Blame and Moral Indifference: Philosophical Problems with Praise-Blame Asymmetry

Exam No. B008602

MSc by Research in Philosophy

2011

Declaration

I have read and understood The University of Edinburgh guidelines on Plagiarism and declare that this written dissertation is all my own work except where I indicate otherwise by proper use of quotes and references.

Abstract

In this paper, I examine the empirically-based folk concept of Praise-Blame Asymmetry. I note that when laypeople make assessments of moral responsibility, they sometimes appeal to the agent's motivations for acting, and they sometimes do not. In order to figure out why they make these judgments and how these judgments are justified, I construct and explore an empirically-based folk theory of moral worth. While examining the folk theory of moral worth, I realize that laypeople inconsistently apply the criteria for an agent's being morally responsible and that it is these inconsistently applied criteria that are fueling asymmetrical judgments about praiseworthiness and blameworthiness. I try to resolve the inconsistencies in the folk theory of moral worth by appealing to the intend/foresee distinction and, eventually, by including new criteria for ascriptions of moral responsibility: reasons for action. Believing that Nomy Arpaly's theory of moral worth can successfully support Praise-Blame Asymmetry; I flesh out her theory, hoping to resolve the inconsistencies and problems that I noted in the original folk theory of moral worth.

Introduction

In recent years, moral psychologists have investigated the underlying cognitive processes that people engage in when they make judgments about the morally relevant actions of other agents. Of particular interest, is how laypeople ascribe praiseworthiness and blameworthiness to others. These evaluations are important because philosophers look to lay intuitions about praise and blame when they develop their theories of moral worth. It is generally taken to be the case that lay intuitions about what is right and wrong or praiseworthy and blameworthy are correct, and it is the endeavor of philosophers to demonstrate how these judgments are justified. When I say that philosophers demonstrate how lay intuitions about morality are justified, I mean to say that philosophers try to demonstrate that these intuitions can be made to fit into a neat and coherent theory of morality that can be more generally applied. Philosophers try to identify and clarify the criteria that laypeople employ when they make these judgments in order to establish generalizable rules for making moral evaluations. It might, however, be a mistake to assume that lay intuitions about moral judgments are reconcilable with any philosophically defensible theories and that they may, therefore, be justifiable.

In one of his early experiments on folk judgments about intentional action, Joshua Knobe discovered what is now known as the Side-effect Effect or the Knobe Effect.[1] Subjects in his experiment were presented with one of the following two scenarios:

> **Harm Condition**: "The vice-president of a company went to the chairman of the board and said, 'We are thinking of starting a new program. It will help us increase profits, but it will also harm the environment.' The chairman of the board answered, 'I don't care at all about harming the environment. I just want to make as much profit as I can. Let's start the new program.' They started the new program. Sure enough, the environment was harmed."[2]

After being presented with the harm condition, subjects were then asked to determine, on a scale ranging from 0 to 6, how much blame the Chairman deserved for his actions, and they were asked to say whether they believed that the Chairman intentionally harmed the environment.

---

[1] Joshua Knobe, "Intentional action and side effects in ordinary language," *Analysis*, 64 (2003): 81 – 87. The phenomenon is called the Side-effect Effect because it is associated with the unintended side-effects of acts and the Knobe Effect because this is the name of the author who discovered this particular phenomenon.
[2] Knobe 191.

> **Help Condition**: "The vice-president of a company went to the chairman of the board and said, 'We are thinking of starting a new program. It will help us increase profits, and it will also help the environment.' The chairman of the board answered, 'I don't care at all about helping the environment. I just want to make as much profit as I can. Let's start the new program.' They started the new program. Sure enough, the environment was helped."[3]

After being presented with the help condition, subjects were then asked to determine, on a scale ranging from 0 to 6, how much praise the Chairman deserved for his actions, and they were asked to say whether they believed that the Chairman intentionally helped the environment.

The two different conditions generated two very different responses from the participants in the experiment. Subjects who evaluated the Chairman in the harm condition largely determined that he had harmed the environment intentionally. Subjects who evaluated the Chairman in the help condition largely determined that he had not helped the environment intentionally. Subjects' judgments about an act's being performed intentionally correlated to their judgments about the blame and praise that they believed the Chairman in each condition deserved. This is to say that when subjects believed that the Chairman had harmed the environment intentionally, they also believed that he was highly blamable. When subjects believed that the Chairman had not helped the environment intentionally, they also believed that he was not praiseworthy.

These results were striking because it was made clear in both scenarios that the Chairman did not directly aim for the outcome (helping or harming the environment) in either condition, but ascriptions of intentional action and praiseworthiness and blameworthiness were assigned differently on the basis of the outcome. The asymmetry in the assessments of intentional action in the two cases, and in the assessments of praiseworthiness and blameworthiness, is what is known as the Knobe Effect. I will also refer to the asymmetrical judgments about praiseworthiness and blameworthiness as Praise-Blame Asymmetry, following a large body of literature that refers to this aspect of Knobe's results in particular.[4]

---

[3] Knobe 191.

[4] Susan Wolf refers to another type of asymmetry in our judgments about praiseworthiness and blameworthiness in her book *Freedom Within Reason*. She notes that we think that a person can be appropriately called praiseworthy for performing a right act even if this person's praiseworthy character is largely due to the luck of a "good up-bringing," but that a person might not be appropriately called blameworthy for performing a wrong act if this person's blameworthy character is largely due to the luck of a "bad up-bringing." Here, she seems to be citing an asymmetry in our ascriptions of who can be accurately named reason responsive (who can be said to act for reasons), and who are, therefore, potential candidates for praise and blame. While I do not discuss it in this paper, it is possible that there is some underlying phenomenon that can explain both types of asymmetry.

Knobe's initial assessment of the results was that people tend to be more willing to say that a side-effect was brought about intentionally when it is harmful than when the side-effect is helpful and that this attribution will be made regardless of whether or not the side-effect was actually brought about intentionally.

My Project

The first part of my project, Sections One and Two, is dedicated to the question of whether laypeople are justified in engaging in Praise-Blame Asymmetry, and, ultimately, whether, the Chairman in the HARM scenario, or any morally indifferent agent who is responsible for harm, is blameworthy. I intend to demonstrate that laypeople are not justified in engaging in Praise-Blame Asymmetry and that agents like the Chairman are not blameworthy based on the current empirically-based psychological model. To clarify, my project is to examine empirical models of the folk concepts that laypeople employ while making judgments about praiseworthiness and blameworthiness and to build a folk theory of moral worth using these criteria. I will then assess this folk theory of moral worth, examining the rational implications of its various components, and determine whether Praise-Blame Asymmetry is justified on this view. It turns out that, if the relevant criteria are applied consistently, the folk theory of moral worth fails to accommodate our desire to engage in Praise-Blame Asymmetry and gets one of the key cases, the case of the morally indifferent agent who performs a wrong act, wrong.

In Section One, I will introduce the cases of four Kantian-variety grocers to better explain Praise-Blame Asymmetry and to demonstrate why this phenomenon is interesting. The cases of the grocers are important because they depict idealized versions of the kinds of agents to whom laypeople apply asymmetrical judgments about praise and blame and can be used, in their simplest form, to demonstrate why Praise-Blame Asymmetry does not work. I plan on describing these grocers in depth because I will appeal to these examples repeatedly throughout the course of this paper. At the end of this section I pose the question that drives the rest of my paper: Why do we sometimes assess agents for praiseworthiness and blameworthiness based on their motivations and why do we sometimes not?

In Section Two, I am going to appeal to literature in psychology and experimental philosophy to describe an empirically-based folk theory of moral worth in order to get a

---

For more on the Asymmetry of the Reason View, see Susan Wolf, *Freedom Within Reason*, (New York: Oxford University Press, 1990) 79 – 81.

better idea of why laypeople make the judgments they do. To do this, I will examine how laypeople make judgments about three separate folk concepts: the evaluative status of acts, the criteria for intentional action, and an agent's candidacy for ascriptions of moral responsibility for performing an action. I am then going to examine this model to determine how to best explain the occurrence of Praise-Blame Asymmetry. I will finish this section by determining whether this model actually supports asymmetrical judgments about praiseworthiness and blameworthiness.

