can limit moral decision making by rendering certain choices impossible. This is what is occurring with Huck when he chooses to help Jim; what appears to him as the morally correct decision is essentially blocked as a practical choice by sheer strength of intuition. Perhaps this is just irrationality. But perhaps that is not a bad thing given our moral aims. Let us go back to the trolley. Could it not be the case that when on the footbridge, one is able to correctly see that, in this situation, one should kill the man, yet find themselves unable to will it? Yes. Could it be, that even though one should will it, one still ought to be unable to will it? Also, yes. How is this possible? Think of it as such: given two options before you, options A and B, both of which are within your capacity to will, A furthers the moral aim more than B. You ought then to choose A. But then, what if A simply cannot be willed, due to an overwhelming intuition? Then, of course, due to ought implying can, there is no obligation to choose A. Whether it ought to be *possible* to will A, however, is a matter beyond just the choice between A and B. That is calling the intuition itself into question, which affects more decisions than just the one given here. In other words, is it in the aim of persistence for this moral intuition against personal violence to be strong enough to block the conclusion produced here by consequentialist thinking? This is a truly difficult question to answer, but there is at least reason to believe this might be true. Are we better off having someone with generally sound intuitions that mess up in this situation, or someone with more aberrant intuitions that pass this situation, but lead them to fail others? It is hard to answer yes or no. It might be the case that situations where killing is in the interest of the greater good are rare enough that the intuition is effective and appropriate. For every trolley problem, there might be other,

poorly guided decisions that are prevented by the intuition, such as with Raskolnikov. And in those, it might be better not to be psychologically able to take the ‘computer line,’ in equivalent chess terms.

But ultimately, that is a matter of the intuition’s justification, in the same way that happiness was judged based on how appropriate it was in the interest of achieving some aim given the stimulus. To make the discussion much simpler, is it in the interest of persistence to have overriding intuitions such that they deem killing someone violently as inappropriate, even when it is five to one? Is the intuition so critical on the tactical level that it cannot be sacrificed to the better situational strategic decision? It is difficult to say, but it seems to be a plausibly good tactic that our psychology does not make an exception here. It really comes down to judging the intuition on risk versus reward. What we are really asking is whether choices like option A be open in other situations as well. When the issue at hand is killing, the risk is high in the sense that it tends to lead to a ‘sharp line.’ Even envisioning the footbridge case to be a real-life situation can illustrate this issue. It would not even seem unlikely that pushing over the man would not stop the trolley, leading to his death and the five on the track as well. In this case, attempting to save the five resulted in six dying instead. Is the potential to save the five worth the risk of creating an even worse outcome by the pointless killing of another? That is a very tough call that requires careful evaluation of the circumstance and of those like it.

Make no mistake though; because it is a thought experiment with a certain outcome, the simple answer to the question of what one ought to do in the footbridge case is to kill the man, but that is only identifying what the strategy demands. There is still the matter of

whether we ought to be psychologically able to or not; in other words, whether we should act in accord with the strategy, or the tactic.

This process can be copy pasted to the other dilemmas. Singer's obligation to distant others gets the same answer: we ought to, but ought we be psychologically able to? In reality, most people simply are not psychologically able to bear such a commitment to distant others, and to the extent they are not, there is no obligation to hold such a commitment, as ought implies can. It could very well be the case that, in the interest of persistence, prosociality should be extended in this radical way to all others. If that is the case, then what we ought to do in practice is give as much as possible within psychological limits, while also expanding the limits themselves. Or, of course, maybe it works out that our current arrangement is already more ideal than that, due to charity and generosity being more effective to those in proximity (though this seems more likely than in the footbridge case). Is it worth the risk of removing our effective but biased preference towards those close to us, both in geographical and emotional terms, so as to extend aid equally to all? The answer is, again, not simple. But that is the question to ask in addressing whether strategy or tactic should be followed.

Even Kant's murderer at the door can be evaluated with this method. This one might be more useful as the result of the process is very clear. If capable of lying here, ought we to lie? Yes. Ought we be psychologically able to lie here? Yes. This is because the reward for enabling lying of this sort seems to far outweigh the risks. Situations where lying is beneficial are much more common than with killing, and at a far lesser risk. Though, even lying can fall into the same problem that killing has however, in that one may overestimate

their psychological capacity for lying, resulting in a much worse result than were one not lie.

