

MORAL RELATIVISM: A SORT OF DEFENSE

EVA GOODWIN
MOUNT HOLYOKE COLLEGE
PHILOSOPHY DEPARTMENT
THESIS FOR HONORS CONSIDERATION

For Dad

Acknowledgments

Many, many people have been an important part of this project. I would like to thank first and foremost my advisor, James Harold, whose intellectual and personal support and encouragement have been invaluable. This project would never have been completed without him. The other faculty of the Philosophy Department have also been marvelously supportive and helpful, always ready to answer my questions and make suggestions. Additionally I would like to thank Tom Wartenberg and Bill Quillian for being my readers, as well as my academic mentors these four years. My friends have been very patient with me, and I am incredibly grateful to them for putting up with my stress, my musings, and my limited appearance at social occasions this year. I would especially like to thank Carmen Guhn-Knight and Katie Omberg for proofreading and offering critical feedback. I am grateful to Tracy Zhu for her love and encouragement. And last but certainly not least, I thank my parents for instilling in me an intellectual curiosity and a sense of ambition, for all the long talks about this project, and above all, for believing in me and rooting for me every step of the way.

Table of Contents

Introduction	5
Chapter One	8
1. Moral Relativism	10
2. Criticisms of Moral Relativism	20
3. Conclusion	29
Chapter Two	31
1. A Word on the Word ‘Reason’	32
2. Nagel, Introduced	33
3. A Subjective Account of Reasons: Harman	38
4. Williams Lends a Hand	41
5. A Second Look at Nagel	47
6. Conclusion	52
Chapter Three	53
1. Metaethical, not Normative	54
2. Justice as Fairness	56
3. Interpretation	68
4. Implications	73
5. Conclusion	76
Chapter Four	79
1. Harman’s Objective Constraints	80
2. Consistency	82
3. Impartiality	86
4. Full Information	90
5. Conclusions	94
Appendix A: Nagel’s Logical Formulation of Objective Reasons	99
Works Cited	101

INTRODUCTION

How do we know ethically what is right and what is wrong? Despite all the scientific progress made since the Enlightenment, little progress has been made toward discovering objective laws of morality. Nevertheless, people make ethical claims all the time. Last year, for example, I was watching on television a political debate about the ethics of same-sex couples having or adopting children. Liberal Guy was spouting all the normal liberal stuff: equal rights for gays and lesbians; benefits of a loving home outweigh any potential disadvantages; they'll have children anyway so for the children's benefit it should be legal; etc. Conservative Guy, however, went straight to the root of the problem: "All children," he said, "have a *natural born right* to both a mother and a father." With this statement, Conservative Guy effectively rendered Liberal Guy's arguments irrelevant; if Conservative Guy was right, it wouldn't matter whether anything that Liberal Guy said were true. It could *all* be true. But it wouldn't outweigh the *fundamental right* that all people have to both a mother and a father.

I smelled a fish. Who was this guy, and who made *him* the authority on natural rights? And even if someone *did* give him the authority, how did *that* person have the authority? For that matter, how does any human being have the authority to determine what all people's natural born rights are? I determined to

investigate. My question was, how do we know what our natural rights are, and who gets to determine them?

Quickly, the scope of my project went beyond that of just natural rights. The question instead became, how do we know anything definite and universal about morality? That is essentially the question with which I embarked on this project. Since I was skeptical about the possibility of any kind of universal morality from the outset, I leave the burden of proof on universalism. Thus, my thesis investigates whether universalism is possible.

Beyond that, my thesis investigates whether there are any *objective constraints* on moral reasoning (moral reasoning is what we all engage in when we ask ourselves moral questions or make decisions about morality). This is more of a procedural question than a substantive one; rather than asking whether universalism is possible based on any specific principles (such as natural rights), this asks whether there is a right or a wrong way to go about thinking morally and making moral decisions. This question is about whether there are any universal *ways* in which individual people must think about and engage in morality. Clearly, this is related to the question of the possibility of universalism, because if there *are* any universal moral principles, they obviously provide constraints on moral reasoning. If there *aren't* any universal moral principles, there may still be objective rules for how to reason morally.

Thus, the first task I undertake in this thesis is to investigate whether moral universalism in general can stand on its own two feet. To do this, I use Gilbert

Harman's theory of moral relativism as a wall that universalism has to try to break down. Judith Thomson and Thomas Nagel are my universalist voices. Once I've shown that the relativist wall remains standing, I look at an alternative to universalism in John Rawls' and Ronald Milo's contractarian constructivism, which suggests that though there might not be any pre-existing moral principles, there might be an objective procedure for arriving at moral principles. I investigate whether contractarian constructivism does in fact provide a non-universalist argument. I argue that it doesn't: in Chapter Three, I try to show that constructivism is essentially similar to Nagel's theory, which means it fails in the same ways Nagel's theory fails. In Chapter Four, I show that, even if my interpretation of Rawls' constructivism is unfounded, and even if I grant that constructivism is essentially different from Nagel's theory, it is still problematic. Interestingly, the problems with it at this level end up being problems for Harman's relativism as well.

Thus, in the end, my thesis attempts to settle once and for all the question of what, if anything, can we know objectively about morality.

CHAPTER ONE:
MORAL RELATIVISM

I am skeptical of moral universalism. Universalism strikes me as a somewhat arbitrary tactic for people with ardent moral beliefs to attempt to show that those beliefs are objectively true. Universalists start out with specific beliefs about morality, and then they set out to prove them. Moral relativism, on the other hand, is much more objective even in its conception, for relativism has no prior moral beliefs and sets out to prove none. Because I consider relativism to be more neutral than universalism, I begin my thesis by introducing a theory of moral relativism (Gilbert Harman's theory, to be exact), which is my theoretical wall in this thesis, and then I see whether there are any satisfactory universalist wrecking balls which can knock down that wall. In this chapter, after introducing relativism, I briefly go over some preliminary objections to it before moving on to more substantial opposition from Thomas Nagel in Chapter Two.

I should start by defining what I mean by universalism and relativism. It is difficult to define universalism, since it applies to many different (and often conflicting) ethical theories (e.g., Kantian and utilitarian). *In general*, however, any universalist moral theory assumes or argues that moral statements are objectively true or false. An "objective truth" is independent of time, individual opinions, beliefs, intents, and circumstances. For example, a statement such as "it

is wrong to kill another human being” is objectively true (or false) according to moral universalism. Human rights such as the right to freedom of speech can also be formulated as moral statements, such as “all humans have the right to freedom of speech.” Such statements are touted as objectively true statements, and thus, are universal rights. Some approaches to universalism find one law or maxim that is (supposedly) objectively true, and the rest of morality (supposedly) derives from that law. For example, a utilitarian might have a single maxim that states that an act is morally right if and only if it causes the greatest good for the greatest number of people. All moral decisions are then derived from that single objective truth. To put it somewhat differently, universalism puts objective constraints on morality through such universal laws, in that it objectively limits how we can think about and act out morality. Kant’s categorical imperative, for example, constrains us morally in that we are objectively required not to use human beings as a means to an end.

Moral relativism rejects such a universal and objective understanding of morality. In moral relativism, no moral statements are objectively true or false; rather, they are true or false only *in relation to* the intents, desires, beliefs, practices, or agreements of certain people or groups of people. Moral relativism is not the same as moral skepticism: it does not claim that no moral statement can be true. It merely rejects the universalist claim that the truth of a moral statement can be objectively determined for all people at all times. In the next section, I give a detailed account of Gilbert Harman’s version of moral relativism. There

are other versions of relativism, but for the purposes of this thesis, I focus on Harman's version as a model of relativism.

1. Relativism

The clearest way of giving an account of relativism is by defining it in opposition to universalism. While universalism asserts that there are (or at least potentially can be) objective, absolute truths about morality, relativism asserts that there are not. As defined above, according to moral relativism, moral statements are not objectively true or false; rather, they are true or false only in relation to the intents, desires, beliefs, practices, or agreements of certain people or groups of people. According to relativism, the examples I used above would not have universal truth. In Harman's version, relativism also implies that there are no objective universal *reasons* for obeying any particular moral principle. (Universalism implies that there *are*.)

In order to clarify this position, Harman invokes an analogy of motion. Motion is not an absolute. Nothing can be said to be moving if it is by itself in a vacuum with no other existing thing to which to relate it. Motion is only determinable in relation to something else, some spatio-temporal framework. As there are a multitude of such frameworks, all of which are equally viable, no single framework can be "objectively privileged in correctly capturing" motion.¹ Similarly, morality is not an absolute. Something isn't simply right or wrong.

¹ Harman (2000), 41-42.

Morality is only determinable in relation to a framework, and this framework is a set of agreements established by and among people and groups of people. This set of agreements constitutes a moral system. Since many such frameworks exist, no single one of these moral frameworks can be “objectively privileged in correctly capturing” morality.

In his paper “Moral Relativism Defended,” Harman formulates a moral example of relativism by inventing a “contented employee of Murder, Incorporated” who sees nothing wrong with killing anyone outside of his company. Harman argues that trying to convince him of the Golden Rule “would imply that our own moral considerations carry some weight with him, which they do not.”² He would simply shrug off our attempts at persuasion. According to moral relativism, the Golden Rule is not rationally required or even necessarily justifiable for all human beings; rather, it is an example of the moral laws of *one moral system* among others. The Murder, Inc. employee is not necessarily wrong for believing that hypothetical thought experiments about “how I would want to be treated in such a situation” are irrelevant to his moral views. Since he is not part of a moral system which contains the agreement that one ought to obey the Golden Rule, he is under no obligation to obey it. If I participate in a moral system in which it *is* morally required to consider the “how would I feel if...?” question, then I can probably justify lack of its consideration being wrong *in*

² Harman (1975), 5.

relation to that moral system. The Golden Rule is not, however, a universal and objective moral principle.

Similarly, I take the contemporary debate around abortion to be a perfect example of conflicting moral systems. Those who consider themselves to be pro-life or anti-abortion consider a fetus to be a person. Thus, abortion essentially kills a person, and since they consider killing a person to be wrong, abortion is clearly also wrong. For those who take a fetus to be as much a person as a living adult, it would in fact be inconsistent not to be anti-abortion. Many people on the pro-choice side of the abortion debate do *not* think a fetus is a human being. Thus, having an abortion is not equivalent to murder. If the fetus is not considered to be a human being, there is room for other values (such as the right of a woman to have autonomy over her own body) to have salience. A person who believes that a fetus is a human being and a person who believes that a fetus is not a human being really don't have much to say to each other, since they are operating within completely different conceptions of a fetus. Thus, neither of their resulting moral statements that "abortion is wrong" and "forbidding abortion is wrong" is objectively right, and neither is objectively wrong; they are each right in relation to their respective, yet incompatible, moral frameworks.

By extension, any moral principle can be shown not to be universal. Suppose it seems that there is such a principle; suppose every single person in the world, all seven billion of us, agrees with and has a reason to agree with this moral principle. This principle would then be descriptively universal, because it would

be a fact that all people agreed with it. This would *not* make it objectively universal, however. Suppose that tomorrow, a baby is born who in twenty years decides she doesn't agree with this moral principle, that she has no reason to agree with it, and that she instead has many reasons to have other beliefs instead. As long as her moral principles are logically consistent with her own desires, intents, motivations and beliefs, and as long as she is rational and fully informed, then, according to Harman, she is not objectively wrong. Notice that here I have discussed morality according to the idea of reasons. There can be no objective truth to moral statements because there are no universal reasons to believe that any particular moral view is true (or if there are, they are only incidental, not rationally required). I will follow up on this idea in Chapter Two.

The fact that there are no objective moral principles does not, however, mean that anyone can run around lying and stealing, and use moral relativism as an excuse, saying "you can't tell me it is wrong of me to do this, because I do not share your reasons." Harman's view is metaethical, not political. He does not argue that legal systems and theories of justice are unjustifiable or conceptually misguided, merely that they are based within specific moral systems which are not objectively true. He *does* hold that there may be ways in which a moral system could improve itself by recognizing internal inconsistencies and incoherence. For example, the Declaration of Independence from 1776 states that "we hold these truths to be self-evident, that all men are created equal." At that time and for almost a century following, however, many black men and women

were enslaved, and thus were *unequal*. At some point, recognizing their own moral inconsistency, people may have started to think, “well black people are still ‘men’ according to our broad conception of ‘men’ as ‘humanity,’ and we do hold these truths to be self-evident, that all men are created equal, so I guess enslaving black people might be inconsistent with our beliefs of equality.” Similarly, when women were campaigning for the right to vote, people may have thought “well, according to our belief that all people are fundamentally equal, and since women *are* people, they probably ought to have the same political rights as men.” In this way, members of moral systems may come to question, criticize, and eventually change their moral beliefs (in these cases, one belief—human equality—leads to the changing of other related beliefs—the morality of slavery and women’s political rights respectively). This does not mean that their original beliefs were objectively false, merely that they were inconsistent with other beliefs. So slavery is not necessarily objectively wrong; it is, however, inconsistent with certain beliefs about equality. That is, particular moral beliefs can be wrong—in relation to a set of moral views. Harman calls this process of questioning, criticizing, negotiating, and changing moral ideas *moral bargaining*. I will return to the idea of moral bargaining in a minute.

People may balk at the idea of slavery not being objectively wrong. If slavery isn’t objectively wrong, couldn’t a society based on slavery just as easily decide to change their beliefs about equality of all people rather than their practice of slavery? Or what about a slave society in which there is no inconsistency because

there is no fundamental belief in human equality? Can we really say that slavery is justifiable in relation to the moral framework of such a society? Theoretically, the answer is yes. In such a society, if there is really just one monolithic moral framework in which slavery was not wrong (or even morally right), and if everyone, including the slaves, lived according to and believed that moral framework, then slavery is not wrong. However, I think it would be highly unlikely that there would be such a monolithic voice of society. In a slave society, I would imagine that there would be those in that society—namely, the slaves—who would not think that slavery was morally right. Thus, while slavery might be morally justifiable according to one moral framework in a society, it is not justifiable according to another in the same society.

This place of conflicting moral frameworks is where moral bargaining comes in. In moral bargaining, different people with different beliefs, desires, or interests form different “intentions,” and “after implicit bargaining, some sort of compromise is reached.”³ In a slave society, if the slaves were unhappy as slaves and believed that their enslavement was wrong, they might eventually challenge the dominant moral framework by doing such things as staging rebellions or talking to their enslavers and finding allies among them. Through a process of negotiation, rebellion, and/or debate, the dominant moral framework might change, or at least, in the interest of peace, the two conflicting moral frameworks might reach a compromise (such as ending slavery but maintaining strict laws

³ *ibid*, 13.

limiting the legal rights of the former slaves). Eventually, the resulting compromise can *also* be challenged and another compromise can be reached; theoretically, the process continues until moral agreement is attained.