As it turns out, there appear to be inconsistencies in the application of the criteria relating to ascriptions of intentional action, specifically the criteria for having an intention to act. In other words, there are inconsistencies in the way that we appeal to the motivations that agent's act on when we make ascriptions about moral responsibility. Here, it is worth pointing out that the distinction between having an intention to act and acting intentionally will play an important role throughout this paper. Judgments about an agent's acting intentionally, as is shown in Knobe's experiments, are related to judgments about the moral responsibility of an agent. Judgments about an agent's having an intention to act are related to judgments about an agent's acting intentionally.

In Section Two, I will show that the layperson's criteria for judgments about an agent's having intentions to bring about certain outcomes keep them from making ascriptions of intentional action and, therefore, logically prohibit them from judgments that certain kinds of agents are morally responsible for their bad acts. If this is true, then laypeople cannot engage in Praise-Blame Asymmetry because the application of their criteria for praiseworthiness and blameworthiness does not accommodate the judgment that morally indifferent agents are blameworthy. This is the specific type of agent that is targeted by the application of Praise-Blame Asymmetry. The upshot of Section Two is that it seems like laypeople are unjustified in sometimes assessing agents for praiseworthiness and blameworthiness based on their motivations and sometimes not.

The second part of my project involves a critical analysis of the empirically-based folk theory of moral worth using philosophical perspectives. In Section Two, I determine that the layperson's criteria for ascriptions of intentional action logically prohibit them from engaging in Praise-Blame Asymmetry. In Section Three, I am going to start by attempting to resolve the inconsistencies in the application of the criteria for intentional actions, specifically in the application of the criteria for having an intention to act, by appealing to the intend/foresee distinction. Here, I am aiming at fixing the folk theory of moral worth. By clarifying the criteria for having an intention to bring about some outcome versus merely

foreseeing it, I hope to suggest that the inconsistencies in the application of Praise-Blame Asymmetry are simply associated with the vagueness commonly attributed to the intend/foresee distinction and that distinguishing these concepts will help us clearly and consistently apply the criteria for intentional action and, ultimately, Praise-Blame Asymmetry.

However, getting clearer on the distinction between intending versus foreseeing does not resolve the inconsistencies in the application of Praise-Blame Asymmetry, and, instead, it presents me with a difficulty. All of my attempts to resolve the inconsistencies in Praise-Blame Asymmetry by clarifying the intend/foresee distinction result in backwards Praise-Blame Asymmetry. Each of these attempted resolutions get the case of the morally indifferent agent who performs a right act wrong; they say that he is praiseworthy. This is in direct conflict with the results indicated in Knobe's experiment.

In Section Four, I adopt a new strategy and suggest that laypeople and the empirical experiments that are recording their judgments are neglecting important criteria for ascriptions of moral responsibility: reasons for acting. It is possible that laypeople are appealing to an agent's motivational reasons for acting, rather than their intentions, when they make asymmetrical judgments about praiseworthiness and blameworthiness. It is possible that laypeople are always appealing to an agent's motivations for acting when they assess his moral responsibility for an act – just not the motivations that I spent the first half of my paper exploring.

In this section, I explore the relationship between reasons and intentions and how the idea of acting for a reason is related to the idea of acting intentionally. I talk about what it means to have a reason and what it means to have reason all-things-considered. From here, I move on to talk about what it means to act for a reason, paying special attention to the views put forward by Nomy Arpaly. Her views play an important role in my paper from this point on. Once I get clear on each of these different ideas, I set out to reformulate the folk theory of moral worth in order to make it accommodate the appeal to reasons for acting. I note that this view looks a lot like Nomy Arpaly's own view of moral worth, so I move into the next two sections, Sections Five and Six, with an eye towards fleshing out Arpaly's more sophisticated theory and showing that theories of moral worth in which Praise-Blame Asymmetry occurs can be justified.

Throughout Sections Four, Five, and Six, however, there is a lingering shadow of the question that I originally posed. It quickly becomes apparent that an appeal to reasons for acting does not satisfy the question of why we sometimes appeal to the motivations of agents

when we make assessments of praiseworthiness and blameworthiness and sometimes do not, and cannot obviously justify the tendency. It seems that, in order to make Praise-Blame Asymmetry work; we have to appeal to normative versus motivational reasons to make claims about the blameworthiness of morally indifferent agents who perform wrong acts. In Section Seven, I spend some time questioning whether adopting this strategy to maintain Praise-Blame Asymmetry is justified on a view like Arpaly's, but I ultimately do not sure that my questions can be resolved. Arpaly's view does seem to successfully accommodate Praise-Blame Asymmetry without being held back by the problems and inconsistencies that lead me to move away from each of the other views I addressed within my paper. Her view seems to be a good place to look for resolving my questions, though I do not, ultimately, find my questions satisfied by my exploration of her view.

One

One might be wondering why I am questioning whether Praise-Blame Asymmetry is justified to begin with. After all, there is not anything obviously wrong with asymmetrical judgments about praise and blame. This question requires further clarification. So far, I have only reported that this phenomenon occurs; I have not really pointed to the reason that I think this phenomenon is worth investigating. In order to show why I think this phenomenon is worth investigating, I will present four different cases about Kantian-variety grocers. I will use these cases to show how Praise-Blame Asymmetry occurs in our judgments about praise and blame, and where I think the potential problem arises.

Before going through my four cases, however, I think it is a worthwhile endeavor to explain why I am choosing to use my own characters rather than Knobe's original cases from the HELP/HARM scenarios. This is to demonstrate that the examples I am using are still relevant to the question of Praise-Blame Asymmetry. In the introduction, I described Praise-Blame Asymmetry as an empirical phenomenon that occurs when laypeople make moral judgments about the morally relevant side-effects of an agent's act. Knobe uses the cases of two Chairmen who either help or harm the environment to draw his conclusions about Praise-Blame Asymmetry. In this section I use the examples of four grocers to explore the concept in question. This is because I frame questions about Praise-Blame Asymmetry a little bit differently than Knobe does, though I do not think that my way of talking about the concept is outside of the original parameters.

In Knobe's experiments, the participants were asked to make judgments about agents who perform acts with unintended side-effects, such as helping or harming the environment while trying to increase a company's profits. Throughout my paper, I will be talking about the intended and unintended harms and goods produced by agents when they act. There is a question as to whether it is appropriate to break down the problem in terms of side-effects without further exploration of what counts as a side-effect of an act. In both my account of Praise-Blame Asymmetry and Knobe's, we are ultimately talking about how we make judgments about agents whose acts have outcomes that they were not specifically aiming at. I think that this question is generalizable to more than just unintended side-effects of acts. I am choosing to use the cases of four grocers because I think it is useful to examine the

parallels of examples that have already been somewhat hashed out in philosophical literature.[5]

Praise-Blame Asymmetry: Four Cases

Imagine a grocer, one who is always honest with her customers, and one who does so because acting in accordance with morality is her primary motive. This grocer cares about profiting from her vending as much as the next grocer, but performing the right act is what motivates her to be honest with her customers. If being honest with her customers were not as profitable as cheating her customers, this *good-willed grocer* would still treat her customers with honesty and fairness where other kinds of grocers might not. Many of us are inclined to think that this good-willed grocer is morally praiseworthy for being honest with her customers; we think that she performs the right act with the intention of performing the right act, and that her decision to pursue this act is commendable.

Now, imagine a second grocer. She is a *malevolent merchant*, who lies to and cheats her customers for the sake of doing so. This malevolent merchant might not even particularly care about profits; she engages in vending activities just because she knows that being a grocer gives her the opportunity to steal from and hurt unsuspecting customers. Many of us are inclined to think that the malevolent merchant is blameworthy for cheating her customers. We might take her to be blameworthy because she is performing wrong acts precisely because they are wrong.