This is the general process that can be applied to any conflict between strategy and tactic. It is, in short, evaluating whether the prosocial intuition that blocks the more circumstantially effective strategic calculation is justified in doing so. It will not always be clear or easy to find the answer, but it is better than preferring the ‘manual mode’ or the ‘automatic mode’ on principle, as is suggested by consequentialism and deontology respectively. And there are general signs to look for in aiding that judgement, such as the risk in failing versus the reward in succeeding with the strategic choice, as well as the impact that a given prosocial intuition has on overall persistence, and so on. That is why, for instance, allowing for stronger intuitions against killing is a better idea than in lying.

To conclude, there can be general advice for how one ought to approach each moral mode. On the strategic level, the most valuable thing one can do is have self-awareness of what one’s psychologically limits *are*. For example, if you are contemplating becoming a ‘Napoleon,’ you should be aware of how your intuitions simply are or are not suited to it (with virtually any well-adjusted person falling into the latter), and act accordingly with the risk in mind. On the tactical level, it is more important to focus on what those limits *ought* to be. It is a matter of ensuring that one’s intuitions are such that they are, to the greatest extent possible, disposed to effective prosociality. It is important to point out that we can in fact alter our intuitions; they are not innate and immutable. That is the essence of virtue under the Aristotelian sense. When these components are combined, it can be claimed that morality is truly about choosing the best path relative to one’s own limits (with

consequentialist thinking) while also expanding and developing those limits (improving non-consequentialist intuitions).

The distinction also maps well to professional and personal life, with strategy guiding the former and tactics the latter. While the best thing you can do professionally is probably something like dedicating your life to humanitarian aid, fighting climate change, and so on, there are psychological limits that determine whether that is truly a realistic choice. For most people, it simply is not. Those are professionally ideal, they are ‘the best’ under strategic morality, but often personally unrealistic due to the nature of our tactical strengths. The solution then is to find the highest impact goal that is within one’s psychological limits (and natural abilities). Simultaneously, however, tactical morality compels one to improve those psychological limitations. It is, in short, the cultivation of virtue.

CHAPTER 6

CONCLUSION

In the introduction, the need for a more practical normative ethical theory was established. Is the theory outlined here practical? First, to characterize it. This is a dual-process morality, though with a fundamentally consequentialist core resulting from the conception of ‘ought’ as identifying the most effective means to an end. However, non-consequentialist results can be, and often are, produced from the strategy-tactic distinction, leading to a more complete and realistic account of moral life, particularly in the social sphere. Even with this, there are not necessarily clear-cut answers for certain moral dilemmas. Nevertheless, the methods described here seem far more applicable in practice than either a purely consequentialist or a purely non-consequentialist theory. In truth, no matter what theory one adheres to, life will involve the usage of both. To the extent the strategy-tactic distinction accounts for this, there is a distinct advantage in the areas where traditional normative theories fail.

REFERENCES

- Alexander, L. (2021). Deontological Ethics. In Moore, M. (Ed), *The Stanford encyclopedia of philosophy* (Summer 2021 ed.). Stanford University (Summer 2021 ed.). <https://plato.stanford.edu/entries/deontological-ethics/>
- Blackburn, S. (2003). *Ethics: A very short introduction* (Vol. 80). Oxford University Press.
- Buncombe, A. (2020, November 3). Trump represents worse threat to humanity than Hitler, claims Chomsky. The Independent. Retrieved November 20, 2021, from <https://www.independent.co.uk/climate-change/news/trump-climate-change-noam-chomsky-book-interview-hitler-robert-pollin-b1374789.html>.
- Cholbi, M. (2009). The murderer at the door: what Kant should have said. *Philosophy and Phenomenological Research*, 79(1), 17-46.
- Corbett, B. (1995, October). Moral Obligations to Distant Others. Retrieved November 20, 2021, from <https://web.archive.org/web/20080325124519/http://www.webster.edu/~corbetre/philosophy/moral/others/distant.html>.
- Dostoyevsky, F. (2017). *Crime and punishment*. Oxford University Press.
- Dubljević V, Sattler S, Racine E. Deciphering moral intuition: How agents, deeds, and consequences influence moral judgment [published correction appears in PLoS One. 2018 Oct 25;13(10):e0206750]. *PLoS One*. 2018;13(10):e0204631. Published 2018 Oct 1. doi:10.1371/journal.pone.0204631