Harman illustrates another important aspect of moral bargaining, which accounts for distribution of power and resources in moral negotiations, with the famous trolley problem. The basic trolley problem is this: a trolley is heading down a track to which six people are tied. I can flip a switch and the trolley will fork off on a different track, saving those six people. However, there is one person napping unaware on the other track, which is never used, and he will die if I flip the switch. Why is it that it seems worse to force the trolley onto another track to kill one innocent and unaware person and save the doomed six than it does to let the trolley stay on its track and kill six people? Another, perhaps more useful version of the trolley problem takes place in a hospital: a healthy person in the waiting room just happens to be a perfect donor match for six patients who will die without organ transplants. Each patient needs a different organ, so cutting up the healthy person (which would obviously kill her) would save *six* people. Utilitarianism (a form of moral universalism) holds that it is probably wrong *not* to force the sacrifice the healthy person, since that is what would adhere best to the “greatest good for the greatest number” principle. This doesn’t seem satisfactory, however. It completely disregards the intuitive feeling many of us have that it is wrong to harm someone actively. Letting the six organ-awaiting patients die would be tragic, but actively killing the healthy person to harvest her

organs would constitute murder. Utilitarianism, then, doesn't seem like a moral system to which many people would be happy about adhering. In fact, it is even written into our legal code in the U.S. that it is illegal to actively harm someone, but perfectly legal to stand by and let people be harmed (as long as the stander-by is innocent and not somehow complicit in the active harm-doing).

Harman's "hypothesis that morality derives from an agreement among people of varying powers and resources" (moral bargaining), however, explains this dilemma quite tidily:

The rich, the poor, the strong, and the weak would all benefit if all were to try to avoid harming one another. So everyone could agree to that arrangement. But the rich and the strong would not benefit from an arrangement whereby everyone would try to do as much as possible to help those in need. The poor and weak would get all of the benefit of this latter arrangement. . . . [T]he rich and the strong . . . would be reluctant to agree to a strong principle of mutual aid. A compromise would be likely and a weaker principle would probably be accepted.⁴

This explanation of moral bargaining explains how moral principles are formed, and also how disagreements surrounding them are reconciled. Moral bargaining, a process that constantly takes place in our society, must also be tied up in the same systems of power and privilege that members of society are tied up in. In negotiating slavery, for example, the slaveholders have much more power and many more resources than the slaves, and will most likely act in their own self-interest; this means that the compromise reached will likely benefit them more than it would benefit the slaves.

⁴ *ibid*, 12-13.

Another of Harman's arguments in favor of relativism is that it provides the best explanation for the existence of moral disagreement. Many cultural anthropologists (such as Ruth Benedict⁵) use descriptive relativism as an argument for metaethical moral relativism; in other words, they use the fact that there *are* a plurality of moral systems as evidence that no single system is any better than any other. Harman's approach is different, however. Rather than taking empirical evidence as proof,⁶ Harman argues that the existence of moral disagreement is most plausibly explained by his version of moral relativism. Moral relativism explains moral disagreement by recognizing that different people and groups of people have different moral frames of reference, and that within a particular moral framework, most views can be explained as an agreement based on moral bargaining.

Moral universalism, according to Harman, doesn't explain moral disagreement as satisfactorily. In the case of utilitarianism, its explanations can contradict our beliefs about what is right and wrong. In the case of the Golden Rule, moral disagreement really isn't explained at all. Suppose I don't like it if someone plays a practical prank on me, but my brother finds it funny. If both of us employ the Golden Rule when making decisions about morality, we might come to very different conclusions about what is right and wrong; I would probably find it morally objectionable to play practical pranks on people, while my brother would

⁵ Benedict (1934)

⁶ I have shown above (page 12) that empirical evidence of moral agreement wouldn't imply moral universalism, so to claim that empirical evidence of moral disagreement implies relativism would be inconsistent.

not find it at all morally objectionable. Golden Rule universalism would have to take a case-by-case look at moral disagreement, and would rationalize it based on context. Moral disagreement might be the result of, among other things, closed-mindedness (“I don’t understand it, so I don’t agree with it”), obstinacy (“I don’t care, I still think I’m right”), or traditionalism (“I’ve always done it this way, so no one can tell me it’s wrong”). All of these things would show different sorts of deficiencies in individuals, rather than disputing the established moral principle.

Harman shows that relativism is more plausible than universalism because its explanation of moral disagreement is a one-size-fits-all explanation. It is simple and straightforward, whereas moral universalism has to explain moral disagreement in a potentially infinite number of ways. A good analogy for why this kind of simplicity is preferred is the heliocentric model of the solar system versus the geocentric model. Though Copernicus hadn’t been into outer space when he invented the heliocentric (sun-centered) model of the solar system, this model was clearly more plausible than Ptolemy’s geocentric model (which had the Earth at the center of the solar system). The heliocentric model was a simple model of the sun at the center and the Earth and other planets orbiting the sun elliptically. The geocentric model, although it did work computationally, was far more intricate than the heliocentric model, and had to account for many bizarre circumstances of observation with individual explanations for each one. The geocentric model had planets and stars following all kinds of outlandish orbits around the earth. Although Copernicus wasn’t *proven* right until the space travel

became possible and the solar system could be observed from beyond Earth's surface, all things considered, his model was far simpler, far more sensible, and thus far more plausible than Ptolemy's model. Similarly, Harman's model, though scientifically unproven, seems to provide a much more satisfactory account of moral disagreement than a universalist model.⁷

Overview

The primary concepts of Harman's version of moral relativism, then, are as follows: people form different moral beliefs based on their individual circumstances, practices, and desires. None of these beliefs can be said to be objectively or universally true; rather, they can be said to be true or false in relation to a specific moral framework. Moral systems are defined and changed by a process of moral bargaining in which individuals with different interests implicitly reach a compromise about their visions of morality. Moral relativism is the most plausible explanation for the existence of a plethora of moral systems that are at odds with one another.

2. Criticisms of Moral Relativism

There are, as one might imagine, criticisms aplenty of Harman's relativism. Though I certainly won't be able to do them justice here, I do think it is important to address some of them before moving on. In this section, I will briefly discuss

⁷ This example comes from Kuhn (1970), p. 68-76.

some preliminary criticisms of moral relativism, including some of Judith Thomson's strategies for countering relativism, from her book with Harman, *Moral Relativism and Moral Objectivity*,⁸ as well as the objection that moral relativism doesn't allow us to have moral convictions.

Thomson

Judith Thomson criticizes moral relativism from multiple angles, some of which I believe are especially worth addressing. One strategy is to argue that, in Harman's version of moral relativism, moral statements end up being no more than statements of moral attitude. Harman argues that a moral statement doesn't have a truth value until it is fulfilled by an elliptical clause, "...in relation to such-and-such a moral framework." For example, the statement "it is wrong to kill dogs" is neither true nor false; however, the statement "it is wrong to kill dogs according to my moral beliefs" can be true. He further argues that although we generally don't say "according to my moral beliefs" when we make moral statements, this ellipsis can, and in fact must, be inferred in order for the statement to have a truth value. In other words, moral statements aren't true by themselves; they are only true in relation to a moral framework.

Thomson, however, disputes the idea of the incompleteness of moral statements. She argues that we rarely relate our moral statements to any other moral framework than our own, and when we do, we *always* do so explicitly.

⁸ Harman and Thomson (1996)

Thus, when a moral framework *isn't* referenced, it should be inferred that the speaker is referring to her own moral framework. If this can be reasonably inferred without being stated, then the moral statement is hardly incomplete. Suppose, for example, I say to my brother, "you shouldn't call girls sluts and whores." I don't say and it would probably sound weird to say "according to my moral framework, you shouldn't call girls 'sluts' and 'whores'." According to Thomson, the "according to my moral framework" bit is already implicit in the moral statement. It is automatically inferred by the audience that it refers to the speaker's own moral framework. If I said, "according to Dad, you shouldn't call girls 'sluts' and 'whores'," then I would be *explicitly* referencing another person's moral framework, but it would never be the case that I would implicitly reference someone else's moral framework. If moral statements that don't refer explicitly to a moral framework can always be inferred to be referring to the speaker's moral framework, then such statements are hardly incomplete. If this is true, then Harman is wrong about moral statements being incomplete. It does not dispute Harman's claim that moral statements as such do not have a truth value; however, if such statements aren't in fact incomplete, then all they do is state moral attitudes. This does not render Harman's entire argument for moral relativism false, but it does render it less plausible.

In summary, Thomson claims that on Harman's view of moral relativism, moral statements are nothing more than statements of moral attitude; this is because he claims that moral statements, in order to have a truth value, must be

completed by an elliptical "... in relation to such-and-such a moral framework." However, says Thomson, no one ever seriously makes a moral claim *outside of* her own moral framework, and if she does, she states so explicitly (not to do so would be absurd). Thus, you can assume that any moral statement without a specific reference to a moral framework is in fact relative to one's own moral framework. If this is true, then such moral statements aren't in fact incomplete; they nevertheless don't have a truth value without the elliptical clause (according to Harman); thus, they are mere statements of moral attitude, not moral statements that can be completed to have a truth value.

This is a somewhat difficult argument to counter. The reasoning is valid; I will argue, however, that this doesn't actually pose a huge problem for Harman. Harman would not disagree that it would be odd for someone to go around saying "you ought to be a vegetarian, relative to my moral framework"; saying that would be strange, because *we assume* that we speak from within our own moral frameworks. The reason this isn't a problem, Harman might say, is that for many of us, it is *part of our moral system* that we believe our morals are (or ought to be) universally accepted; thus, we wouldn't want to acknowledge that a moral statement such as "you ought to be a vegetarian" might only be true in reference to our own moral beliefs. To go back to Harman's spatio-temporal framework analogy, we never *state* that something is moving in relation to something else, because in our spatio-temporal framework, that is taken for granted. There is no need to say every time that "the earth orbits the sun, in relation to my spatio-

temporal framework” because we all assume that when you say “the earth orbits the sun,” you are speaking from an established framework (which is, I might add, one that we share). This doesn’t change the fact that the statement “the earth orbits the sun” *is not objectively and universally true* unless completed with information about the spatio-temporal framework.

Additionally, many people don’t know anything about, or just don’t understand, Einsteinian theories of motion, and so they don’t *know* that they’re making claims that are in fact only true in relation to a particular spatio-temporal framework. Before Einstein came along and invented his theories, clearly no one believed them either—they weren’t even an option. Nevertheless, it is still the case that they were, all along, talking about motion in relation to a specific spatio-temporal framework—their own. Similarly, many (if not most) people may not understand that there are a multitude, perhaps even an infinite number, of moral frameworks, and thus they *believe* that they are speaking universally, although their statements don’t have an absolute truth value unless they are taken relative to a particular moral framework. Almost without exception, people do make moral statements that relate to their own moral frameworks; this doesn’t end up refuting Harman’s claim that moral statements have no truth values until they’re “completed,” and it doesn’t change the fact that such statements are incomplete. Thus, moral statements such as “you ought to be a vegetarian” or “everyone ought to have the right to freedom of speech” *are not* mere statements of moral attitude,

and they *do* still require a relation to a moral framework in order to have a truth value.

Another of Thomson's strategies begins with this sentence: "It is a good heuristic in philosophy to be suspicious of views that would shock your grocer."⁹ She follows this statement with a scenario in which her grocer would be shocked and disturbed by her unwillingness to pay her grocery bill, which she won't pay because "I have no wants that would be met by my paying your bill, and I have wants that would be met by my not paying the bill." She understands Harman to be arguing that if she has no wants met by paying the bill, and has wants that *are* met by not paying, that morality doesn't require her to pay the bill. Thomson thinks this "idea looks crazy" and further states that

...surely it can't at all plausibly be thought that I am relieved of a moral requirement whenever it so happens that I have no wants that would be met by abiding by it, and have wants that would be met by not abiding by it, and my mental state is not due to an epistemic failure.¹⁰

Moral relativism in this light seems flighty, unreliable, highly individual, and counterintuitive. This seems so implausible, says Thomson, and her earlier arguments have already shown serious flaws in Harman's relativism, so all things considered, universalism is the better system.

Her argument, however, is based on a misunderstanding of the "mental state." Thomson refers explicitly to *wants*. She says, "JT has no wants that would be met by paying the bill and has wants that would be met by not paying the bill, and

⁹ *ibid*, 211.

¹⁰ *ibid*, 212.

JT's being in that mental state is not due to an epistemic failure." Put that way, this situation seems pretty reasonable. Clearly, by paying the grocery bill, she would be a little bit worse off financially; by not paying the bill, she would have both the groceries *and* would still have the money she otherwise would have spent. Obviously, this would be beneficial to her. Considering goods, JT definitely would be better off not paying the bill. This seems to be the understanding Thomson has of the "mental state." This is a very narrow conception. Harman would acknowledge many other aspects of the mental state that would enter into a person's decision to either pay or not pay a bill. For example, JT might consider beliefs she has in human integrity or ideals she holds that it is wrong to swindle people. Such beliefs and ideals would almost certainly prevent her from cheating the grocer out of sheer whimsy. She would also probably consider her conscience, and how she would feel about herself if she didn't pay—she might recognize that she *would* feel like a rotten thief. Indeed, it seems, based on Thomson's incredulity about the "crazy idea" that morality might not require JT to pay the grocery bill, that Thomson believes it *would* in fact be wrong not to pay it. If that is the case, then not paying the bill *would* be wrong and irrational, because it would be inconsistent with her moral beliefs that one ought to pay one's bills. JT might also, from a position of self-interest, consider what the repercussions would be if she didn't pay the bill—a hefty fine? Jail? Loss of respect from the grocer, from the community? In other words, a mental

state includes a whole lot more than just wants—it also includes motivations, beliefs, ideals, and many other subjective factors.

If JT actually *did* have sufficient reason not to pay her grocery bill, and if her entire mental state *had* been taken into consideration and it would not be irrational for her not to pay the bill, then Harman would say that no, morality does not, in fact, require her to pay the bill. Since this does not appear to be the scenario which Thomson is considering, however, her argument does not go very far. Harman would probably agree with her that JT would be wrong not to pay her grocery bill—only, he would add that it would be wrong in relation to her moral framework, not *objectively and universally* wrong.