It is worth noting here that both of these grocers act with the goal of helping or harming their customers in mind. This is to say that, even though the acts of these two grocers differ in that they are right or wrong, we can see that they act on similar types of intentions. They both specifically try to bring about the positive or negative outcomes of their acts. The cases of the good-willed grocer and the malevolent merchant are different from the cases of the Chairmen in the HELP and HARM conditions in that the positive or negative outcomes (helping or harming the environment) were not specifically among the intended outcomes of the Chairmen's acts.

While many of the acts performed by moral agents are in accordance with morality, not all of them seem to commend themselves to praise in the way that the good-willed

[5] Two of the examples that I will be discussing, the good-willed grocer and the prudent grocer belong to Kant. The intuitions that I appeal to regarding their candidacy for praiseworthiness or blameworthiness are Kant's own, and I try to keep my discussion of them in line with his views. Further, my inspiration for the other two grocers I introduce, the malevolent merchant and the beguiling grocer, is fueled by Kant's grocer examples. See Immanuel Kant, *Grounding for the Metaphysics of Morals* (Indianapolis: Hackett Publishing Company, 1993).

grocer's act does. Kant suggests that we imagine a third grocer: one who is always honest with his customers, but who is only honest with his customers because it is prudent to do so.[6] If this *prudent grocer* were to cheat his customers, he would probably lose business; he is always honest with his customers because it is good for his profits. If cheating one's customers were not bad for business, this grocer might not care about fair pricing, but he cares about it for now because it is beneficial for him to do so. By this, I mean to say that the prudent grocer has no qualms about cheating his customers, and that he would do so if being honest with his customers put him at a financial disadvantage. It is simply the case that the prudent grocer currently thinks that the most profitable move is to treat his customers fairly.[7]

The prudent grocer clearly acts in accordance with morality, but, as Nomy Arpaly notes, we cannot help feeling that his performing the right act is somehow accidental.[8] For the prudent grocer, honesty is the means to some end, rather than the end itself. The fact that he is honest has more to do with his interest in profits than his interest in morality and, while we are happy that the prudent grocer is fair to his customers, we do not believe he is praiseworthy for his behavior. We think that the prudent grocer performed the right act, but we also think that his motivations were insufficient for ascribing praiseworthiness. After all, under slightly different conditions, the profit motive would have motivated the prudent grocer to cheat his customers rather than treat them fairly. I take the prudent grocer to be a type of morally indifferent agent. A morally indifferent agent is an agent who, on my view, acts without regard for the rightness or wrongness of the act. This is to say that a morally indifferent agent is an individual who does not specifically intend the positive or negative outcomes of his act. The prudent grocer is similar to the Chairman in the HELP condition.

When the prudent grocer is honest with his customers, he does not regard the rightness of being honest with his customers as a relevant motivation for acting. The prudent grocer knows that his act will result in his being honest with his customers, and he probably

---

[6] Kant 10.

[7] I must quickly distinguish a morally indifferent agent, one who does not take moral concerns as relevant to motivating him to act when he acts, from an agent who does think about moral concerns for acting when she acts but is not sufficiently motivated by these moral considerations to actually act on them. This would be a grocer who, for example, weighs moral considerations in favor of treating her customers fairly against relevant prudential considerations but ultimately takes prudential concerns to be the best for motivating her to act.

At the behavioral level, she might perform the same act as the prudent grocer. However, the prudent grocer never considered the evaluative or moral status of the act as a relevant reason for acting, whereas this other agent did. I will not give this second type of agent very much time in this paper, but it seems likely that laypeople would ascribe the same level of praiseworthiness or blameworthiness to this type of agent, based on her actions, as they would to a morally indifferent agent. In this paper, I focus on morally indifferent agents because the Chairmen in Knobe's experiments claim that they do not care at all about the environment, indicating that the relevant moral considerations for their acts are irrelevant to them. In other words, the Chairmen are indifferent to the relevant moral concerns, making them types of morally indifferent agents.

[8] Arpaly 71.

even knows that being honest with his customers is the right thing to do; he just does not take these facts into consideration when he chooses which act he is going to pursue. The relevant moral concerns in this situation simply do not exist on his cognitive horizons.

By contrast, the good-willed grocer has these relevant moral concerns on her cognitive horizons, and it is precisely these considerations that move her to treat her customers fairly. The prudent grocer is comparable to the good-willed grocer in that they both perform the right act, but he is different in that he does not care about morality when he acts, while the good-willed grocer does seem to care about morality. For Arpaly, this regard for morality, or this propensity for responding to conscious or unconscious moral considerations, when performing an act is key to an agent's being a potential candidate for praise or blame.[9]

So far, my accounts of the grocers have not been all that interesting. Good-willed grocers are individuals whom we consider praiseworthy and malevolent merchants are individuals whom we consider blameworthy. The prudent grocer we regard with a general attitude of neutrality. I would like to compare the case of the prudent grocer to the case of a fourth grocer: he is a *beguiling grocer*. The beguiling grocer is much like Kant's prudent grocer; he has a desire for profit, and he acts on prudential motives, but his act is ultimately the morally wrong act. This grocer desires profit, and, while pursuing profit, he cheats his customers.

The beguiling grocer is like the malevolent merchant in that they both perform the wrong act; they both cheat their customers, but the beguiling grocer is like the prudent grocer in that, when he acts, he does so without regard for moral considerations. This is to say, the beguiling grocer, like the prudent grocer, is a morally indifferent agent. When the beguiling grocer cheats his customers, he does not regard the rightness of being honest with his customers as a relevant motivation for acting, nor is he motivated to hurt his customers. The beguiling grocer knows that his act will result in his harming his customers, and he probably even knows that cheating his customers is the wrong thing to do; he just does not take these facts into consideration when he chooses which act he is going to pursue. The relevant moral concerns in this situation simply do not exist on his cognitive horizons. The fact that he cheats his customers has more to do with his interest in profits than his interest in morality. The beguiling grocer is similar to the Chairman in the HARM condition. The prudent grocer

[9] Arpaly 72.

and the beguiling grocer are equally responsive to moral reasons when they act in the sense that both are equally non-responsive to moral considerations when they act.

Moral Luck and the Profit Motive

One topic in philosophy that seems relevant to this discussion about the potential praiseworthiness and blameworthiness of agents like the prudent grocer and the beguiling grocer is *moral luck*.[10] Questions about moral luck occur when an agent is thought to be praiseworthy or blameworthy for an act in spite of the fact that some features of the act appear to be outside of the agent's control. Both the prudent grocer and the beguiling grocer act from the same set of motivations; they want to make profits. Their situations, however, differ, and the prudent grocer is able to acquire profit by being honest with his customers, while the beguiling grocer is only able to acquire profit by cheating them. It seems possible to suggest that the good or bad acts of the prudent grocer and the beguiling grocer are in some ways accidental. I noted earlier that Arpaly states as much about the good outcome of the prudent grocer's act. The profit motive itself does not assure that an agent who acts on it will commit good or bad acts. It appears to be the situations that these agents find themselves in that dictate what kinds of acts they carry out and their moral worth. If the scenarios were reversed, the two grocers would have performed reverse acts – making the beguiling grocer act fairly and the prudent grocer cheat.