- Greene, J. D. (2008). The secret joke of Kant's soul. *Moral psychology*, 3, 35-79.
- Greene, J. D. (2014). Beyond point-and-shoot morality: Why cognitive (neuro) science matters for ethics. *Ethics*, 124(4), 695-726.
- Harman, E. (2015). Morality Within the Realm of the Morally Permissible. *Oxford Studies in Normative Ethics*, 5, 221-224
- Hume, D. (2003). *A treatise of human nature*. Courier Corporation.
- Janaway, C., & Robertson, S. (Eds.). (2012). *Nietzsche, naturalism, and normativity*. Oxford University Press.
- Jarvis Thomson, J. (1985). The trolley problem. *Yale Law Journal*, 94(6), 5.
- Joyce, R. (2021). Moral Anti-Realism. In Zalta, N. (Ed), *The Stanford encyclopedia of philosophy* (Summer 2021 ed.). Stanford University (Summer 2021 ed.).
<https://plato.stanford.edu/entries/moral-anti-realism/>
- Kagan, S. (1990). *The limits of morality*. Clarendon.
- Kant, I. (2012). *Kant: Groundwork of the metaphysics of morals*. Cambridge University Press.
- Kant, I. (1949). On a supposed right to lie from altruistic motives. *Critique of practical reason and other writings in moral philosophy*, 346-50.
- Kant, I. (1908). *Critique of pure reason*. 1781. Modern Classical Philosophers, Cambridge, MA: Houghton Mifflin, 370-456.
- Koenigs, M., Kruepke, M., Zeier, J., & Newman, J. P. (2012). Utilitarian moral judgment in psychopathy. *Social cognitive and affective neuroscience*, 7(6), 708-714.
- Mercier, H., & Sperber, D. (2017). *The enigma of reason*. Harvard University Press.
- Mill, J. S. (2000). *Utilitarianism*. (R. Crisp, Ed.). Oxford University Press.

- Nussbaum, M. C. (1999). Virtue ethics: A misleading category?. *The Journal of Ethics*, 3(3), 163-201.
- Sartre, J.-P. (2007). *Existentialism is a humanism* (C. Macomber, Trans.). Yale University Press.
- Scheffler, S. (1986). Morality's demands and their limits. *The Journal of Philosophy*, 83(10), 531-537.
- Schwitzgebel, E., & Cushman, F. (2012). Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers. *Mind & Language*, 27(2), 135-153.
- Schwitzgebel, E., & Rust, J. (2013). The moral behavior of ethics professors: Relationships among self-reported behavior, expressed normative attitude, and directly observed behavior. In *Philosophical Psychology* (Vol. 27, Issue 3, pp. 293–327). Informa UK Limited. <https://doi.org/10.1080/09515089.2012.727135>
- Sidgwick, H. (2007). *The methods of Ethics*. Hackett.
- Singer, P. (1972). Famine, affluence, and morality. *Philosophy & public affairs*, 229-243.
- Twain, M. (1985). *The Adventures of Huckleberry Finn*. 1884. Available from a variety of publishers.
- Wise, R. A. (2002). Brain Reward Circuitry. In *Neuron* (Vol. 36, Issue 2, pp. 229–240). Elsevier BV. [https://doi.org/10.1016/s0896-6273\(02\)00965-0](https://doi.org/10.1016/s0896-6273(02)00965-0)

BIOGRAPHICAL INFORMATION

Gage Martinez entered university in 2018 and is graduating in 2021 with an Honors Bachelor of Arts in Philosophy along with History and English minors. His research has been primarily concerned with ethics, which has led to ranging from philosophy to classical literature to geopolitics. He also has other academic publications in those topics. After graduation he plans to enlist in AmeriCorps, a federal volunteer program, and after that enter the non-profit sector.