Moral convictions

A third criticism of moral relativism comes from a completely different angle. If I believe that moral relativism is true, how can I have any moral convictions? How can I hold any moral beliefs? For example, suppose I grew up in a feminist household, and now, as a young adult, I hold the conviction that women, and not the law, ought to have control over their bodies; thus, I am pro-choice. Suppose further that I meet someone who grew up in a conservative Christian household and has learned that abortions are wrong because they are akin to murder; thus, she is anti-abortion. Moral relativism tells us that neither of us is more “right” than the other, unless one person’s view is more consistent with her entire value set than the other’s; supposing our views both *are* consistent with our individual

value sets, it seems we have come to a stalemate. If I believe Harman's version of moral relativism, I have to acknowledge that she is also "right," and if that is the case, then I might feel swindled: my beliefs and convictions are seemingly rendered meaningless. Since I can't know that I'm right, and I have to acknowledge that the opposite view isn't objectively wrong, then how can I have any positive moral beliefs or convictions? Moral relativism paralyzes me.

I have several things to say about this. First of all, I maintain that it is because of a shared moral principle within most of Western thought that our system of morality is universally true (a prime example of this is the Universal Declaration of Human Rights from 1948) that we even think our moral principles ought to hold universally at all. The West has a history of imperialism and colonialism in the rest of the world, and along with physical aspects of imperialism come cultural and intellectual aspects. The fact that so many of us feel the need to have our moral beliefs be universally true is in fact a moral belief itself—the moral belief that *everyone ought to have these beliefs*. It is entirely possible that there exist moral frameworks in which the people within the frameworks are completely comfortable with the fact that contradictory beliefs to theirs also exist and are equally justifiable. I am not saying that we ought to stop believing that our beliefs ought to apply universally. I only mean that the *belief* that our beliefs ought to apply universally does not make it *objectively and universally true* that our beliefs apply universally.

Moral relativism is a metaethical explanation of morality. It makes no claims about what we ought or ought not to believe; we can all continue as before with our beliefs and convictions. Harman's moral relativism might give us a reason to be more understanding of conflicting moral beliefs, but otherwise it doesn't have to change the way we think at all.¹¹ It provides a good explanation of *how* we come to have our moral beliefs, how we make moral decisions, and how we can have moral disagreement. Recognizing and acknowledging how the metaethical system of morality works does not mean we should stop playing our parts in it. I can still believe just as strongly that I am right, and I can still try to convince you that I'm right, because by trying to convince you I am just adding to your subjective motivational set, and who knows? Maybe you'll change your mind.

3. Conclusion

In this chapter, I have introduced the concept of moral universalism and the central arguments of Gilbert Harman's version of moral relativism. I have tried to show that, all things considered, relativism is more plausible than universalism. My arguments are by no means exhaustive, and it is not my intent to refute universalism definitively; rather, I have tried to show that and why I think relativism is right. There is one major argument for relativism that I have *not*

¹¹ It is worth noting that Wong (1984) disagrees; he applies a Kantian ethic of humanity as an end in itself to moral relativism such that interfering with another's ends purely for the sake of one's own ends is a "fail[ure] to treat them with the respect due rational beings" (181). Investigating the validity of this argument would be another whole project which I don't have the space for here; however, I suspect he errs on the side of making universal claims about all of humanity which Harman would probably not support.

fully addressed in this chapter; that is the argument about *reasons*. Chapter Two is devoted to a discussion of reasons as they relate to an understanding of morality, and the take that both Harman and Thomas Nagel (a proponent of universalism) have on them. I will show that reasons are at the crux of moral reasoning and deliberation. The rest of the thesis then investigates whether there are objective procedures for moral reasoning (rather than objective laws or principles) that can get around the difficulty that reasons pose for universalism. Specifically, I look at Rawls' constructivism as an example of a procedure for determining principles of justice. But first: how does Nagel fare in the face of relativism?

CHAPTER TWO:
NAGEL, WILLIAMS, HARMAN, AND REASONS

In the previous chapter, I introduced the general concepts of Harman's relativism, as well as some preliminary arguments against relativism and responses to them. In this chapter, I will take on Thomas Nagel's version of universalism. In doing so, I will focus on the concept of a reason, as this concept is integral to both Nagel's and Harman's arguments. I use Bernard Williams' view of reasons, as elaborated in his essay "Internal and External Reasons," to supplement and elaborate on Harman's view.¹² In this discussion about reasons and the role that Nagel and Harman take them to play in their arguments for universalism and relativism respectively, I will show that Harman's conception of a reason (and its role in his arguments for relativism) is more plausible than Nagel's.

I begin the chapter with a clarification of the word "reason" as it is used by Nagel and Harman (and in this thesis), since there are several different ways in which the word is used in philosophical discourse. Once that is out of the way, I give a positive account of Nagel's central arguments for universalism as outlined in his book *The Possibility of Altruism*,¹³ which include his arguments regarding subjective and objective reasons. I follow this with an exposition of Harman's

¹² Williams (1981)

¹³ Nagel (1970)

contrasting view of reasons, for which I incorporate Williams' essay. After this, I will return to Nagel's argument to show how Harman's and Williams' conception of reasons can be used to argue against Nagel's entire argument for universalism.

1. A Word on the Word 'Reason'

First of all, it is important to clarify that in this thesis, when I use the word "reason," I am *not* referring to an explanatory reason, but rather to a justificatory reason. An explanatory reason might take the form "he did x for reason y" in which y is some sort of explanation of the action x. For example, "Bob fell asleep because he was tired" would be an example of a statement with an explanatory reason, in which the reason for falling asleep was that Bob was tired. Another example of an explanatory reason would be psychological, e.g. saying that the reason that Bob is afraid of the dark is that he experienced some childhood trauma in the dark. Explanatory reasons are like underlying causes.

When I (and Williams, Harman, and Nagel) use the word "reason," however, I mean something different. I mean that a reason is justificatory. A justificatory reason might take the form "she has reason y to do x" in which y is some sort of justification for performing some action x. An example would be "Because she is opposed to the war in Iraq, Joan will vote for Barack Obama." In this case, although the reason "being opposed to the war" does explain Joan's voting for Obama, it also justifies her vote. Joan *has a reason* to vote for Obama, and the reason is that she is opposed to the war. Justificatory reasons indicate that one has

undergone some sort of deliberative reasoning in order to come to a certain conclusion or to perform a certain action.

When I speak of *objective reasons*, I mean justificatory reasons that apply objectively and universally. An objective reason, if such a thing exists, is a reason that everyone has to act or think a certain way. A Nagel-inspired example of an objective reason might be “I wouldn’t want someone to do that to me”; if this were an objective reason, it would be a reason that everyone has to justify not doing something (e.g. stealing an umbrella). Even if someone is unaware of that reason or doesn’t think she has that reason, if it is an *objective* reason, it is regardless a reason for her to act or think a certain way. It is a reason that is not internally determined, but rather applies to us independently of our internal deliberative processes.

2. Nagel, Introduced

Nagel takes a Kantian approach to moral universalism, though he focuses more on arguing for a specific moral law than on defending universalism in general. He asks the question, “Where will we find a reason that everyone has not to hurt other people, even those we don’t know?”¹⁴ His answer comes in the form of another question: “how would you like it if someone did that to you?” According to Nagel, reason requires that people obey the “Golden Rule” for the sake of consistency.

¹⁴ Nagel (1987), 64.

His argument is illustrated by the following example: Suppose someone trips me by sticking out his foot as I'm walking past unaware. I am indignant and resentful. My resentment leads me to believe that he ought not to have tripped me, that he actually had a reason *not* to trip me, and thus that he has the same reason not to trip *anybody else*. If he has this reason not to hurt me, then consistency insists that this reason must apply for anyone; there is nothing special about *me* that makes tripping me wrong. This means that I also have this reason not to trip anyone. So when I am contemplating tripping an unaware passer-by, I have to recognize that I have a reason not to—and that reason is, it would be inconsistent of me *not* to want someone to trip me, but for *me* to trip someone else.

This argument relies on a very central specific conception of the self: “a conception of oneself not merely as *I*, but as someone.”¹⁵ I am a person among other persons; the concept of personhood does not apply uniquely to me. This informs the step of the above argument that there is nothing special about me that makes it wrong to trip me: it is not an insult to *me*, but rather an insult to my personhood that makes tripping me wrong, and in order to take the step that it must be equally wrong to trip anybody else, I must see myself, in terms of personhood, no differently from anyone else. Although I can allow that I have individual thoughts and experiences, this does not in any way separate me from other persons (who *also* have individual thoughts and experiences). When I feel

¹⁵ Nagel (1970), 100.

resentment towards someone for tripping me, and once I have acknowledged that anyone else would feel a similar resentment had *they* been tripped, then I am forced (according to this particular self-conception) to acknowledge that there is nothing special about *me* that allows me to separate myself from all other people. So the next time I am waiting to board a plane and out of sheer boredom consider tripping a hapless traveler, the fact that I would be resentful if *he* tripped *me* is enough to give me a reason not to trip him (he is, after all, no less a person than I am).

You will notice that this argument does not rely on any *feelings* of altruism, such as generosity, compassion, good will, empathy, etc; nor does it rely on any feelings of self-interest or self-preservation. It is instead what Nagel calls “pure altruism,” which “has genuinely the status of a rational requirement on human conduct.”¹⁶ This rational altruism may be (and in fact often or even usually is) accompanied by such feelings, but these feelings are not necessary components of Nagel’s altruism. Even if everything about my feelings and experiences directly go against altruism (if, for example, my sheer hatred for an individual leads me to feel absolutely nothing like compassion or generosity towards her), I still have a reason to act in a “purely” altruistic way, according to Nagel, because reason and rationality require it of me, and I would be irrationally inconsistent if I didn’t act altruistically. Thus when I decide not to trip my fellow traveler, I am not compelled to have any particular feelings (e.g. compassion for fellow human

¹⁶ *ibid*, 80.

beings), and in fact, it can still be true that I have a reason not to trip him, even if I have no *desires* not to trip him.

All of this ties in very closely with Nagel's conception of a reason. Nagel offers a logical justification for objective reasons which I find to be incredibly unhelpful and not all that comprehensible. I provide a brief exposition of my understanding of it in Appendix A, but since it doesn't add much to his argument that subjective reasons are incomplete without objective reasons to back them up, I won't include it here. Essentially, the idea behind the argument is that subjective reasons aren't complete without objective reasons that stand behind them. Subjective reasons, if they could stand by themselves, would endorse the idea that there is something special about me that is different from others.

More useful than the logical justification is his discussion of solipsism.¹⁷ In his discussion of solipsism, Nagel asserts that there is no meaningful difference between the claim that "I have brown hair" and the claim that "she has brown hair"; the content of both claims is the same, and the difference is one of perspective only. And although the circumstances that bring about the two claims may be different, this "should not obscure the fact that different persons . . . of the same statement can be employed from appropriately different standpoints to *say something common* about a given situation" [emphasis mine].¹⁸ In the same way, Nagel takes *moral* claims to be impersonally motivated; although the claim "I have a reason not to trip the traveler" may come about under different subjective

¹⁷ *ibid*, 99-115. Solipsism is the idea that I can only be certain that I exist, that MY self exists.

¹⁸ *ibid*, 102.

circumstances from the claim “she has a reason not to trip the traveler,” the claims both say the same thing about the same situation. The only difference is one of perspective. This assertion that judgments about reasons from differing perspectives have the same content supports Nagel’s earlier arguments that what one says about someone else’s reasons (“she has a reason not to trip me”) must apply equally to oneself. If you disagree that the only difference is one of perspective, you are, according to Nagel, a “practical solipsist” in that you think that YOU are somehow different from HER, or that there is a difference between “I” and “SHE.”¹⁹ Clearly, this contradicts his notion of personhood, which is that there *is* no moral difference between persons (see page 30).

Overview

Nagel’s most essential argument is that everyone must conceive of herself in every respect as merely a person among other people. His universal moral law that is based on this conception of the self is the Golden Rule: “how would I like it if someone did that to me?” If I decide that I would feel resentment towards someone for a particular action, then that gives that person a reason not to perform that action on me, and since I have no reason to suppose that there is anything special about *me* that gives this person a reason not to trip me, I conclude that she has a reason not to trip *anyone*. Again, since there is nothing special about me, I realize that I also have a reason not to trip anyone. Thus, since

¹⁹ *ibid*, 113-14.

I am a person among others, acknowledgment that I would feel resentment on account of a particular action gives me a reason not to perform that same action. In this way, all reasons are objective, because they are not merely reasons that *I* have, but rather reasons that *all people* have. At a meta-level (not concerning the particular things we may or may not have reasons to do, but rather the reasoning that guides us in such decisions), Nagel's argument is that everyone has a reason (an objective reason) to obey the Golden Rule. The Golden Rule, then, is an objective constraint on moral reasoning; it requires that we engage in moral reasoning in a particular way.

3. A Subjective Account of Reasons: Harman

Nagel's argument rests on an entirely different conception of reasons from Harman's. In this section, I will first give an overview of Harman's conception of reasons, and some particular passages in which he addresses Nagel. Then I will refer to Bernard Williams' paper "Internal and External Reasons" in order to give a more complete account of a reason according to Harman's conception. I will wrap up the section with an explanation of how Williams' account is useful to and supports Harman.

First, however, I will briefly recount Harman's response to Nagel's argument against practical solipsism. Essentially, the argument is that there is no substantial difference between "I" statements, "she" statements, and "you" statements (or statements from any other perspective); thus, saying "she has reason *y* to do *x*"

means essentially the same thing as “I have reason y to do x.” Harman responds to this directly by stating that Nagel’s so-called “objectivity principle” (which states that statements which differ only in perspective still share the same content) “is not perfectly general and some argument is needed to show it should be applied to what Nagel calls ‘motivational content.’”²⁰ In other words, just because the statements “I have brown hair” and “she has brown hair” share the same content, it is not necessarily the case that the same obtains for judgments with “motivational content” (i.e., judgments about whether someone has reason to do something); if it does obtain, it demands careful justification, which Nagel does not provide. This is not so much a counter-argument as it is a nudge to Nagel that the burden of proof is on him, and he doesn’t follow through.

With that out of the way, I will turn to giving a positive account of Harman’s conception of a reason.²¹ Harman distinguishes between a “sufficient” reason and a “compelling” reason. Someone has a sufficient reason to do something if it would be “(at least) reasonable” for her to do it. She has a compelling reason if and only if it would be reasonable for her to do it and unreasonable for her not to do it.²² This distinction is necessary because in the rest of his argument, he specifically refers to *compelling* reasons.

Harman next discusses what it means to say that “there are reasons for someone to do something”: essentially, it means the same thing as “there are

²⁰ Harman and Thomson (1996), 52-3.

²¹ *ibid*, 45-8.

²² *ibid*, 45-6.

considerations the awareness of which would provide [someone] with compelling reasons to [do that thing].”²³ So, for example, I have a compelling reason to use an umbrella if there are certain considerations (I will get wet if I don’t) the awareness of which *give* me a compelling reason to use an umbrella (since I don’t want to get wet).