In some ways, it is strange to say that it is okay to judge two agents differently when the only difference between them is something that is outside of one agent's control. Other things being equal, it seems that the only difference between the case of the prudent grocer and the case of the beguiling grocer is that acting on the profit motive in the first case results in a good act and that acting on the profit motive in the second case results in a bad act, and yet we are inclined to blame the beguiling grocer and not praise the prudent grocer.[11]

Kant is known for stating that morality is immune from luck. He states that the will alone, specifically the good will, has moral worth.[12] A good-willed agent, like the good-willed grocer, would be praiseworthy no matter what outcome she actually produced when she acted. Kant believes that it is not the utility of the outcomes of an agent's acts that create moral worth; it is the will with which the agent acts that has it. I think, then, that Kant would

---

[10] For more on the topic of Moral Luck, See Bernard Williams and Thomas Nagel, "Moral Luck," *Proceedings of the Aristotelian Society, Supplementary Volumes*, 50 (1976): 115 – 135; 137 – 151.
[11] This is what Nagel calls circumstantial luck (Williams & Nagel 140).
[12] Kant 7.

agree with me that the good-willed grocer and the malevolent merchant would be praiseworthy and blameworthy respectively because of the quality of the wills with which they act. Kant, however, believes that the will with which the prudent grocer acts is morally insignificant even though he produces the good or right outcome. He might also be committed to the idea that the will with which the beguiling grocer acts is morally insignificant even though he produces a bad outcome. If it is the will of the agent and not the outcomes of the agent's acts that define the moral worth of an agent's act when he acts, then Kant would probably have to regard the prudent grocer and the beguiling grocer similarly.

I think that we often want to agree with Kant. While we care about what kind of act an agent performed, right or wrong, when we assess her praiseworthiness or blameworthiness, we are more interested in the reasons for which the individual in question performed that act.[13] This is to say that we appeal to an agent's mental states when making assessments about that agent's praiseworthiness or blameworthiness when she performs some act. When we consider the good-willed grocer, the malevolent merchant, and the prudent grocer, we take a look at the relevant reasons for which they were actually acting when we make judgments about their praiseworthiness and their blameworthiness.

Another way of saying this is that we examine the intentions that the agents in question acted on when we assess them. We ask the question: What outcome was the agent pursuing when she acted? The malevolent merchant and the good-willed grocer do in fact intend to produce the wrongs or the goods that are the consequences of their acts, and we make claims about their blameworthiness or praiseworthiness only after assessing their mental states. The prudent grocer very clearly does not intend the good that comes about as a result of his being fair to his customers; he only intends to acquire profits. Thus, after assessing his mental states, we do not think that he is praiseworthy for his act.

However, when we make judgments about the blameworthiness of the beguiling grocer we are not attending to the reasons for which he actually acted. In other words, we do not appeal to the agent's actual intentions. The beguiling grocer does not intend to cause harm to his customers; he only intends to acquire profits. If we were to attend to the reasons for which the beguiling grocer was actually acting, we would have to regard him with the same moral attitude that we regard the prudent grocer with. After all, they acted on identical

[13] In this section, I use the term "reasons" vaguely. I mean to refer to the factors that motivated the agent to act. These could be reasons, desires, intentions, or any other relevant considerations that could motivate an agent to act and explain her doing so. Later I will use the term more specifically and will take the time to define the term there.

profit-driven motives, and neither agent intended the good or bad consequences of their actions.

If we adhered to some principle that stated that it is not okay to treat two agents differently when the morally relevant features of one agent's act is largely due to luck, then we would not get away with treating the beguiling grocer and the prudent grocer differently. If, like Kant, we believed that it was the will alone that has moral worth, then, once again, we would not get away with treating the beguiling grocer and the prudent grocer differently. We can also be sure that it is not the outcome of the act, its ultimate utility, or the beguiling grocer's unfortunate circumstances that get him into trouble because we know that it is not the rightness or the wrongness of the act alone that makes an agent praiseworthy or blameworthy. Otherwise we would make no distinction between the praiseworthiness of the good-willed grocer and the prudent grocer. Herein lays the asymmetry in our judgments about blameworthiness and praiseworthiness: we sometimes assess agents for praiseworthiness and blameworthiness based on their motivations and we sometimes do not.

Specifically, we appear to make a special exception for agents like the beguiling grocer and the Chairman from the HARM condition, or morally indifferent agents who perform wrong acts. The fact that this asymmetry occurs means that we adopt some alternative strategy for making the assessment that the beguiling grocer is blameworthy when he cheats his customers.

As a result, it is not clear what principle laypeople are using to make judgments about praiseworthiness and blameworthiness.

Thus far, I have been able to point out where Praise-Blame Asymmetry arises; it is the asymmetry in how we regard morally indifferent agents who perform right versus wrong acts. I have taken the time to explore how Praise-Blame Asymmetry arises by showing that it falls out of a tendency to sometimes morally judge agents based on their motivations and sometimes not. Ultimately, I am left questioning how laypeople justify this way of making judgments. They are not following any obvious principle. In the next section, I am going to construct a folk theory of moral worth from the relevant empirical literature and examine it to try to determine whether Praise-Blame Asymmetry can actually be justified through the lay perspective.

Two

In the previous section, I introduced and discussed the phenomenon called Praise-Blame Asymmetry. In order to acquire a better understanding of how Praise-Blame Asymmetry arises, we must understand how these asymmetrical judgments about praiseworthiness and blameworthiness fit into a larger folk theory of moral worth. In this section, I am going to appeal to literature in psychology to describe an empirically-based folk theory of moral worth. I am then going to examine this model to determine how to best explain the occurrence of Praise-Blame Asymmetry, and I will follow this by looking to see whether this model actually supports asymmetrical judgments about praiseworthiness and blameworthiness.

Thinking again about the HELP and HARM scenarios introduced by Knobe, we can note that there are three key features of these scenarios that are taken into account by laypeople. First, it seems that we attend to the evaluative status of the act in question. This is just to say that, in order to make moral judgments about these scenarios, we have to consider whether the act in question was right or wrong. Second, we assess whether the act in question was performed intentionally. While I will later offer a list of some necessary criteria for ascriptions of intentional action, I will quickly state here that an intentional action is an action that is performed for some purpose or an action that is performed with some goal in mind. Third, we make assessments about the moral responsibility of the agents in question and determine whether they are praiseworthy or blameworthy for the acts that they have performed. I think that judgments about these three concepts are interrelated but distinct. This is to say that I think that judgments about moral responsibility are separate from judgments about an act's being right or wrong and that judgments about these two concepts are distinct from judgments about an agent's performing an intentional action.[14] Meanwhile, I will continue my exploration of the folk theory of moral worth by first examining how laypeople make judgments about the evaluative status of acts.

Evaluative Status of Acts

The *evaluative status of an act* specifically refers to the *rightness* or *wrongness* of an act, or the extent to which an act is in accordance with morality or the extent to which an act

[14] By enumerating the features of the HELP/HARM scenarios that are attended to by laypeople, I am not currently trying to establish that this is the order in which these judgments are made. This is an important point, and I will get back to it later.

is in conflict with morality. References to rightness and wrongness throughout this paper are meant to be appealing to this distinction. In the experimental cases related to Praise-Blame Asymmetry, it seems that lay judgments about the rightness and wrongness of acts are based on whether the act in question actually resulted in a positive or negative outcome.[15] This is to say that acts are determined to be wrong when some harm occurs as a consequence of the act, and acts are determined to be right when some good occurs as a consequence of the act. This is probably because the assessments of the moral responsibility of agents in Knobe's experiments also tend to be outcome-oriented. This is to say that the participants in Knobe's experiments are asked to assess whether the Chairmen in the HELP/HARM scenarios are potentially praiseworthy or blameworthy by appealing to the positive or negative (helpful or hurtful) outcomes of the scenarios. Thus, the evaluative status of the act (its rightness or wrongness) and the evaluative status of the agent (the agent's level of moral responsibility for the act) seem to be assessed by appealing to the harms and the goods produced by the acts.[16]

One experiment, which shows how this is likely to be the case, employs the following two scenarios:[17]

> "DOCHARM: A patient is suffering from a potentially terminal disease. A doctor judges that death is very likely, but that an operation has some chance of saving the patient's life. The doctor also knows that the operation itself has a good chance of killing the patient. The doctor operates and this kills the patient."
>
> DOCHELP: A patient is suffering from a potentially terminal disease. A doctor judges that death is very likely, but that an operation has some chance of saving the patient's life. The doctor also knows that the operation itself has a good chance of killing the patient. The doctor operates and this saves the patient.

For both scenarios, the participants were asked to determine whether the act was good, bad, or neither. In the DOCHARM scenario, the participants largely decided that the doctor's actions were bad. In the DOCHELP scenario, the participants largely decided that doctor's actions were good. While the judgment that the doctor's actions in the DOCHELP scenario seems pretty straightforward, the assessment that the doctor's actions were bad in the

---

[15] Knobe 193.

[16] The *evaluative status of an act* is distinct from the *evaluative status of an agent*. The evaluative status of an act is the extent to which an act is right or wrong, and the evaluative status of an agent is the extent to which an agent is praiseworthy or blameworthy for some right or wrong act she performed. Here I am saying that both rightness and wrongness and praiseworthiness and blameworthiness are assessed by appealing to the helpfulness or hurtfulness of an act, but that appeals to the outcome of an act are sufficient for judgments about an act's being right or wrong but not for an agent's being praiseworthy or blameworthy for bringing about that act.

[17] J. Wright and J. Bengson, "Asymmetries in Judgments of Responsibility and Intentional Action," *Mind and Language*, 24.1 (2009): 24 – 50.

DOCHARM scenario may not be. This judgment was made even though it is clear that the doctor was trying to save the life of the patient in question. This means that participants think that performing acts with negative outcomes are bad or wrong regardless of what the agent in question was actually trying to achieve. I will later show that when an agent performs bad acts or wrong acts, she becomes a potential candidate for blame.

Here I will quickly note that the authors' use of the word "bad" throughout their experiments generated assessments of moral responsibility in the same way that the moral use of the word "wrong" does. This is to say that there was a relationship between the judgment that some act was bad and the judgment that some agent was responsible and potentially blameworthy for bringing about that bad act. The same was true of assessments about good acts and assessments of praiseworthiness. While it is not obvious that the judgment that an act was bad in the DOCHARM scenario resulted in the judgment that the agent was potentially blameworthy for his act, this connection is more clear in other experiments by Wright and Bengson and in experiments by other researchers. One factor that appears to mitigate the assessment that an agent is blameworthy for bringing about a bad or harmful or wrong act (these terms become pretty synonymous in these experiments) is whether that act was performed intentionally. I will get back to this issue later.

Fiery Cushman examined the roles of desire, belief, causation and consequences as determinants in our moral judgments about the wrongness and permissibility of acts as well as the punishability and blameworthiness of agents who perform those acts. The results of the experiment indicated that most people appeal to "culpable mental states," rather than negative outcomes, when making moral judgments about the wrongness of acts. This means that there is at least one experiment with results that contradict this understanding of how assessments of the evaluative status of acts are made, and the author of the experiment himself admits that his results are in direct conflict with the larger body of work associated with this topic.[18]

Culpable mental states include beliefs, desires, and intentions. It is important to note that Cushman's participants divided "intentions" into beliefs and desires, meaning that his participants believed that intentions are derived from beliefs and desires. Thus, on this view, individuals with culpable intentions are individuals with morally problematic beliefs and desires when they act.

Culpable mental states are generally possessed by two types of agents. The first type of agent is a person who has a belief that some act A will cause harm to someone and a desire

[18] Fiery Cushman, "Crime and punishment: Distinguishing the roles of causal and intentional analysis in moral judgment," *Cognition*, 180 (2008): 353 – 380.

to cause harm to someone when she performs the act. This agent is much like the malevolent merchant I described previously. The second type of agent is a person who has a belief that some act A will cause harm to someone but who has no similar malicious desire to cause harm when she performs the act. This agent is much like the beguiling grocer that I described previously. The first agent is one who acts from the motive of maliciousness and the second agent is one who is taken to be acting out of indifference to morality.

Cushman's results demonstrated that our judgments about the wrongness of an act are more closely associated with an agent's having the belief that an act will cause harm to someone than with an agent's having the desire for an act to cause harm to someone. Hence, the fact that we determine malicious individuals and morally indifferent individuals to have performed equally wrong acts even though one, the morally indifferent individual, lacks the desire that her act causes harm.

Imagine, once again, the malevolent merchant. The malevolent merchant has the belief that her act (cheating her customers) will count as an act that will harm her customers. She also has a desire to harm her customers. Similarly, the beguiling grocer, as we shall recall, has the belief that his act (cheating his customers) will count as an act that will harm his customers. He does not however have the desire to cause harm to his customers when he acts (his only desire is profit). The desires of an agent cannot reliably predict how we would assess the wrongness of an act given that we would make the same judgment about the wrongness of both acts even though one agent has the desire to cause harm and the other agent does not. This means that we appeal to an agent's beliefs about the potential harm-causing nature of an act when we make judgments about the wrongness of the act. As Cushman demonstrates, it is the agent's beliefs about the fact that an act will cause harm that reliably predicts our actual moral judgments about wrongness.

While it does matter how assessments of the evaluative status of acts are made by laypeople for a complete folk theory of moral worth, I am more interested in determining how the evaluative status of acts is connected to ascriptions of praiseworthiness or blameworthiness. For the purposes of this paper, I will rely on the view that lay judgments about the rightness and wrongness of acts are based on whether the act in question actually resulted in a positive or negative outcome, and stick to the results indicated by the larger

body of empirical literature. It is simply worth noting that there is some conflicting empirical evidence regarding how these judgments are made. [19]

Quickly, however, I would like to suggest that Cushman's understanding of how assessments of the evaluative status of acts are made does not seem to work as a defensible philosophical position.[20] This is because beliefs, just like desires, cannot reliably predict our judgments about an act's being right or wrong. To show how beliefs about an act's being harmful do not predict our judgment that an agent's act is actually wrong when she acts, I will consider the case of Mark Twain's Huckleberry Finn.[21]

Huck Finn is a boy who has internalized the values of his era, and among these is the belief that slaves are property. Huckleberry runs away from home at the same time as a slave, Jim, is running away from captivity. Huck finds himself in a moral predicament because he believes that he ought to turn Jim in to the authorities, and he believes that failing to turn Jim in causes a serious harm to Jim's owner. Huck struggles with the idea of doing the act that he overtly believes is the right act (turning Jim in). After spending some time with Jim, Huckleberry is reluctant to turn Jim in, determines that he is unable to do so, and eventually resigns himself to performing the act that he believes is harmful: helping the slave Jim escape.

Even though Huck Finn believes that failing to turn Jim in is a harmful act, we do not think that he performs the wrong act when he helps Jim escape. To the contrary, we think that Huck clearly performs the right act even if we do not think, as Nomy Arpaly does, that he is probably a potential candidate for praise.[22] The mere fact that Huck Finn believes that he is performing a harmful act is not enough for us to say that he performed a wrong act. Thus, culpable mental states when acting do not seem to be sufficient criteria for ascribing rightness or wrongness to an act. Given the philosophical indefensibility of this position and the fact that this research is not supported by the larger body of literature, I will take for granted the view that wrongness, on the lay perspective, is largely dictated by an act's resulting in a harm and that rightness is largely dictated by an act's resulting in a good.

Intentional Action and Moral Responsibility

---

[19] I am not trying to make the claim here that this is the only way that judgments about rightness and wrongness are made. This just seems to be the strategy employed by participants who are making judgments about the evaluative status of acts in these experimental scenarios.

[20] Hence, my decision to stick with the larger body of literature. His research does not follow most results, and what he proposes does not seem philosophically defensible

[21] Mark Twain, *Adventures of Huckleberry Finn* (USA: Quill Pen Classics, 2008).

[22] Arpaly 76.

Now that I have talked a little bit about what makes an act right or wrong according to laypeople, I will examine the folk criteria for intentional action. As I noted at the beginning of this section, judgments about intentional action appear to be one of the three key features that laypeople attend to when making attributions of moral responsibility to agents who bring about morally relevant outcomes.[23] Wright and Bengson suggest that judgments about intentional action connect the evaluative status of the act in question to the responsibility of the agent who performed it.[24] This is to say that we assess whether some act was right or wrong and whether it was performed intentionally in order to determine whether the agent who performed the act is praiseworthy or blameworthy.