The crux of Harman’s view is summed up in this sentence, in which *S* is some person and *D* is some action: “if *S* does not *D*, *S* is aware of all relevant considerations, and *S* is not in any way unreasonable, then it is not the case that there are compelling reasons for *S* to *D*.”²⁴ What this means is that, using the above example, if I *don’t* think I have a compelling reason to use an umbrella despite awareness of any and all relevant considerations that might convince me otherwise, if I *don’t* use an umbrella, and if my decision not to use an umbrella isn’t unreasonable, then I do not, in fact, have a compelling reason to use an umbrella. It is a simple matter to show that my decision not to use an umbrella is not unreasonable. Consider:

- (a) If I don’t use an umbrella, I will get wet.
- (b) I don’t care whether I get wet.
- (c) I don’t use an umbrella.

I can use a simple modus ponens to show that if (c) is true, then I will get wet.

This clearly does not contradict (b), so it is not unreasonable not to use an umbrella.

²³ *ibid*, 46.

²⁴ *ibid*, 46.

What about moral reasons? Harman maintains that the same thing holds for moral reasons that holds for all other reasons. He claims that arguments “purporting to show that there are reasons for such a person to observe one or another moral requirement, reasons that the person is allegedly failing to appreciate . . . do not work.”²⁵ Instead, the reasons a person has for particular moral decisions or actions depend entirely on “that person’s principles and values.” Harman’s reasons, then, are entirely *subjective*, in contrast to Nagel’s *objective* reasons.

4. Williams Lends a Hand

At this point, I think it is useful to consider Bernard Williams in support of Harman’s views on reasons. Williams’ essay “Internal and External Reasons” performs a much more detailed deconstruction of objective reasons (what he calls “external reasons”) than Harman performs anywhere.²⁶ He develops Harman’s idea that an individual’s reasons rely on subjective reasoning alone. I will do my best to recreate his arguments here.

In order to talk about external reasons, it is first necessary to discuss briefly “internal” reasons, for which Williams initially provides the following definition:

²⁵ *ibid*, 47.

²⁶ There is a bit of a difference between what Williams refers to as an external reason and what I have until now been referring to as an objective reason. By “external reason,” Williams means any potential “reason” that supposedly comes from some source outside the self. Thus, a reason that any individual has for me to do something might also be an external reason, even if it is not an objective reason. All objective reasons (in the sense that I mean them) are external reasons, so the distinction is not important for my purposes. All of Williams’ arguments concerning external reasons also apply to objective reasons.

“*A* has a reason to φ iff *A* has some desire the satisfaction of which [*A* believes] will be served by his φ -ing.”²⁷ An internal reason must have some relationship to *A*’s “subjective motivational set,” (*S*), which includes all of *A*’s background, personal experience, desires, beliefs, et cetera, and *only* these subjective things.

Williams amends this definition over the course of the paper to make it more “adequate.”²⁸ His amendments, in summary, add the following information to the definition of an internal reason:

- 1) An internal reason statement can be false if it is based on a false belief; thus, we may not know what our *true* internal reason statement is.²⁹ However, “it will, all the same, be true, that if he does φ in these circumstances, there was not only a reason why he φ -ed, but also that that displays him as, relative to his false belief, acting rationally.”
- 2) Internal reason statements are based on “deliberative reasoning,” which means they are logical conclusions based on premises in the subjective motivational set.³⁰

What about external reasons? Consider the following example: you and I are in an airplane, and we are going skydiving. You are getting ready to dive, and I am inspecting the parachute one last time. I discover that the parachute has a malfunction, and so if you dive, you will die (unless you are extremely lucky). I haven’t told you any of this yet. Don’t you still have a reason not to jump? It

²⁷ Williams (1981), 101.

²⁸ *ibid*, 102.

²⁹ *ibid*, 103.

³⁰ *ibid*, 104.

seems that the fact that you will in all likelihood die if you jump with this parachute *ought* to be a reason for you not to jump, even if the reason doesn't arise out of your subjective motivational set, i.e. is external. Is this really an external reason, though? I have broken down this example into three cases in order to better evaluate:

- 1) I don't tell you that the parachute is broken. You don't find out, and you jump as planned. You die. Before you jumped, your internal reason statement for jumping would have probably sounded something like: "I have everything I need to jump, and it works properly, so I'm going to jump!" In this case, one of your beliefs is clearly false; the parachute does *not* in fact work properly. Thus, according to Williams, your internal reason statement is false.
- 2) I tell you that the parachute is broken, and you decide to jump anyway. In this case, you have the information, it becomes a part of your subjective motivational set, and yet, right before you jump, you say "well, I study philosophy, and so I am convinced that my life is meaningless, so it doesn't really matter if I die; additionally, the fun I will have by jumping will be worth dying for. Thus, I don't care if the parachute is broken, I'm jumping anyway!" You have all the necessary information to make a well-informed decision, and this is the decision you make. *I* may have a reason for you not to jump (namely, I like you, and I'd rather you live), but this is hardly a reason

you have, since you're not me. In this scenario, your internal reason statement is not false.

- 3) I tell you that the parachute is broken, and you decide not to jump. In this situation, as soon as I tell you about the parachute, you say "I'd rather not die, and if I jump with a broken parachute, I might die. That's a good reason for me not to jump!" The knowledge of the broken parachute becomes a part of your subjective motivational set; you have the information, and its state of dysfunction *becomes an internal reason* for you not to jump.

In cases 2 and 3, what I may have originally thought of as an external reason turns out to be an internal reason. Case 3 is probably the clearest; I believe that the parachute's state of malfunction is a reason for you not to jump, and as soon as you find out, you accept it as a reason. Now that you know, it is a part of your subjective motivational set, and can no longer be considered external.

Case 2 is slightly more complicated. You know the parachute is broken, you know you will probably die, and yet you jump anyway. Here, the broken parachute is a reason for me not wanting you to jump, but it seems as though it isn't a reason for you not to jump. It's not as simple as that, however. It is probable that the broken parachute does give you a reason not to jump. It might just not give you *compelling* reason. All things considered, you would still rather jump, and since you have arrived at this decision based on deliberative reasoning that is not based on false beliefs, it is a true internal reason statement. One might

argue that the broken parachute might still be an external reason not to jump, but Williams would reject such an argument: once the false belief is corrected, you have all the necessary information in your subjective motivational set to make a rational decision. What seems like an external reason is just *my internal reason* for you not to jump based on my own subjective motivational set.

Suppose some combination of 2 and 3 is the case: I inform you of the broken parachute, and you decide you will jump anyway for the reasons given in case 2. However, I plead with you and share with you *my* reason(s) for thinking you shouldn't jump. If I thus convince you not to jump, then a keen universalist might jump up and say "ha! See? An external reason!" Williams argues, however, that this is no longer an external reason, because my reasons have also become a part of your motivational set—now you do subjectively care about (some of) the same things I do. Clearly I was convincing; this is analogous to case 3. If, however, my pleading and reasoning amounts to nothing, and you decide to jump anyway, then the case is analogous to case 2.

Case 1 is probably the most complicated of these three cases. In this case, you don't find out that one of your beliefs is false, so your entire internal reason statement is also false. Here is where another aspect of Williams' argument comes in: even though your internal reason statement is false, it cannot be said that there was an external reason for you not to jump, because a reason also has to *explain an action*, and "no external reason statement could *by itself* offer an

explanation of anyone's action."³¹ In this case, since you have no knowledge of the broken parachute, it cannot be an explanation of your action in any way; if it *were* an explanation, then you would have to know about it, and then either case 2 or case 3 would have to be considered instead.

According to Williams, then, external reasons don't actually exist. There isn't any such thing. He says that "the sort of considerations offered here strongly suggest to me that external reason statements, when definitely isolated as such, are false, or incoherent, or really something else misleadingly expressed. ... If this is so, the only real claims about reasons for action will be internal claims."³² If external reasons are incoherent, then clearly objective reasons are also incoherent, since objective reasons are a universal sort of external reasons.

Suppose I make my moral decisions based on the question "how would I feel if someone did that to me?" and you don't make your moral decisions based on that question. If, after all things have been considered, you still don't think you have a reason to accept my moral system, and none of your beliefs are empirically false, then you do not have a reason, external, objective, or otherwise.

It is very likely that Harman would agree with this discussion of reasons, specifically since he holds that reasons are internal and based on subjective beliefs and motivations. This concept of reasons, however, is quite different from Nagel's. Nagel's arguments for moral universalism in the form of the Golden Rule rely on *external* reasons, and Williams has shown that Nagel is

³¹ *ibid*, 106.

³² *ibid*, 111.

insupportable in believing that it is possible for there to be a reason that applies universally and objectively to everyone, a reason that everyone must consider in order to be reasonable.

5. A Second Look at Nagel

Let's take another look at Nagel's argument for the Golden Rule in order to evaluate it more carefully according to Harman and Williams. I will break down Nagel's argument into steps:

- (a) I am a person among other people, not just an *I* but a *someone*.
- (b) I am considering stealing an umbrella from the entry of the library, because it is raining and I do not have my own umbrella. I think, "How would I like it if someone stole *my* umbrella (if I had an umbrella)?"
- (c) I recognize that I would feel resentment towards whomever stole my umbrella, no matter who it was.
- (d) This resentment leads me to the conclusion that this person would have a reason not to steal my umbrella.
- (e) Since (a) is true, there is nothing special about *me* that everyone *else* has a reason not to steal *my* umbrella but I don't have a reason not to steal anyone else's, I conclude that I as well as everybody else *do* have a reason not to steal an umbrella from the foyer.

Harman and Williams might not even agree with (a), but suppose, just for the sake of argument, they do. (I will return to this issue in a moment.) What about the rest of the argument? Harman would have no problem with (b). He would say, terrific, you operate within a moral framework in which you (apparently) try to put yourself in other people's shoes by imagining that you are in a similar situation. This is not problematic for Harman, because this step is still about *me*, about how *I* am considering the moral situation from my point of view. (c) similarly offers no problems. By imagining myself in a similar situation, I am able to realize that I would feel resentment. Great. If the conclusion from this were that I shouldn't steal another person's umbrella, Harman would have no problem with the argument at all; at this point, the argument is still entirely subjective. The problem is, Nagel doesn't stop here.

Nagel wants to argue from (c) that (d). This is where Harman would jump ship. Harman would say, how does the fact that *I* feel resentful give *this other person* a reason to do or not to do something? Why should this other person necessarily care what I might feel? Suppose this other person happens to be an employee of Umbrella Stealers, Inc. It is built into his moral framework that he explicitly *doesn't* care about the feelings of other people outside of his employing company. There is no way I or anyone else can convince him that he ought to care about others' feelings. Even if *I* might think or claim that he has a reason not to steal my umbrella, according to Williams' arguments that external reasons are

not supportable, this definitely does *not* give him a reason not to steal my umbrella.

Harman would probably not have much of a problem with (e), if (a) were true and if (d) were not part of the argument. In other words, he would agree that if you think of yourself as just one person among others equally persons, and if you think it's important to consider their potential feelings as seriously as your own, then you probably would have to conclude that you have no reason to steal someone else's umbrella, and that you would be inconsistent (and thus wrong) if you did. However, this step is quite contingent on (a) being true. It is also contingent on the belief that I should not do to other people what I would not want done to me, because if I feel resentment about something, it means that others probably would too (which is also, Harman might point out, not necessarily the case). For this to be the case, it is important to clarify what resentment means. I understand Nagel to take resentment to be not only a feeling of anger or hostility, but also a feeling that somehow *includes* in it the judgment that "that person should not have done that." Thus, if I were resentful towards someone for stealing my umbrella, then I would think that that person should not have stolen it; so if someone *else* would be resentful that *I* stole *her* umbrella, then that means that *I* should not steal it. Even if it were the case that the definition of resentment necessarily does include this clause (which I see no justification for in Nagel's arguments), and even if it were the case that all people would feel resentful about the same exact things (which is highly unlikely), this by no means implies that it

is necessary that everyone accept the Golden Rule, that we should treat others the way we want to be treated, or that we must care about how others feel.

It almost seems pointless at this point to return to the question of whether (a) is necessarily the case, because even if it is, I have already shown that Nagel's argument is flawed. I may as well cover all the ground, though, and I will actually return to this question briefly later in Chapter Four. The question is whether it is the case that I must conceive of myself as a person among other people, and thus as no more special or privileged than any other person in terms of my humanity and personhood. It seems to me unlikely that Harman would agree with this; he would probably say, there is no *reason* why we all have to conceive of ourselves this way. One might very well be a part of a moral system in which certain people *do* have more humanity and privileges over others, and that some people in this system are less human than others. Even if he did agree, however, that all people are equally human and that we must conceive of ourselves and others that way, I doubt that he would agree that this has any implications for how we engage in moral reasoning and for the kind of moral beliefs we ought to have.

These are the arguments that Harman might use to respond to Nagel. Even *if* Nagel could somehow respond adequately to all of them, however, and even if we were to grant him the chance to show that reasons are objective, Nagel runs almost immediately into a huge problem: how can he account for moral disagreement?

Suppose, for example, someone comes along who *wouldn't* feel resentment at having his umbrella stolen. In this case, Nagel's purported "objective" reason falls flat. Clearly this is just one example of umbrella-thieving, but the same can be said of any other moral issue—murder, for example. Harman imagines Murder, Inc., a company in which employees kill company outsiders without hesitation or feelings of guilt.³³ An employee of such a company may very well *not* feel resentment about being murdered himself, as long as his murderer was a company outsider. In fact, he might think (contemplating the circumstances from Heaven, or Hell, or Purgatory, or wherever he may be) that his murder was skillfully executed, and he might even *admire* his murderer. He might still be unhappy at the prospect of being dead, but he harbors no resentment towards his murderer, and thus, he harbors no feelings of guilt or regret about murdering people himself.

What this shows is that it may be impossible, using Nagel's formula for discovering an objective reason, to actually find any normative reasons that are objective. And even if we somehow *did* manage to poll every single one of the residents of the planet, even if that poll revealed that there is a moral principle that *every single person* believes in, some act that *every single person* would resent, were it performed on them, there is no way to prove that the empirical fact that all people agree on that principle implies that all people always have and always will agree on it (I argued this same point in Chapter One; see page 4).

³³ Harman (1975), 5.

Nagel might just reply that clearly, some people are right and some people are wrong. But that just begs the question, how do we know who is right and who is wrong? And if we have to ask that, then we're no better off than we were before hearing Nagel's initial argument for a universal moral law.