In *The Folk Concept of Intentionality*, Malle and Knobe determine necessary criteria for folk ascriptions of intentional action through a series of experiments.[25] These criteria include:

1. A desire for some outcome.
2. A belief that some act will result in that outcome.
3. An intention to perform the act.
4. Awareness of performing the act.
5. Evidence of skill to successfully complete the act.[26]

It is important to note that laypeople make a distinction between having an intention to act and acting intentionally. Namely, they believe that intentions to perform an act are necessary criteria for one's having acted intentionally. It is also worth noting that the results of these experiments indicate that laypeople believe that intentions are derived from beliefs and desires. This, at least, is consistent with the understanding of intentions demonstrated in Cushman's article. If an agent fails to have an outcome-oriented desire or a belief that some act will result in that outcome, then the agent in question does not have an intention to

[23] It is possibly confusing that I am now making the move to talk about intentional action, when my objection to Cushman was that an agent's mental states, her beliefs, desires, and, ultimately, her intentions do not matter. Reexamining my objection to Cushman, I stated that intentions and their component parts, specifically beliefs, are not relevant to our judgments that certain acts are right or wrong (recall that Cushman's results indicated that laypeople break intentions down into beliefs and desires). I did not say that they were not relevant to our judgments about the moral responsibility of agents who perform right or wrong acts.

[24] Wright & Bengson 25.

[25] B. F. Malle and J. Knobe, J, "The Folk Concept of Intentionality," *Journal of Experimental Social Psychology*, 33 (1997): 101 - 121.

[26] The authors of the article do not take this to be an exhaustive list of necessary criteria for having performed an intentional action (Malle & Knobe PAGE NUMBER HERE).

perform a relevant act, and if the agent lacks any of these three criteria, then he does not act intentionally.[27]

It might be argued that one does not need a desire for some outcome in order to have an intention to perform some act. One might intend to clean one's room while having no desire to tackle the overwhelming mess. However, it seems strange to say that the agent in question has no desires regarding the cleaning of his room. He just might have a desire to be in a cleaner space, a desire to not be embarrassed when his mates stop by for a visit, or a desire not to be scolded by his mum when she walks into the room. Any of these desires, coupled with the belief that cleaning one's room will achieve that outcome, should be sufficient for developing an intention according to Malle and Knobe's folk criteria for intentional action.

It is worth noting that this conception of intentional action is contended by some philosophers. Desires are often not taken to be necessary criteria for ascriptions of intentional action, or, at least, specific outcome-oriented desires are not taken to be necessary criteria for ascriptions of intentional action. While many philosophers do seem to agree that there is a distinction between having an intention and acting intentionally, some philosophers, like Bratman, do not think that having a specific outcome-oriented anything (*e.g.* desires, or intentions) is a necessary criterion for ascriptions of intentional action. By this I mean that he thinks that it is not a necessary criterion for intentional action that an agent either desires some specific outcome or aims at it when she acts. In the next section of this paper, I will talk about how philosophers try to resolve the issues I bring up associated with the folk concept of intentional action and with the folk theory of morality in general. Meanwhile, I will continue my examination of the folk theory of morality.

Folk Theory of Moral Worth: A Uni-directional Model

So far, I have taken the time to explain how laypeople make judgments about the evaluative status of acts and how laypeople make judgments about intentional action. I have noted that judgments about intentional actions seem to make the connection between the evaluative status of an act and an agent's moral responsibility, his praiseworthiness or blameworthiness, for bringing about that act. With this, I have my first iteration of the folk theory of moral worth:

[27] With regards to the folk criteria for intentional action, I will primarily discuss beliefs, desires, and intentions throughout the rest of this paper. It is the problems surrounding these criteria that will fuel my discussion about intentional action from this point on.

Right/Wrong Act + Intentional Action = Praiseworthiness/Blameworthiness

I will start my examination of this first iteration of the folk theory by considering it as a *uni-directional model*. On a uni-directional model, praiseworthiness and blameworthiness are ascribed to an agent when we have first determined that she has performed a right or wrong act and then determined that she has performed that act intentionally. If the agent in question has either failed to perform a relevant right or wrong act (an act with a positive or negative outcome) or failed to have performed that act intentionally, then that agent is not a potential candidate for praiseworthiness or blameworthiness. To assess this model, I will appeal to the cases of the four grocers. Since the only relevant differences between the grocers, besides the rightness or wrongness of their acts, are their beliefs and desires, I will assume that all other necessary criteria apply to the grocers, and I will assess whether their actions were performed intentionally on the basis of their beliefs and desires alone.[28]

In the case of the good-willed grocer, she has a belief that her act (being fair to her customers) will do some good for her customers, and she has a desire to do such goods for her customers when she acts. As a result, the layperson would determine that the good-willed grocer is praiseworthy for being fair to her customers because she performed an act with a positive outcome and she performed that act intentionally. Similarly, the malevolent merchant has a belief that her act (cheating her customers) will do some harm to her customers, and she has a desire to do such harms to her customers when she acts. As a result, the layperson would determine that the malevolent merchant is blameworthy for cheating her customers because she performed an act with a negative outcome, a wrong act, and she performed that act intentionally.

On the other hand, we have the prudent grocer who has a belief that his act (being fair to his customers) will do some good for his customers, but he has no such desire to do goods for his customers when he acts; he only desires to acquire profit. As a result, we would determine that the prudent grocer is not praiseworthy for being fair to his customers because, although he performed an act with a positive outcome, he cannot be said to have performed that act intentionally. So far, the uni-directional model is in line with empirically demonstrated folk ascriptions of praiseworthiness and blameworthiness. However, this model fails to allow for Praise-Blame Asymmetry because if relevant beliefs and desires are necessary criteria for ascriptions of intentional action, then we cannot say that the beguiling grocer harms his customers intentionally.

[28] Here I am referring the necessary criteria for intentional action put forward by Malle and Knobe.

The beguiling grocer has a belief that his act (cheating his customers) will do some harm to his customers, but he has no such desire to harm his customers when he acts; he only desires to acquire profit. As a result, we would determine that the beguiling grocer is *not* blameworthy for cheating his customers because, although he performed an act with a negative outcome, he cannot be said to have performed that act intentionally; he lacks the necessary desire to cause harm to his customers.

If praiseworthiness and blameworthiness are ascribed to an agent when we have first determined that she has performed a right or wrong act and then determined that she has performed that act intentionally, then we consistently get the case of the beguiling grocer wrong and Praise-Blame Asymmetry is unjustified on this view. This is to say that when the beguiling grocer fails to have a desire to cause harm, he does not meet the criteria for having an intention to cause harm, and he thereby fails to meet the criteria for having brought about that harm intentionally. If his bringing about a harm does not meet the criteria for intentional action, then he cannot be considered blameworthy for having brought it about. If this is how we typically make judgments about the blameworthiness of agents like the beguiling grocer, then we consistently get the case of the beguiling grocer wrong, and Praise-Blame Asymmetry seems unjustified. What is puzzling about this is that the layperson's assessment that the beguiling grocer (and similarly the Chairman from Knobe's HARM scenario) is blameworthy is very obviously wrong according to this model. It is strange that we would so consistently fail to make an accurate judgment about a case that is so straightforward. The beguiling grocer simply fails to meet the necessary criteria for ascriptions of intentional action, and thus fails to meet the criteria for ascriptions of blameworthiness.

Folk Theory of Moral Worth: A Bi-directional Model

To explain this consistent failure on the part of laypeople at getting the case of the beguiling grocer (and morally indifferent agents in general) wrong, Wright and Bengson suggest a *bi-directional* folk theory of moral worth.[29] This is to say that sometimes laypeople ascribe praiseworthiness and blameworthiness to agents for performing right or wrong acts only when they have performed those acts intentionally and that laypeople also sometimes judge that a right or wrong acts was done intentionally once they have decided that the agents in question are praiseworthy or blameworthy.