6. Conclusion

There is no one sweeping argument here for Harman's relativism. However, there are counter-arguments for many of the criticisms and difficulties posed by Nagel, and I think moral relativism accounts for moral difference and disagreement much better than universalism does. Relativism has also shown that there aren't any objective reasons, and thus, that moral theories (such as Nagel's) that rely on them are insupportable. The next question is, are there any objective *constraints* on moral reasoning? In other words, is there anything that rationality requires in terms of moral deliberation? In the next chapters, I will investigate John Rawls' constructivist theory, and in doing so, I use my discussion of Nagel as a guide, for there are some striking similarities between the two of them (as well as striking differences). I pose the following question: does Rawls effectively argue that there are any aspects of morality or moral reasoning that are rationally required, or is Rawls, like Nagel, just another metaethical theorist whose arguments depend on faulty assumptions about human nature and universal moral law?

CHAPTER THREE:
RAWLS AND CONSTRUCTIVISM

In the previous chapter, I showed how Harman's arguments for moral relativism and Harman's and Williams' arguments against external or objective reasons provide solid refutations to Nagel's version of moral universalism. In the next two chapters, I investigate whether there are any objective moral *requirements* (rather than specific moral principles or rules). The question of whether relativism is true or not relies on this step; if specific normative universal laws turn out not to be universal, the question becomes: are there any requirements for moral *reasoning*, or, how we think about morality? First, I present Rawls' theory of justice (constructivism). Rawls, as I explain below, is a prime example of someone who provides a procedure for determining laws of morality, rather than providing specific laws of morality. Then, I evaluate the premises of his arguments, assumptions he makes, and intuitions he relies on for his theory according to how Harman might evaluate them. In doing so, I take a two-pronged approach: in this chapter, I try to show how constructivism is, despite substantial differences from Nagel, also in many ways similar to him; thus, many of my criticisms of Nagel will work for constructivism as well. In Chapter Four, I offer another argument against constructivism as an alternative to moral relativism; even if my analysis of constructivism in Chapter Three is

flawed, Chapter Four will show that nevertheless, constructivism, which aims to provide an objective procedure for determining principles of morality, fails to respond adequately to Harman's relativism.

1. Metaethical, not Normative

Before I embark on recounting Rawls' theory, I would like to say a word about how I read Rawls. Rawls' theory of justice as fairness (which I explain in the next section) can be understood and interpreted both as a normative theory and as a metaethical theory (although Rawls himself may have been reluctant to explicitly discuss the metaethical aspects of it—in "Justice as Fairness: Political not Metaphysical," Rawls argues for a political interpretation of his theory, saying that he "avoid[s] pre-judging these larger questions [of whether the theory of justice as fairness can be extended beyond a political conception based in a constitutional democracy] one way or the other"³⁴). As a normative theory, it explains the principles of justice that would be determined in an ideally democratic society. It could also attempt to determine what might be morally right or wrong, or offer solutions to moral quandaries which perpetually surface in society. As a normative theory, it would *not* claim objective truth for any moral principles identified. In "Justice as Fairness: Political not Metaphysical," Rawls explains it to be "a moral conception worked out for . . . political, social, and

³⁴ Rawls (1985), p. 225.

economic institutions”³⁵ and “a political conception of justice for a democratic society.”³⁶ A normative reading of justice as fairness is steeped in a history and intellectual tradition of democratic thought, which is specific to a certain time (post-Enlightenment) and place (Europe and the West).

Ronald Milo, however, in a later article, builds on Rawls to develop a conception of Rawls’ constructivism as a metaethical theory.³⁷ “Rawls’ disinclination to become embroiled in metaphysical/metaethical issues, and his preference for addressing more practically oriented normative questions, prevented him from developing this conception himself,” Milo states.³⁸ As the discussion in this thesis is metaethical, I rely on and expound upon this interpretation of justice as fairness, using Milo at times to fill in the metaethical approach to Rawls’ theory in places where Rawls himself hesitates to do so. There *is* considerable ambiguity in Rawls, and Milo offers a reasonable interpretation and expansion of him. From this point on, I will use Rawls’ term “constructivism” to refer to the metaethical theory based on his theory of justice as fairness.

What does it mean for constructivism to be a metaethical theory? It means that it is a theory about the *nature* rather than the *content* of morality. As a metaethical theory, it is no longer about universal laws, but rather about objective procedures for establishing systems of morality. It doesn’t determine what is right

³⁵ *ibid*, 224.

³⁶ *ibid*, 225.

³⁷ Milo (1995)

³⁸ *ibid*, 185.

or wrong; rather, it determines the nature of concepts such as right or wrong. This *might* be compatible with Harman; such a procedure might allow for a certain amount of relativity of laws. On the other hand, it might fall into the same trap Nagel falls into—it might presume the existence of objective reasons.

Constructivism, as a universal procedure, would put some objective requirements, or constraints, on moral reasoning, but the question for Harman is whether these objective requirements in fact rely on objective *reasons*.

Thus, I consider constructivism, like Harman's and Nagel's theories, to be a theory about moral reasons. Making a moral decision or performing a moral action relies (just like making any non-moral decision or performing any non-moral action) on having a reason to do so. For Harman, there aren't any objective reasons, so moral reasons are entirely subjective and thus entirely relative. For Nagel, there is one significant objective reason, and this reason (the "Golden Rule") means that morality is universal. For constructivism, then, the question is whether there are objective reasons or not, and thus whether there are any rational requirements for moral reasoning. That is the question that I will investigate in this chapter. But first, I must present Rawls' theory.

2. Justice as Fairness

Rawls summarizes the "guiding idea" of his theory in several sentences at the beginning of *A Theory of Justice*.³⁹ His goal is to

³⁹ Rawls (1971)

present a conception of justice which generalizes and carries to a higher level of abstraction the familiar theory of the social contract [in which] the principles of justice for the basic structure of society are the object of the original agreement. They are the principles that free and rational persons concerned to further their own interests would accept in an initial position of equality as defining the fundamental terms of their association. These principles are to regulate all further agreements; they specify the kinds of social cooperation that can be entered into and the forms of government that can be established.⁴⁰

There are several different ideas in this abstract, which I will pull apart. First, I will explain Rawls' notion of the original position as it relates to social contract theory. Next, I look at the conditions Rawls considers necessary in the original position, primarily the veil of ignorance. Finally, I will consider the concept of personhood which the original position relies on, which includes freedom, equality, rationality, and self-interest (the elements mentioned in the above extract), among other things.

A significant portion of *A Theory of Justice* is devoted to working out the principles of justice which would be decided upon in the original position. At that point, the theory is a normative one, and so it doesn't help my thesis at all to focus on them. My object is not to critique the principles which Rawls believes the hypothetical contractors would create, but rather to be a critic of the rules of the game (and the game itself) Rawls invents. Thus, when summarizing Rawls' argument here, I will not discuss any of the normative principles of justice.

The Original Position

⁴⁰ *ibid*, 11.

The original position is the “initial position of equality” cited in the above extract, which is analogous to the state of nature in traditional social contract theory. It is not meant to be historical (in fact it is distinctly ahistorical, as will become clear); nor is it meant to be achievable (again, as will become clearer as this chapter progresses, many of the elements of Rawls’ original position are not possible in reality). It is, as Rawls puts it, “the appropriate initial status quo,” a situation which (for the purposes of constructivism) is ideally fair.⁴¹ It is a purely hypothetical position.

The essential qualification of the original position is that of equality. In “Justice as Fairness,”⁴² Rawls conceives of justice as “essentially the elimination of arbitrary distinctions,”⁴³ which helps to explain his “initial position of equal liberty.”⁴⁴ Any initial inequality between individuals is, according to Rawls, arbitrary, at least for the purposes of establishing principles of justice. This does not mean that there is no descriptive inequality; nor is it a normative claim that there ought not to be any descriptive inequality. Rather, Rawls holds that equality of the hypothetical contractors in the original position is a necessary condition for establishing the principles of justice. Any inequality in the original position automatically favors the viewpoint of one party over another, and therefore might unfairly skew the resulting principles of justice. For example, if men and women

⁴¹ *ibid*, 12.

⁴² Rawls (1958)

⁴³ *ibid*, 165.

⁴⁴ *ibid*, 166.

are unequal in the original position, principles of justice might turn out biased in the favor of either men or women.

But what *is* the original position? It is the position from which principles of justice are decided upon. It is a hypothetical situation in which people come together and work out a hypothetical “contract” for a just society. Since it is hypothetical, it can be ideal; in *A Theory of Justice*, Rawls makes the rules for an ideal hypothetical “game,” or procedure for establishing principles of justice. One of the rules is the veil of ignorance.

*The “Veil of Ignorance”*⁴⁵

One crucial condition of the original position is to take away every person’s knowledge of her own situation and put everyone behind a so-called “veil of ignorance.” If the agreements between the people establishing the principles of justice are to be fair, all knowledge of an individual’s own situation must be abstracted. By way of explanation, I will give an example (within the context of the United States). Suppose three people are trying to decide on principles of justice for their society: 1) a wealthy, white, agnostic, middle-aged gay man; 2) a young, working class, straight Muslim man with immigrant parents; and 3) an elderly, middle class, straight, Christian white woman. If they were each motivated by self-interest, how would they agree on any principles? Assuming that all three have to agree on the principles, there might be many situations in

⁴⁵ Rawls (1971), 12.

which those with more power (either due to numerical majority or due to privileged status in society) somehow sway the other(s) to agree with them. The principles agreed upon by this group of three people might end up privileging, for example, heterosexuals, whites, men, non-immigrants, working, young, and able-bodied adults, the middle/upper-middle class, and the Christian mainstream.

According to Rawls, the privileging of these groups (or any groups) is not fair.

In order to ensure fairness, it is necessary that none of these three people know the particulars of their own situations, or the particulars of the others' situations. (The particulars are, among other things, race, age, gender, sexual orientation, religion, wealth, etc.) If, for example, none of the three people above knew whether they were male or female, but knew that they each had a 50% chance of being male and a 50% chance of being female, it would benefit each of them to agree on a principle that would equally benefit men and women. The same thing holds for all other contingent identities. If no one knew his religion, but knew there was a chance that he might be non-Christian, he would advocate a society in which all religions were welcome. If no one knew their own particular situations, the decisions they would reach would be impartial. And the decisions reached would be impartial *because* of the impartiality of the *process itself*—the veil of ignorance forces all the hypothetical contractors to be impartial when determining their principles of justice. Rawls evidently holds impartiality to be necessary for fairness in matters of justice.

In addition to not knowing their contingent social identities, such as race, class, age, and gender, the hypothetical contractors cannot know the particulars of their own dispositions (such as a disposition to taking risks, or a disposition to being extremely generous). They also can't know anything about any historical or political circumstances of their own societies, which might additionally influence their thinking and decision-making. If they knew the history of the struggle for equal rights for women, for example, they might end up being biased against women, just by having some internalized belief that women must somehow be inferior to men if they weren't always treated equally. If they knew about slavery in America, they might think that blacks were inherently subordinate; if they knew about the Middle East crisis, they might have suspicions about Muslims or Jews. Thus, it is important that those who are making decisions about justice are completely uninformed about anything which might give them a bias towards or against any people or groups of people. "They must choose principles the consequences of which they are prepared to live with" no matter who they turn out to be.⁴⁶

So far, I have only described what the contractors *cannot* know because of being behind the veil of ignorance: individual situations and historical and political circumstances. What, then, *can* the contractors know? They do know "the general facts of human society,"⁴⁷ which include the general structures and theories of sociology, psychology, and the economy, so that they have a

⁴⁶ *ibid*, 137.

⁴⁷ *ibid*, 137.

framework in which to base the principles. They are also aware of what Rawls calls the “primary goods,” that is, “things that every rational man is presumed to want.”⁴⁸ Primary goods are the means that are essential to each achieving his good, things such as “rights and liberties, powers and opportunities, and income and wealth.”⁴⁹ It is necessary that everyone have knowledge of the primary goods, so that each can know what exactly their principles will be regulating and distributing. Essentially, the contractors know just enough to be able to establish principles of justice, but not enough to allow any individual bias. The veil of ignorance renders all contractors unequivocally equal.

Personhood

The original position and the veil of ignorance both require a fairly specific concept of a person. These specifications and qualities are further rules of Rawls’ constructivist game. Paraphrasing from the extract at the beginning of this section, the principles of justice are those that free, equal, and rational persons would agree on (there are other qualities as well, which I will address later). I have already addressed equality in my initial discussion of the original position; in addition to equality, then, the people who are making the decisions about the principles of justice must also be free and rational. This does not, of course, mean that all people descriptively *are* free and rational. It is obvious that in society, some people hold more power and have more privileges than others, and some

⁴⁸ *ibid*, 62.

⁴⁹ *ibid*, 62.

people have more freedom than others (freedom that is a result of having certain powers and privileges, such as financial wealth, which enables people to be free of debt and have fiscal autonomy). Additionally, there may be many people who are not rational—children, for example. Rather than contending that all people *are* free and rational, then, Rawls holds that for the purpose of establishing principles of justice, those persons involved (the hypothetical contractors) must be free and rational. For the purpose of deciding which principles of justice are most fair (which is what justice as fairness requires), lack of freedom restricts some from following their own thoughts or inclinations, and irrationality hinders the entire group of people from operating and deliberating on the same plane.

In addition to freedom and rationality, each person must have “a conception of [his] good.”⁵⁰ To have a conception of the “good” means that each person must have an idea of what *he* wants, and what he strives to attain (his sense of “good”). For one person, this might mean attaining financial security; for another, it might mean having and successfully raising children. It is necessary for a person to have a conception of her good so that she can identify what she needs in order to attain it (e.g., access to a well-paying job, childcare, healthcare, public education, etc.). A further specification of a person that goes along with this is that each person is out to further her own interests. Each person has an individual conception of the good based on her circumstances (including knowledge, upbringing, and individual situation), and thus each person wishes to have the

⁵⁰ *ibid*, 19.

ability to achieve her own good. For example, if it is someone's conception of the good to be a devout Muslim, then it is necessary that there be no restrictions on practicing Islam in her society. Thus, she will advocate for a society in which practicing Islam is allowed; additionally, she will *not* want to allow her society to be a Christian nation, even if the majority of the citizens of her country are Christians.

While each hypothetical contractor in the original position has an individual concept of the good, and each wishes to pursue her own conception, once behind the veil of ignorance, these qualities must transform into something else. The contractors behind the veil of ignorance by definition do not know what their own conceptions of the good *are*, and nor do they know any particulars of their own circumstances; thus, behind the veil of ignorance, these qualities translate into the qualities of self-interest and caution. This means that, whatever the principles of justice are, they want to ensure that the principles will benefit them as individuals, no matter who they as individuals turn out to be. They are not going to be willing, for example, to endorse a principle of justice which would give women power over men, because that is not in their best interest should they turn out to be men. Nor can they have any feelings of generosity or charity or self-sacrifice which might influence them to endorse biased principles. Instead, they will recognize that it is to their own advantage to endorse principles that at least give everyone the opportunity to benefit equally.