[29] Wright & Bengson 26; for more on bi-directional folk models of moral worth, see also Thomas Nadelhofer, "Blame, Badness, and Intentional Action: A Reply to Knobe and Mendlow," *Journal of Theoretical and Philosophical Psychology*, 24.2 (2004): 259 – 268.

By appealing to the bi-directional model, we can see that laypeople are still wrong when they assign blameworthiness and intentional actions to morally indifferent agents who perform wrong acts, but they are less obviously wrong. In first ascribing blame to the agent and then making judgments about intentional actions, they basically commit a logical fallacy similar to affirming the consequent. Wright and Bengson note that this is something like what Gilbert Harman refers to as "inference to the best total explanation."[30] Mark Alicke refers to this as the "blame validation assumption" in his culpable control model.[31]

According to Alicke, laypeople use a harm's occurring as an indicator that somebody is potentially blameworthy. The beguiling grocer and the Chairman elicit "strong negative spontaneous reactions" for their harmful actions. Once these reactions occur, laypeople assess the relevant information about the situation in a way that justifies their desire to blame the agent, and, as a result, they make *post hoc* ascriptions of intentional action in order to make their judgments about the agent's blameworthiness seem coherent. These judgments are not dependent upon logical processes; rather they are dependent upon affective biases. This is probably not a complete explanation of why laypeople so consistently get the case of the beguiling grocer wrong, but it does successfully demonstrate why laypeople do not easily notice that they are getting the case wrong. It is worth restating here that evidence supports the idea that judgments about an agent's acting intentionally can be made on the basis of both ascriptions of praiseworthiness and blameworthiness. While a number of researchers, like Alicke, talk explicitly about the bi-directionality of ascriptions of intentional action based on immediate assessments of blameworthiness, some experiments demonstrate that the same occurs when immediate assessments of praiseworthiness are made.[32]

Further Discussion

I will quickly make note of a concern the reader may have had while working through this section. It is possible to think that I have been trying to make the argument that laypeople cannot accurately ascribe praiseworthiness and blameworthiness an agent because they do not have access to the agent's internal mental states. This is very much not what I am trying to argue. I am merely making the claim that laypeople can be mistaken about the mental states that they are ascribing to a particular agent. In the case of Praise-Blame Asymmetry,

[30] See Gilbert H. Harman, "Inference to the Best Explanation," *Philosophical Review*, 74.1 (1965): 88 – 95.
[31] See M. Alicke, "Culpable Causation," *Journal of Personality and Social Psychology*, 63 (1992): 368 – 378; see also M. Alicke, "Blaming Badly," *Journal of Cognition and Culture*, 8 (2008): 179 – 186; M. Alicke, "Culpable Control and the Psychology of Blame," *Psychological Bulletin*, 126 (2000): 556 – 574.
[32] Wright & Bengson 26.

laypeople consistently ascribe the wrong mental states to morally indifferent agents who perform wrong acts. They determine that these agents have a desire to cause harm and that they are therefore acting intentionally when they perform wrong acts, even though this is clearly false.

There are some problems with the current folk theory of moral worth. One problem seems to have to do with a distinction between bringing about an outcome intentionally and intentionally performing some act. In Knobe's original experiment, he asked his participants to determine whether the Chairmen brought about the positive or negative outcomes of their acts intentionally. In the HELP scenario, the Chairman is determined to have not brought about the positive outcome intentionally. In the HARM scenario, the Chairman is determined to have brought about the negative outcome intentionally. While we may disagree about which outcomes were brought about intentionally, it is strange of us to say, in either case, that the acts in question were not performed intentionally.

Surely the Chairman in the HELP condition intentionally performed whatever act it was that brought about the positive outcome of helping the environment, even if he did not intentionally bring about that outcome. Similarly, it would be strange to say that the prudent grocer did not intentionally treat his customers fairly, even if he did not intentionally do them some good. The key to solving this problem seems to be associated with which desires or intentions we attend to when we make attributions of intentional action. Because all of the grocers and the Chairmen have some relevant desire associated with performing their acts, it would be fair to say that each of them performed their acts intentionally.

The good-willed grocer and the malevolent merchant have desires specifically associated with the positive or negative outcomes of their acts. Thus, they both intentionally performed their respective acts and intentionally brought about their respective outcomes. The prudent grocer, the beguiling grocer, and both of the Chairmen all act from the profit motive, meaning they all have desires for some outcome (the acquisition of profit) and beliefs that their acts will achieve that outcome, Thus, each agent does, in fact, act intentionally. So, the question of whether some positive or negative outcome was brought about intentionally really must be a question of whether these agents have relevant desires for these outcomes when they perform the act.

Methodologically speaking, this seems to be an issue with the above experiments. Knobe noted in his early work that desires appear to be largely outcome-oriented and that

intentions appear to be act-oriented.[33] This means that when he focused on whether and outcome was brought about intentionally in later experiments, he should have been aware that he was asking questions about desires and not intentions or even intentional action.[34] This means that the language that Knobe uses in his experiments is unclear and that it may be causing confusion on the part of his participants. Laypeople very obviously want to say that agents like the beguiling grocer and the Chairman in the HARM scenario are blameworthy when they act, as in the case of Praise-Blame Asymmetry, even though the current folk model of moral worth does not seem to successfully allow for it. This is either because the model that laypeople employ is actually philosophically indefensible or because researchers have simply not yet fully explained how laypeople make judgments of praiseworthiness and blameworthiness.

In this section, I have laid out and examined an empirically-based folk theory of moral worth. I have determined that the folk theory of moral worth, as it stands, does not support Praise-Blame Asymmetry. In the next section, I will look at some philosophical discussions on the concepts related to the folk theory, and I will see if the philosophical literature can help resolve the inconsistencies that render the folk theory incapable of supporting Praise-Blame Asymmetry.

[33]Malle & Knobe 48.

[34] Refer back to the questions that Knobe asks about the HELP and HARM scenarios. He asks his participants to make judgments about whether the Chairmen sought to bring about particular outcomes rather than particular acts (Knobe 191). If he makes the claim that desires are outcome-oriented and intentions are act oriented, he might just be asking questions about desires that are masked by questions about intentions.

Three

In the previous section I laid out and examined the criteria for an empirically-based folk theory of moral worth. As I have shown, the current folk theory of moral worth does not successfully account for our asymmetrical judgments in praise and blame. There is a straightforward inconsistency in the way that laypeople apply the criteria for intentional action in order to achieve Praise-Blame Asymmetry. While psychologists and other empirical researchers appear to be able to explain the phenomenon, they have not been able to describe a folk theory of moral worth in which Praise-Blame Asymmetry is philosophically defensible. In spite of the fact that it seems like actual lay judgments about praiseworthiness and blameworthiness are philosophically indefensible, laypeople seem committed to continue making these judgments. This indicates that laypeople believe that these judgments are correct. Thus we come to the role of the ethicist.

I noted at the very beginning of this paper that moral philosophers tend to examine actual judgments and intuitions about praise and blame in order to develop their theories of moral worth. Further, it is one of the roles of philosophers to demonstrate how lay judgments and intuitions about praise and blame are justified. Thus, we can look to philosophy to try to come up with a philosophically defensible theory of moral worth that successfully accommodates Praise-Blame Asymmetry, an actual phenomenon that occurs when laypeople make moral judgments.

We can start this exploration of the problem of Praise-Blame Asymmetry by looking to the philosophical literature in order to discover where some of the tension lies in the concepts related to Praise-Blame Asymmetry. Since I determined that the inconsistencies in the folk theory of moral worth occur in the application of the criteria for intentional action, specifically the criteria for having an intention, we can look to some of the complexities and problems with this concept and how philosophers have dealt with them as a way of resolving the inconsistencies related to Praise-Blame Asymmetry.