Furthermore, the hypothetical contractors must be cautious. Although they are self-interested, they cannot be so self-interested that they become willing or even eager to take risks or gamble. It cannot happen, for instance, that a contractor, seeing that he has a 50% chance of being a man, advocate principles that would give power and privileges to men over women.

Thus, it is necessary that each person be self-interested and cautious so that there are no resulting imbalances or inequalities in the principles of justice that are an effect of risks taken or generosity exercised. The poor might think it unfair if they were merely told to pull themselves up by the bootstraps without any assistance whatsoever, and the wealthy might think it unfair if they were told they had to sacrifice all their wealth to the poor. The idea behind the veil of ignorance and the restrictions on people in the original position is that *everyone* in society be able to acknowledge that the resulting principles of justice are fair, regardless of any power structures which might be in place. Here you can see embodied the principle of impartiality: if everyone is impartial, then the principles of justice will be fair and will be to everyone's equal advantage.

One final rule is that each person in the original position must be "capable of a sense of justice."⁵¹ To be "capable of a sense of justice" is somewhat difficult to clarify. At the very beginning of *A Theory of Justice*, Rawls says that all people "have a conception of justice," which means that "they understand the need for, and they are prepared to affirm, a characteristic set of principles for assigning

⁵¹ *ibid*, 19.

basic rights and duties and for determining what they take to be the proper distribution of the benefits and burdens of social cooperation.”⁵² So, no matter what the specifics of a person’s conception are, as long as she has a conception of and recognizes the need for justice, she is potentially capable of helping to establish the principles of justice. Rawls makes the further clarification that “institutions are just when no arbitrary distinctions are made between persons in the assigning of basic rights and duties;”⁵³ thus, being “capable of a sense of justice” would involve having this understanding of the word “just.” For the purposes of establishing principles of justice, then, no arbitrary inequalities are to be established. (It does seem fairly obvious that in order to be involved in establishing principles of justice, one must be capable of conceiving of justice first.)

People are free, equal, and rational; they must be capable of conceiving of justice; and they must have a conception of their good and be interested in pursuing it. In the original position, where they *don't* know their own conceptions of the good, they are nevertheless always self-interested, with the idea that they should promote a society in which any conception of the good can be pursued. These are the qualities of a person which Rawls lays out in *A Theory of Justice*. There is one further rule of the game of constructivism (introduced by Ronald Milo) which I will now discuss.

⁵² *ibid*, 5.

⁵³ *ibid*, 5.

Universality

The rules of the game stipulate that the principles that are decided upon by contractors in the original position are true and universal. However, as Milo notes, “constructivism conceives of moral truths as truths about an ideal social order, rather than truths about the natural order of things.”⁵⁴ Moral truths are not on a par with scientific or physical truths, and they are not discoverable in the same way that scientific truths are discoverable. That is, the hypothetical contractors do not *discover* any fundamental facts about the universe. Rather, they *determine* them. What makes an act right or wrong is not that there is something naturally and really right or wrong about the act, but that the hypothetical contractors in the original position have determined it to be right or wrong. Thus, as Milo states, “constructivism claims ... that the moral wrongness of acts is *constituted* by their being prohibited by the norms chosen by the hypothetical contractors.”⁵⁵

In metaethical terms, what constructivism does is set up the prerequisites, so to speak, for determining moral norms and principles. The goal is fairness; the procedure is designed so that the principles eventually determined will be maximally fair. In order to accomplish this, he sets up the procedure such that all players are rational, equal and impartial (though in this sense, impartiality does not mean that they have no investment in the outcome, but rather that they lack knowledge of any circumstances which would *make* them partial). The rules of

⁵⁴ Milo (1995), 185.

⁵⁵ *ibid*, 202.

the game have been set: now, the game of establishing a system of normative ethics can be played. The principles which are decided on through this procedure are universal and objective because they are established by a (supposedly) universal and objective procedure. The principles are not pre-existing, and are not objective in the sense of being discoverable, pre-established, natural laws of the universe.

Altogether, this is what the metaethical framework for Rawls' theory of justice as fairness looks like. It includes the original position, including the condition of the veil of ignorance and the concept of a person necessary for Rawls' theory. Rawls presents all of this as a pure procedure for determining principles of justice; is it, however, really pure procedure? In the next part of the chapter, I will offer an interpretation of Rawls, and show the ways in which constructivism teeters dangerously on the edge of falling into universalism.

3. Interpretation

According to Milo, the principles determined in the original position are universal and true. This seems to advocate a system of moral universalism. Now, I will take a look at the extent to which constructivism does rely on universalist tactics to make its arguments. As Chapter Two demonstrates, I think that the pivotal difference between Harman's relativism and Nagel's universalism is the requirement of external, or objective, reasons. Harman does not require them and

in fact argues that there is no such thing. Nagel's argument is that objective reasons are the only viable form of moral reasons. I will argue that constructivism relies on objective reasons as well, and thus does put inappropriate constraints on moral reasoning.

I take Rawls to be arguing indirectly in favor of objective reasons because of the moral weight he places (through the veil of ignorance) on impartiality. He argues that a system of justice must be fair, and in order to be fair, all members must be initially equal. Furthermore, in order to determine fair principles of justice, all members must not know the particulars of their own individual situations; that is, they must be impartial. If they are knowledgeable about their own circumstances, they may knowingly or unknowingly be partial to them. Since Rawls constructs the veil of ignorance as a means to avoid the problem of partiality, it is clear that Rawls views partiality as interfering with fairness, and thus by extension with justice.

I take this to mean that impartiality is objectively favorable and is, in fact, a rule of reason in a similar way that modus ponens or the law of non-contradiction are rules of reason. I also take this to mean that impartiality is not only an ideal of *reason* but also one of *morality*. In other words, Rawls holds that impartiality, in addition to being a rule of reason, is also a rule of morality and justice. Impartiality is not just reasonable, it is also objectively morally ideal.

This can be explained through Rawls' explanations of the Reasonable and the Rational.⁵⁶ The Rational is that which people in the original position promote, and it consists in "sensible principles of rational choice," such as coherence, consistency, proper balance of consequences, and all the other methods by which people deliberate rationally. The Rational turns out to be reasons of self-interest. The Reasonable, on the other hand, are the *circumstances of the original position* which constrain the hypothetical contractors to be *reasonable*. The Reasonable establishes fairness (which one can presume to be reasonable); it enables the hypothetical contractors to fully practice rationality; and finally, "its principles limit ... the final ends that can be pursued."⁵⁷ The original position, then, is the representation of the Reasonable, and the methods of deliberation that the hypothetical contractors use in the original position, with all its constraints, are the Rational. Each person in the original position is completely rational, but the constraints of the original position (such as the veil of ignorance) make each person *reasonable* as well.

How does this show that Rawls takes impartiality to be objectively reasonable? Rawls does not at any point *argue* that fairness (which takes the form of impartiality in his arguments) is objectively reasonable. He merely *states* that fairness is reasonable. I take this to mean that it is objectively reasonable, since it would be strange and unbeneficial to his entire scheme if he took it to be merely subjectively reasonable. Since impartiality is the stance that is required to create

⁵⁶Rawls (1980), 528-30.

⁵⁷ *ibid*, 530.

fair principles, I take it that impartiality (at least in this theory of hypothetical contractors creating fair principles for a just society) is *also* objectively reasonable.

Is impartiality, then, an objective requirement for individual moral reasoning? In other words, is impartiality (or the attempt to achieve it) a moral requirement for an individual, in the same way that it apparently is for a society? Should an individual who is considering whether to steal an umbrella contemplate what an impartial judge of the situation might say? If so, then that would mean that contractarian constructivism requires a notion of objective reasons. If impartiality is an objective ideal, everyone has an objective reason to adhere to it, and thus to prefer a maximally impartial system of principles over any other system. Rawls himself does not advocate a normative principle of impartiality. At no point does he argue or claim that an impartial solution to a normative moral dilemma is the objectively right solution. At no point does he argue or claim that it is everyone's moral duty to consider what an impartial, unbiased spectator might say about any given situation. However, the metaethical framework he proposes supposes that impartiality is the fairest (and thus the most just) way of determining morality; the normative principles of justice he determines (which I haven't discussed in this paper) are based on this supposition; and thus I think it reasonable to interpret constructivism such that impartiality is an objective ideal of moral reason. Furthermore, though Rawls himself never took up the task of extending his theory beyond the level of democratic society, others have, notably Milo.

To illustrate how a principle of impartiality might operate at a micro-level, I will give an example. Imagine a situation in which there is one piece of cake and two people, each of whom wants to eat it. The fairest solution to this dilemma would be for them to share it. In order to ensure that the proportions are as equal as possible, the best thing to do would be for one person to cut the cake, and the other to choose the slice. This ensures that the cutter tries his utmost to cut equal slices; he supposes that the chooser will choose the larger slice if he cuts inaccurately, so it behooves him to be as fair as possible so that he will get the largest slice possible for him—half of the piece of cake. If he cuts unevenly, the chooser will choose the larger slice, leaving the cutter with less than half—clearly not an outcome the cutter desires. It is easy to see the similarities here with the original position—both individuals, not knowing what the outcome for them will be, strive to make the outcome as fair as possible so that they will each benefit the most they can. Neither is willing to risk ending up with a smaller piece of cake, so they agree on a procedure that will ensure their own self-interest. Fairness is the outcome that will satisfy *each* person that he has benefited maximally without causing strife. This is a perfect example of a commonplace situation which might call for impartiality as a rationally and morally required operating principle.

Impartiality, then, is an objective moral principle; it governs moral deliberation (rather than direct action, like Nagel's Golden Rule), but it is a principle that applies to everyone in situations of moral deliberation. The fact that it governs

moral deliberation rather than direct action is not problematic, however; deliberation is, after all, a specific kind of action. The principle of impartiality governs moral action taken in situations of moral dispute. As such, this means that the principle of impartiality is an *objective reason* for moral action. In a reason statement, it might take the form “I am going to cut the cake and you can choose, because it is the fairest and most impartial path of action” or “you have a reason to try to see the situation from outside of your own perspective, because everyone should attempt to be impartial.” (In this way, the impartiality principle ends up sounding awfully similar to the Golden Rule.) This is critical: if impartiality is an objective reason for moral deliberation or action, there are heavy implications. I will discuss these implications in the next section.

In summary, I have interpreted constructivism to include, in addition to the things stated outright by Rawls and Milo, three additional conditions: 1) impartiality is a rationally required aspect of fairness and thus of justice and morality; 2) as such, it is a moral principle as well as a rational principle; and 3) as an objective principle, it can be formulated as an objective reason for moral deliberation or action.

4. Implications

If some of the above sounds eerily familiar, it is because it is in some striking ways similar to Nagel’s theory about the Golden Rule as a universal law. In this section, I investigate these similarities; if they turn out to be significant, it is likely

that constructivism does not, in fact, offer a sustainable alternative to Nagel, and that it falls prey to the same arguments that Nagel comes up against.

First, the notion of personhood that Rawls relies on has fundamental commonalities with Nagel's. Nagel, you will remember, requires that each person conceive of herself not merely as an "I" but also as one person among other people. Though Rawls does not articulate his conception of personhood in quite this way, I argue that the implications of Rawls' conception are in end effect the same as Nagel's. Rawls constructs the veil of ignorance, citing the necessity of all people being impartial, and in doing so he takes away all of the things that make people individuals, all of the things that make a person an "I." The hypothetical contractors *are* mere persons among other persons; they don't have anything unique or particular that enables them to prefer their own points of view. Thus, Rawls implicitly shares with Nagel the idea that people, in order to be able to make the "right" moral decision (that which is most just), need to understand themselves not as individuals, but as people among other people. One might say that in Rawls' case, it is a stipulation merely for the hypothetical contractors that they have this conception of themselves, and since they are just that—hypothetical—it is not necessary that this conception extend to real people. However, in light of my interpretation of constructivism in which impartiality is an objective principle (and not merely a hypothetical principle), a similar interpretation of this notion of personhood is entirely reasonable. In fact, I contend that impartiality by definition *requires* this conception of personhood.

Impartiality, insofar as it demands that things that distinguish individuals be put aside, relies on a conception of personhood in which people are first and foremost persons among others equally persons.

Of course, there are differences in their conceptions of personhood. Nagel insists that people consider the Golden Rule because not to do so would be inconsistent with the conception of personhood. Rawls, on the other hand, constructs the veil of ignorance because he assumes that people are self-interested, and so even behind the veil of ignorance, all people supposedly act selfishly, despite being unbiased. Nevertheless, the implications of the details of their conceptions end up being roughly equivalent: an individual has limited flexibility in both her self-conception and in her conception of morality in general.

Starting from this conception of personhood, then, both Rawls and Nagel develop theories of morality which have *objective moral requirements*. For Nagel, this conception of personhood leads him to require all people to hold the same moral standards for themselves that they would for other people, since consistency with the conception shows that there is nothing special about *me* that allows me to be an exception. This leads him to the conclusion that everyone has an *objective reason* to consider the so-called Golden Rule in moral deliberation. For Rawls, in comparison, the conception of personhood in addition to his conception of justice as fairness leads him to require that all people *deliberate together* to determine fair principles of justice. According to my above

interpretation, this leads to the objective moral requirement that one consider the least partial course of action in moral deliberation. This could be a version of the Golden Rule, in which one is required to think not “what would I feel if I were that person?” but “what would I think if I were impartial?”

There are some major differences between Rawls and Nagel, of course. For one thing, Rawls begins with more initial assumptions than Nagel does. Rawls appeals to some already existing collective sense that fairness is an objectively desirable ideal, and proceeds from there with values like equality and impartiality, hardly arguing for them so much as asserting their reasonableness. Nagel, on the other hand, assumes less about human nature; his argument in favor of his objective reason is more logical than Rawls’, if equally misguided. Another obvious and significant distinction between the two is that for Nagel, the objective principle *exists* in itself; he takes his argument to prove its existence. For Rawls, on the other hand, there are no eternal and somehow “pre-established” moral laws to be discovered. Rather, his hypothetical contractors *determine* them, and they are true *because* they are the laws determined by the hypothetical contractors.

5. Conclusion

In this chapter, I have shown how constructivism, despite hoping to avoid the universalist trap by claiming to offer an objective *procedure* for arriving at moral principles rather than offering objective principles themselves, does not succeed at providing an alternative to Harman’s moral relativism. According to my

interpretation of Rawls, constructivism still requires that individuals hold impartiality as an objective principle to adhere to when engaging in moral reasoning, and I have argued that this means that the principle of impartiality becomes an objective reason. Once there is an objective reason, all of Harman's and Williams' arguments can jump in with full force to show that objective reasons are insupportable. Rawls' supposed pure procedure for determining principles of justice evidently has substantive consequences.

If Rawls would vehemently disagree with my interpretation of his theory as a metaethical one, that is not a problem. If his theory is not metaethical, it certainly doesn't provide an alternative to Harman's relativism, which *is* metaethical. Whether or not Rawls (or any other constructivists) have intended constructivism to be interpreted in this manner is irrelevant to my thesis. What I have done is articulated how constructivism *as* a metaethical theory doesn't really work as an alternative to moral relativism. It ends up formulating a supposedly objective and neutral procedure for determining principles of justice, a procedure which would constrain our moral reasoning, in a way that the procedure has substantive consequences and endorses objective principles. The consequences for such a theory are similar to the consequences for Nagel's theory: it crumbles under the argument against objective reasons.

Rawls would want to resist an association with Nagel and universalism. I think he would argue that my interpretation of impartiality as an objective principle that governs moral deliberation and thus moral action is misguided;

perhaps he would say that impartiality isn't meant to be an objective *principle*, but rather a mere objective *constraint* on reasoning. A constraint on reasoning is different from a principle in that, rather than providing an objective reason for taking particular action, it merely restricts the ways in which we can reasonably engage in moral reasoning. If I disregard a constraint on moral reasoning, I am not breaking a moral law, but I *am* being unreasonable. In the next chapter, I will explore this idea in more depth.

CHAPTER FOUR:
OBJECTIVE CONSTRAINTS ON MORAL REASONING

In the previous chapter, I argued that Rawls does not differ enough from Nagel to make his arguments acceptable to Harman. I argued that constructivism's requirement of impartiality amounts to the same thing as Nagel's requirement of the Golden Rule. Interpreted thus, constructivism doesn't succeed in offering an alternative to moral relativism because it is itself squarely universalist. This is a pretty strong argument, and there are probably many constructivists (perhaps Rawls and Milo included) who would not support this interpretation. Whether my interpretation is supportable, however, is irrelevant. Even if it is not supportable, even if impartiality is not an objective principle that implicitly endorses objective reasons, impartiality is still an imposed *constraint* on moral reasoning (a requirement about how we engage in moral reasoning). Objective constraints on moral reasoning imply that there are objective ways in which to think morally; there are conditions or limitations on how we reason morally. Objective constraints on moral reasoning are supposedly void of content claims about morality; they are supposedly merely procedural. For constructivism, impartiality is the primary constraint. In this chapter, I investigate whether objective constraints on moral reasoning are as neutral as they seem, and I show that impartiality as an objective constraint on moral reasoning fails.

In order to do this, it's important to take another look at Harman: Harman, too, proposes certain objective constraints on moral reasoning. It seems as though if my attack on constructivism succeeds, Harman's relativism would suffer in turn. I will show, however, that there are different types of constraints on reasoning. What I call *first-order constraints* are objective; *second-order constraints* are not objective. First-order constraints have two properties: 1) they are valueless; and 2) they govern subjective (rather than inter-subjective) reasoning, that is, they do not govern social deliberation but only internal, personal processes of moral reasoning. Second-order constraints might contain value judgments, or they might govern how we navigate moral disagreement between people rather than governing purely internal processes of reasoning. I will develop these concepts more fully throughout the chapter. In the course of this chapter, I will show that impartiality is a second-order constraint, and thus is not objective. I will also take a look at some of Harman's constraints and evaluate them according to the same standards. I will show that at least one of Harman's constraints is first-order, so it is in fact objective, and at least one of his constraints is second-order, so it is not objective and is not in fact a valid constraint on moral reasoning.

1. Harman's Objective Constraints

Recall Harman's argument about relativism: one can't be objectively morally right or wrong, but one can be morally right or wrong in relation to one's own moral framework. In other words, one *can* be wrong if one is inconsistent with

one's own moral framework. For example, if I hold that tripping a person for my own amusement is wrong, and then I go and trip someone for my own amusement, my behavior is morally wrong because it is inconsistent with my own beliefs. Thus, for Harman, consistency—an ideal of reason—is also an objective constraint on moral reasoning. Consistency is not the only ideal, either. In chapter four of *Explaining Value*, Harman states that “having sufficient reasons to *D* consists in the situation's being such that one would *D* if one were fully informed, fully rational, aware of relevant arguments and able to follow and understand these arguments, not suffering from the weakness of will, and so forth.”⁵⁸ In other words, all of those principles (being fully informed, rational, etc) seem to be constraints on reasoning for Harman. I can be morally wrong if I don't follow a set of standards; these standards *seem* to be simply standards of reason, but because they can render my moral decisions wrong if I fail to follow them, they are also objective constraints on moral reasoning, whether or not Harman intends them to be. In other words, if Rawls' standards for moral reasoning are unjustified (which I will argue they are), so must Harman's be; they rely similarly on some collective intuitions about what moral reasoning entails. If Harman wishes to show that impartiality is unjustified, he seems to have fallen into his own trap for objective reasons.

Is there a difference between Harman's constraints and Rawls'? Harman would not support Rawls' constraint (fairness/impartiality), but clearly, he does

⁵⁸ Harman (2000), 67.

support his own (though he is very vague about defining them). Harman would want, if he didn't so assiduously avoid the problem altogether, to argue that his constraints *are* different in nature from Rawls'. Taking a look at two of the constraints that he more clearly proposes—consistency and full information—as well as Rawls' proposed constraint, I will, with the use of examples, try to tease out the differences between the three proposed constraints, and work out whether any of them are actually objective constraints on moral reasoning.⁵⁹ I approach this on the level of reasons, with the goal being to decide whether there are, in fact, objective constraints on reason of any sort that should always govern our moral deliberations.

2. Consistency

The first supposedly objective constraint on moral reasoning that I will consider is consistency. Harman's theory relies heavily on consistency as an objective constraint of reason. Here's why: his argument about moral relativism isn't that it is impossible for someone to be wrong, but rather that it is impossible for someone to be *objectively* wrong. In other words, it is possible for someone to be wrong in relation to one's own system of moral values. Suppose, for example,

⁵⁹ Though Harman does, at various points in his writings, refer to other potential constraints as well (he mentions, among others, the ability to understand relevant arguments/information and the state of "not suffering from weakness of will" (*Explaining Value and Other Essays*, 67)), I am going to stick with consistency and full awareness of relevant information for the following reasons: 1) consistency is a crucial aspect of Harman's theory as a whole, since he says that one *can* be wrong in relation to one's own moral framework, i.e. one can be wrong if one is inconsistent with moral values one holds to be true; and 2) Williams makes quite a fuss about the need for full information, even if Harman tries to avoid it, so I think it is important to address it here.

I hold the value that it is wrong to steal other people's things. One day, it is pouring as I am leaving the library, but I've forgotten my umbrella, and so I furtively snatch an umbrella from the foyer and run away with it. This action is morally wrong because I hold it to be true that stealing is wrong. Since I acted in a manner clearly inconsistent with my own values, I was wrong. This is a vital element of Harman's argument, because it makes it clear that people can (and probably often do) do immoral things; moral relativism doesn't mean "anything goes." However, this clearly sets up consistency as an *objective constraint* on moral reasoning, because no matter what, in all situations, if I act in a manner that contradicts my beliefs or values, I am wrong. This is the case not just for me, but for everyone.

I consider consistency to be a restatement of the logical law of non-contradiction: it cannot be the case that both A and \sim A. It is a logical impossibility. This does transfer to moral reasoning, as far as I'm concerned, because moral statements are still statements. It cannot be the case that "it is wrong to steal according to my moral framework" and "it is not wrong to steal according to my moral framework" are both true. If I hold that the first statement is true, and then I steal something, either of two things could be the case: 1) it is not wrong to steal according to my moral framework; or 2) I am doing the wrong thing by stealing. Option 1) cannot, according to the law of non-contradiction, be the case; thus, it must be the case that I am doing the wrong thing according to my

moral framework—and “wrong” here does not just mean *morally* wrong, it means *rationally* wrong.

Consistency is a rational principle; in order to be consistent with my own moral values, which is what Harman requires, I need only two things: my own values, and a rational capacity. It is difficult to provide an argument in favor of the need for rational capacity in moral reasoning; indeed, it would be rather like trying to define a word by using that same word. Using a rational argument to argue for rationality wouldn't help with anything. If you reject rationality, then you're irrational by definition, and that's that. This brings to mind a famous essay by Lewis Carroll, “What the Tortoise Said to Achilles,” in which Achilles is trying to use modus ponens to prove modus ponens to the turtle, and it doesn't work—and quickly becomes absurd.⁶⁰ Using logic to try to prove logic doesn't work if the audience doesn't accept logic.

However, it becomes clear that consistency is a justifiable constraint on moral reasoning when you try to think of counter-examples. Suppose that Janet firmly believes that it is wrong to plagiarize. One day, she is up late writing a paper, and she has some unpublished writing from her mother on the very topic of her paper, so, in the interest of getting it done, Janet plagiarizes her mother's work. No one will ever know, except that I know, because Janet told me. “Janet,” I say, “don't you think it is wrong to plagiarize?” “Yes, I think it's always wrong,” Janet says. “Do you grant yourself moral privileges that no one else has, then?” I ask. “No, I

⁶⁰ Carroll (1895).

don't—I think it's wrong for everybody, including me, to plagiarize," Janet says. "So you did the wrong thing?" I ask. "No," Janet says, "I wasn't wrong to plagiarize."

There is something wrong here (no pun intended). It is a clear contradiction that Janet thinks both that it is wrong (for everyone, including herself) to plagiarize, but that she wasn't wrong to plagiarize. It seems absurd to try to say that even if it's wrong, it might not necessarily be wrong. It's a simple case of the law of non-contradiction: it cannot be the case that both A and $\sim A$, that it's both wrong to plagiarize and not wrong. Here you jump in and say, "well, maybe it's wrong in some ways, but not wrong in others; it's wrong, for example, in that it's essentially stealing someone else's work and lying about your own, but it's *not* wrong in that it allows Janet to get enough sleep and focus her mental energy on something much more important than a single academic paper." This may well be the case, from *your* perspective, but you forget that we're talking about Janet—that's the whole point of moral relativism. The point is that *she* thinks it's unequivocally wrong to plagiarize, and even admits that it's wrong for *her* to plagiarize, and thus she can't possibly reasonably claim that she wasn't doing the wrong thing. She may very well have had compelling reasons to plagiarize, and maybe those compelling reasons were more compelling to her than doing the morally right thing, but that has *no* effect on the *moral value* of her actions.

It is pretty clear, then, that consistency is a first-order constraint on moral reasoning—it is void of value because it is a basic rational principle, and it clearly

governs subjective reasoning as opposed to inter-subjective reasoning. In other words, the only claim it makes is that an individual can't reasonably hold two directly contradicting beliefs to both be true. It makes no claims about anything external to the individual person's own mind. Consistency is a first-order constraint on moral reasoning, and thus it is objective.

3. Impartiality

Rawls claims that fairness is an objective ideal; in his original position, which is necessary for achieving fairness, he requires impartiality as a constraint on reason. Rawls says little about the implications of his theory on *real* people in *real* moral situations (as opposed to hypothetical contractors establishing a hypothetical order), but Ronald Milo does; he argues implicitly that impartiality is an objective constraint for real people in real situations.⁶¹ Thus, for the purposes of this paper, I consider impartiality to be an objective constraint on reason according to contractarian constructivism. I will argue, with the use of an example, that impartiality is *not* an objective constraint on moral reasoning.

⁶¹ Milo (1995) argues that moral truths are "truths about an ideal social order" and that it is the case that "a certain kind of act is wrong, for example, just in case a social order prohibiting such acts would be chosen by rational contractors under suitably idealized conditions" (184). The "suitably idealized conditions" he refers to here are the conditions of Rawls' original position, which include impartiality. If moral truths are determined under such circumstances, and if individuals in a non-idealized, non-hypothetical, and real society are trying to determine what's morally right and wrong, then they would be objectively required to try to determine what truths would be determined under the idealized conditions. Thus, they would be required to strive for impartiality, since one of the conditions for determining moral truths in the original position is impartiality. In this way, I take Milo to be arguing for an interpretation (or extension) of Rawls in which impartiality is an objective constraint on individual moral reasoning.

Suppose Paul is the parent of a third-grader. He learns one day that his daughter's teacher is a lesbian. Paul demands a meeting with the principal of the school, Ruth, and insists that the teacher be removed from the school. He is an evangelical Christian and thinks homosexuality is immoral and disgusting; he also believes that the teacher could "rub off" on his daughter or might teach her that there is nothing sinful about homosexuality, that it is perfectly normal, and that she should be open-minded about it. All of this goes directly against Paul's religious and moral beliefs. Ruth, however, vehemently refuses to fire the teacher; she thinks it is bigoted, discriminatory and just deeply wrong to fire someone based on her sexual orientation. (Suppose, for the sake of argument, that it would not be illegal, as it would not be in many states.) She also insists that the teacher will not teach the children anything other than acceptance and open-mindedness towards others; she will *not* be encouraging the students to be gay themselves. Paul will not relent, and the two reach an impasse. A guidance counselor happens to come into the office at this time, and after listening to the two of them, he says, "Can't you pretend you have no personal investment in the outcome of your decision? Can't you try to let go of your respective biases for the benefit of making a decision? Can't you try to decide what an impartial arbiter might say?" The guidance counselor wants them each to let go of their particular positions and beliefs, and consider what would be the fairest outcome to all involved, including those of the teacher in question and her students.

Paul and Ruth consider briefly, and then they both decide that even if they *can*, they *will not*. They are each so completely and utterly convinced that *they* are right that they would consider it wrong to compromise. For Paul, an impartial decision would mean that he might end up having to agree to something that he *personally* would consider a sin; for Ruth, an impartial outcome might mean agreeing to something she finds bigoted and discriminatory. They each have too much personal investment in the outcome that an attempt to find an impartial solution might be sacrificing something central to their very personhood. In this situation, though I clearly have a personal bias and believe (according to my own moral framework) that Ruth is right, I also maintain that *neither* of them is compelled by reason to attempt to be impartial. Impartiality, in this case, might be a helpful negotiation technique and it might move them forward in their conversation, but I do not see how it can possibly be an *objective constraint* on reason. In other words, neither Paul nor Ruth is wrong, irrational, or unreasonable to refuse to strive for impartiality.

Iris Marion Young⁶² argues along these lines, stating that a theory of justice such as Rawls' "masks as disinterested truth what really expresses the interests and values of the dominant class."⁶³ Impartiality, then, isn't really entirely impartial; instead it may be assuming certain dominant (potentially unjust) structures or norms (such as capitalism). As Young says, there is "no 'justice in itself' independent of the particular economic forms and social relations which

⁶² Young (1981)

⁶³ *ibid*, 293.

engender and embody particular conceptions and principles of justice.”⁶⁴

Young’s arguments amount to a similar indictment of impartiality as an objective constraint on moral reasoning to my own: impartiality is not void of content. The notion of impartiality as an objective constraint on moral reasoning itself is, ironically, not impartial. It is not neutral. It carries in it a moral judgment (that people ought to be unbiased) which, as Young notes, may in fact be a value judgment of a particular group or groups of people. Impartiality is not objectively *wrong*, but, as Harman would point out, it is relative to a particular moral framework. Young claims that it is the moral framework of a “dominant class”; though this language is vague, it might refer to any group or combination of groups of people in dominant positions in society, such as (for example) men, the upper class, or white people.

Impartiality differs from consistency, then, in that it involves *values*. In some or even many cases, impartiality might require sacrificing values or moral beliefs for the sake of the principle of impartiality. Impartiality may even *be* a value itself; it may very well be the case that there are people for whom fairness and impartiality outweigh other moral considerations, but if there are people for whom it doesn’t, it hardly seems justifiable to call them unreasonable or irrational, in the way that people who are inconsistent with their own beliefs are irrational.

Consistency, as I said above, does not involve anything other than one’s own values and an iota of rationality. Consistency governs one’s *subjective* moral

⁶⁴ *ibid*, 293.

reasoning. Impartiality, on the other hand, involves others' values aside from one's own. It governs social, or *inter-subjective*, moral deliberation. As such, impartiality is a second-order constraint on moral reasoning, and is not objective.

Another way in which the difference between consistency and impartiality can be explained is as follows: impartiality relies on a specific conception of personhood, one that Harman wouldn't necessarily agree with. Rawls and Nagel both think that a crucial aspect of personhood is recognition of oneself as one person among other people; one is no more or no less a person than others, and this affects how one treats them morally. For Nagel, the fact that one is just a person among other persons leads him to the conclusion that in moral deliberations, we ought to consider others' feelings the same way we consider our own. For Rawls, a similar conception of personhood leads him to the conclusion that we ought to consider each other all originally equal, and that in order to be fair, all contingent aspects of an individual ought to be cast aside for the benefit of all. Harman would not agree that we are rationally forced to admit that we find ourselves (practically speaking) in a world with other people. And even if he *did* support that notion, the pure *fact* that we live in such a world would certainly not lead him to any conclusions about morality. In other words, Harman would not agree that the fact that one is a person among many people ought to imply that there are any objective constraints on one's moral reasoning.

4. Full Information

I have discussed both consistency and impartiality. Now I turn to full information as a potential objective constraint on moral reasoning. Both Harman and Williams at different points in their writings suggest or argue that full information is a constraint on moral reasoning.⁶⁵ The idea is that in order to be right or make the right decision (“right” according to your own moral framework), you need to be aware of all the facts. Remember my skydiving example from Chapter Two: if the parachute is broken, and if you *wouldn't* jump out of the plane with a broken parachute, but you don't realize that the parachute is broken so you jump anyway, then your decision to jump is based on false reasoning. It is based on false reasoning because one of your reasons for jumping (whether implicit or explicit) is that the parachute is functional. Since the parachute is broken, however, your reason for jumping is wrong. Thus, in order to reason properly in moral situations, you need to have full information about the situation and be knowledgeable about any relevant facts or considerations.

Is full information really an objective constraint on moral reasoning, however? If full information is to be a true objective constraint, then it seems to imply that one ought to try to obtain all necessary or relevant information; after all, I *do* want to be right. It would be inconsistent of me *not* to want to be right. It may not be that simple, however. In order to help shed light on the notion of full information, I will consider an example. This will help to determine whether the requirement

⁶⁵ Harman (2000), 67; Harman and Thomson (1996), 46-7; Williams (1981), 103-4.

of full information is a first-order requirement like consistency, or a second-order requirement like impartiality.

Martha hates politics. If she is watching television and anything about the upcoming presidential comes on, she turns off the TV. Politics bore her and make her despair about the nation and the world. She would much rather spend her time reading William Faulkner or doing yoga. However, Martha isn't entirely apathetic. There are political issues she cares about; for example, Martha is a pacifist and is morally opposed to the war in Iraq. If she *knew* a particular candidate had voted for the war, she definitely would not vote for that candidate. Every time anything political comes up, though, Martha avoids it, so she never finds out which candidates voted for the war and which didn't. November rolls around and Martha votes for Hillary Clinton. Since Martha would not have voted for Hillary Clinton had she known Clinton had voted in favor of the war, it seems that Martha's vote is based on false reasoning; after all, she *refused* information that would have changed her mind. Additionally, since Martha is morally opposed to war, it seems as though in refusing full information, she has made a decision which is actually inconsistent with her moral beliefs, although she is not aware of that fact. Is Martha wrong for turning off the TV or not reading the newspaper?

At first glance, it seems as though she might be. There is easily accessible information available that she is actively refusing, the attainment of which would make her act in a way consistent with her moral beliefs, and the unawareness of

which makes her act in a way *inconsistent* with her moral beliefs. Everyone has a reason to desire not to be inconsistent with one's own moral beliefs, because *not* to desire it would be irrational.

However, I think the fact that Martha is *unaware* of her inconsistency makes all the difference in this scenario. Desire for full information, like impartiality, carries value. It has moral weight in the same way that impartiality does, and unlike consistency, which is a pure rational principle. It may very well be the case that the result of Martha's reasoning is inconsistent with what she *would* have done had she known something else, but the fact is that that is a hypothetical situation. Martha does not value political curiosity, and in fact she makes a point of valuing a lack of political curiosity, so to say that there is an objective constraint in place that requires her to obtain political information is to say that there is an objective constraint that goes against her values. Additionally, since it is impossible to be always fully aware (no one can know everything about everything), the form that the requirement would have to take would be that we are all objectively required to seek information. Think of a simple farming society, for example, in which people live and plan according to an instinctive understanding of the earth, as well as stories, myths, and lore that are passed from generation to generation. There may very well be scientific knowledge that such a society lacks which might change the way they choose to farm certain crops, but the fact that they lack the desire to seek out such information is neither irrational nor morally unreasonable.

The requirement for full information, then, is not valueless. The value that it carries with it is curiosity, or desire for more knowledge. It is also intersubjective, in that it clearly demands outside information, information that isn't merely subjective, or part of the subjective motivational set. Thus, the requirement for full information is clearly a second-order constraint, not a first-order constraint; it fulfills both stipulations of a second-order constraint. The requirement for full information is *not* objective. Rather, it is a constraint that is relative to a particular moral framework.

5. Conclusions

In this chapter, I have argued that there are different sorts of supposed constraints on moral reasoning, only one of which—first-order constraints—can be objective. I used consistency as an example of a first-order constraint; there may be others. What is important is that the constraints that Rawls and contractarian constructivism offer are not, in fact, objective. They fall under the category of second-order constraints, and as such they cannot be universal and objective constraints on reasoning. They are not morally neutral.

If first-order constraints are the only truly objective constraints on moral reasoning, what *are* second-order constraints? Second-order constraints are constraints on reasoning that must be relative to a moral framework. Impartiality may well be a requirement on moral reasoning that people who subscribe to particular beliefs about fairness as a moral ideal adhere to. Desire for full

information might be a requirement for those who hold curiosity as a particular value. Since there are no objective reasons to strive for or believe in either of these two (or any other) moral values, however, these second-order constraints are not truly objective.

So what is the difference between a second-order constraint on moral reasoning and an objective moral principle? Maybe there is no difference. An objective reason arises out of an objective moral principle, such as Nagel's Golden Rule; the idea is that the objective moral principle gives everyone an objective reason to take certain moral actions, make certain moral decisions, or deliberate morally in a specific way. Since objective reasons are insupportable, though, as both Harman and Williams argue, and since the only way moral principles *could* be objective would be if they were backed by objective reasons (as Harman also argues), there is no possibility for moral principles. Similarly, second-order constraints on moral reasoning must also be backed by objective reasons in order for them to be objective; since they are not purely reason, but contain moral value, there would need to be an objective reason that everyone would have to hold those particular values. First-order constraints do *not* rely on objective reasons, because they are purely rational constraints. Think again of Achilles and the Tortoise;⁶⁶ there does not need to be a *reason* for someone to be rational. Rationality justifies itself, and it is impossible to justify it by any other means.

⁶⁶ Carroll (1895).

The primary conclusion from this is that contractarian constructivism does not objectively apply to situations of individual moral reasoning. I do not argue that justice as fairness is invalid as a political ideal for a system of justice; rather, I have come to the conclusion that Rawls' project cannot be extended to the realm of ordinary people doing ordinary things. I have argued that I just don't think it is unreasonable for people not to want to strive for impartiality. Rawls, then, does not offer a solution to the relativism vs. universalism debate. On the one hand, I have shown that impartiality seems to be nothing less than an objective moral principle, and thus, an objective reason. On the other hand, I have used a different line of argument to show that even if you *don't* use similar attacks on Rawls as on Nagel, impartiality still turns out to be merely a second-order constraint on moral reasoning, which may end up being an "objective" moral principle after all. Either way, constructivism is not a happy metaethical alternative to moral relativism. I remain just as convinced by Harman's moral relativism now as I was when I started.

There are things that Harman got wrong, as this chapter has made clear. Harman, whether aware of it or not, *also* imposes supposedly "objective" constraints on moral reasoning. In evaluating two of those constraints, consistency and the necessity for full information, I have come to the conclusion that consistency is a first-order constraint but full information is, in fact, a second-order constraint and is therefore just as insupportable as impartiality or Nagel's Golden Rule. (Harman, it seems, has some inconsistencies himself!) I do not

reject the whole of Harman's theory on this basis, however. In fact, I think he is right that there are no objective moral principles because there are no objective reasons everyone must have to abide by or believe any given moral principle. Harman doesn't just think morality is a free-for-all, though; there are constraints on moral reasoning just as there are constraints on any other kind of reasoning (mathematical, political, etc.). I splinter off somewhat from Harman, however, and offer my own argument for what kind of constraints there are, since Harman has some inconsistencies. I have argued that only first-order constraints, or constraints that govern only subjective or internal deliberation and are content-neutral, are actual objective constraints on moral reasoning; they are the only kind of constraint on moral reasoning that are purely rational.

I have said that I am just as convinced by moral relativism now as I was at the beginning of this thesis. That does not, however, mean that I think moral relativism is necessarily true. What is clear to me is that contractarian constructivism and Nagel's moral altruism ARE NOT true. I do think Harman's moral relativism is much more supportable and plausible than altruism or contractarian constructivism, but as I end I should make one thing clear: this thesis is about *reason*. It investigates what *reason* objectively demands of morality. I have come to the conclusion that reason as such demands very little of morality, or perhaps better put, morality is constrained very little by objective reason. This leads me to the conclusion that morality must extend far beyond the scope of reason, and that *if* there are in fact universal moral laws, we cannot

discover them in the same way that we can discover laws of science, by using logic and pure reason. But that is a matter for another thesis.

APPENDIX A:

NAGEL'S LOGICAL FORMULATION OF OBJECTIVE REASONS

Nagel takes a reason to have the following form: “every reason is a predicate R such that for all persons p and events A, if R is true of A, then p has prima facie reason to promote A.” He gives an example of the philosopher G. E. Moore standing in the middle of a street with a truck approaching. In this case, it is probably the case that Moore has a reason to move out of the middle of the street. The following are all various forms that, according to Nagel, the reason predicate (R) can take:

- (a) that the act will prolong G. E. Moore's life;
- (b) that the act will prolong his life;
- (c) that the act will prolong someone's life.

So what's the difference between (a), (b), and (c)? In (a) and (c), “G. E. Moore” and “someone” can be understood as constants, while (b) contains what Nagel calls a free agent-variable, which means that the reason predicate contains the variable p: “(p, A) (If A will prolong p's life, then p has a reason to promote A)”. This is a free agent-variable because the “his” is incomplete—it begs the question, “whose?” It needs a reference to an agent in order to be a well-formed logical sentence (in logical terms, it needs a quantifying p or $\exists p$ so that the p in Rp is not a dangling variable). The agent that it would refer to would end up being

either G. E. Moore or someone, the constants in (a) and (c). Thus, in order to be comprehensible, (b) relies on (a) or (c).

Translated into English, this simply means that the inclusion of p in R makes the reason subjective because p doesn't refer to any external agent. (In order for the sentence "the reason for moving out of the middle of the road is that it will save his life" to make sense, "his" has to refer to an agent.) The agent it can refer to could be, among others, "G. E. Moore" or "someone." The subjective free agent-variable "his" relies on (a) or (c) to be complete. Thus, (b) by itself forms a logically invalid sentence, and it can't by itself be a reason. So, for Moore, the reason that he should move out of the middle of the street is not that it will save his own life; it is that it will save the life of someone, G. E. Moore, a person. Subjective reasons (such as doing something for the reason that it will promote MY interests) don't stand on their own.

WORKS CITED

- Benedict, Ruth. "Anthropology and the Abnormal." The Journal of General Psychology 10 (1934), pp. 59-82.
- Carroll, Lewis. "What the Tortoise Said to Achilles." Mind 4.14 (April 1895), pp. 278-280.
- Harman, Gilbert. "Moral Relativism Defended." The Philosophical Review 84.1 (Jan 1975), pp. 3-22.
- Harman, Gilbert. *Explaining Value and Other Essays in Moral Philosophy*. Oxford, Oxford UP: 2000.
- Harman, Gilbert and Judith Thomson. Moral Relativism and Moral Objectivity. Oxford: Blackwell Publishers. 1996.
- Kuhn, Thomas. The Structure of Scientific Revolutions. 2nd ed. Chicago, University of Chicago Press, 1970.
- Milo, Ronald. "Contractarian Constructivism." The Journal of Philosophy 92.4 (1995), pp. 181-204.
- Nagel, Thomas. The Possibility of Altruism. Oxford: Clarendon Press, 1970.
- Nagel, Thomas. What Does It All Mean? New York: Oxford University Press, 1987.
- Rawls, John. "Justice as Fairness." The Philosophical Review 67.2 (1958), pp. 164-194.
- Rawls, John. A Theory of Justice. Cambridge, MA: Harvard University Press, 1971.

Rawls, John. "Kantian Constructivism in Moral Theory." The Journal of Philosophy 77.9 (1980), pp. 515-572.

Rawls, John. "Justice as Fairness: Political not Metaphysical." Philosophy and Public Affairs 14.3 (1985), pp. 223-251.

Williams, Bernard. "Internal and external reasons" in Moral Luck. Cambridge: Cambridge University Press, 1981. 101-113.

Wong, David B. Moral Relativity. Berkeley: University of California Press, 1984.

Young, Iris Marion. "Toward a Critical Theory of Justice." Social Theory and Practice 7 (1981), pp. 279-302.