One issue in the literature is the distinction between outcomes that are intended and outcomes that are merely foreseen. This distinction is related to questions of whether outcomes are brought about intentionally or as side-effects of an act, how we tell the two apart, and how we make judgments about the moral responsibility of an agent based on how we categorize the outcome of an act. It is possible that laypeople are misapplying the criteria for intentional action because the distinction between intended versus foreseen outcomes is

vague. I noted this vagueness in Section Two as an obvious place to start criticizing the folk theory of moral worth. As we saw in the previous section, laypeople believe that agents like the prudent grocer do not act so as to bring about the positive outcomes of their acts intentionally, yet it seems strange to say that the prudent grocer did not act intentionally when he was fair to his customers because we do think that he was intentionally acting so as to acquire profit. It is possible that laypeople are making such odd claims about the prudent grocer's behavior because they are failing to be clear on the intend/foresee distinction.

In this section, I will talk about this distinction and focus on intentions more closely because it seems like the problems with the folk theory of moral worth and, more specifically, intentional action could be tied up in this concept. I will also take the time to explore the idea that no such distinction is necessary and that the problem with our ascriptions of intentional action lies in the desire criterion presented by Malle and Knobe. First, I will explore how the concept of intention, and thereby intentional action, can be defined via the intend/foresee distinction, and then I will look at some specific issues that philosophers would take up against the folk concept of intention. As it turns out, no matter how we define or apply our concepts of intention, we get a view that is inconsistent with Praise-Blame Asymmetry.

Intending versus Merely Foreseeing

Sometimes we want to say that an outcome is *merely foreseen* rather than *intended*. This is to say that sometimes an agent can know that her act will have some outcome without saying that her aim was to bring about that outcome. An outcome is foreseen when an agent performs some act while she has a relevant belief about the outcome of the act. By this I mean that if the outcome of an act is to be categorized as foreseen then the agent must have some knowledge, some relevant belief that the act in question will bring about some specific outcome for which she is causally responsible. If she lacks such a belief, if she does not know what she is doing when she is doing it, we would say that she has brought about that outcome accidentally.

Suppose, for example, that there is a big red button on the wall, and that there is an agent, Matilda, who enjoys pressing such buttons. Matilda walks into the room, sees the big red button, and she presses it – not really thinking about what purpose the big red button serves. As it turns out, the big red button is the On/Off switch for a feeding mechanism to hundreds of tiny bunnies in a lab, and that, when she presses the big red button, she turns the

feeding mechanism off and leaves the tiny bunnies to starve. Now, she had no relevant beliefs about the potential outcome of pressing the big red button. As far as she could tell, it served no purpose at all. Thus, it would be fair to say that she did not foresee the deaths of the hundreds of tiny bunnies when she pressed the big red button (though we might think that she should have foreseen something, and this is a different issue). So it seems that having a relevant belief about the outcome of an act is a necessary, and most likely a sufficient, criterion for an outcome's being foreseen.

We could suppose that Anastasia, Matilda's friend, does have a belief about tiny bunnies starving to death if she happens to push the big red button. However, we will also suppose that there is a malfunction with the wiring in the feeding mechanism that is causing the insulation to heat up and catch fire. Anastasia knows that if she does not push the big red button in order to stop the feeding mechanism, the whole lab will burn down, killing the tiny bunnies and everybody else as well. In this case, we would want to say that the starving of the tiny bunnies is a foreseen consequence of Anastasia's pressing the big red button; she has the relevant belief, and she knows what will happen. However else we would want to characterize the outcome of Anastasia's act, we know that Anastasia knew that pressing the big red button would result in hundreds of tiny bunnies starving to death.

Now that we know what it means for an outcome to be foreseen, we need to take a little time to discuss what it means for an outcome to be intended. It seems pretty clear that an outcome's being foreseen is a necessary criterion for an outcome's being intended by an agent when he acts. This is to say that an outcome can only be intended when an agent performs some act while he has a relevant belief about the outcome of the act.[35] However, beliefs about the outcomes of acts are not sufficient for an outcome's being intended. Intentions require something else. Intentions involve our motivations. Which motivations we appeal to while ascribing intentions to an agent is a controversial topic, so for now I will say that intentions require some sort of pro-attitude towards a particular act, and they are, as Michael Bratman would put it, related to our aims and goals.[36]

Returning to the case of Anastasia, we would not want to say that the starving of the tiny bunnies is intended because starving the bunnies is not her aim nor does she seem motivated to starve the bunnies. Her goal is to avoid the creation of a massively destructive fire in the lab. We might even praise her, commend her for preventing the destruction, in

[35] G. E. M Anscombe, *Intention* (USA: Harvard University Press, 2000).
[36] For more on the concept of pro-attitudes, see Donald Davidson, "Problems in the Explanation of Action," *Problems of Rationality* (Oxford: Oxford University Press, 2004) 101 – 117; for Bratman's views on intention, see Michael Bratman, "Two Faces of Intention," *The Philosophical Review*, 93.3 (1984): 375 – 405.

spite of the fact that she also caused hundreds of tiny bunnies to die. If, on the other hand, Janet, a friend of Matilda and Anastasia, believed that pressing the big red button would result in the deaths of the hundreds of tiny bunnies, and if killing the bunnies was somehow among her goals – the bunnies, we can say, were all infected with a highly communicable disease and killing them was the only way to prevent its transmission from bunny to human – then we would say that she intended to starve the tiny bunnies when she pressed the big red button. The death of the bunnies was her aim.

An interesting point is that we probably think that Janet is more morally responsible for starving the bunnies than Anastasia. This is because we think that agents are more morally responsible for outcomes that they actively pursue or intend to bring about than those which are the side-effects of their acts.[37] An upshot of this phenomenon is that we tend to think that harms that are intended are worse than harms that are merely foreseen consequences of acts. The same is true of goods that are brought about intentionally versus goods that are merely foreseen consequences of acts. Recall the good-willed grocer and the prudent grocer. We think that the good-willed grocer is praiseworthy where we believe that she intended the good outcomes of her act. We think that the prudent grocer is not praiseworthy where we believe that he did not intend the good consequences of his act – we think that the goods that the prudent grocer brings about are merely foreseen outcomes of his act. The fact that we make the intend/foresee distinction seems to entail that motivations matter to judgments about moral responsibility when we perform morally relevant acts.

Intending, Foreseeing, and the Doctrine of Double Effect

The intend/foresee distinction is featured prominently in the Doctrine of Double Effect [DDE], which has been discussed by Philippa Foot in the benchmark piece "The Problem of Abortion and the Doctrine of Double Effect".[38] The DDE states that it is permissible for an agent to perform an act with an outcome that is harmful just in case the harm that results from the act is a negative foreseen side-effect of the act and not the aim of the act itself. It is also required that the negative foreseen side-effect of the act is less than or at least neutral with regards to the good produced by the act. By this I mean to say that we sometimes think it is okay to perform an act that we know will result in a harm so long as the

[37] Yohsuke Ohtsubo, "Perceived Intentionality Intensifies Blameworthiness of Negative Behaviors: Blame-Praise Asymmetry in Intensification Effect," *Japanese Psychological Research*, 49.2 (2007): 100 – 110.
[38] Philippa Foot, "The Problem of Abortion and the Doctrine of Double Effect," *Virtues and Vices, and Other Essays in Moral Philosophy* (Berkeley: University of California Press, 1978).

purpose of our acting is to bring about some good and the harm that is caused by our action is less than the good we bring about.

The DDE has a place in our daily moral interactions. Suppose, for example, that Mordred, from Arthurian Legend, found out that his best friend's wife was cheating on him. He approaches King Arthur, knowing that what he has to say will cause him harm but believing that the truth will result in the most good. He tells Arthur that Guinevere and Lancelot are hiding the truth from him, and, sure enough, Arthur feels pain as a result of hearing this fact. We want to allow the Mordreds of the world to be honest and to share truths in spite of the fact that being honest is sometimes hurtful. So, according to the DDE, it is permissible for Mordred to cause Arthur pain in order to be honest with him, that pain being a negative foreseeable side-effect of his bringing about a good – giving important information to Arthur.[39]

Vagueness in the Intend/Foresee Distinction

I do not know if it is apparent yet, but there is a problem with the intend/foresee distinction. What distinguishes the outcome of an act as intended versus merely foreseen is not clearly defined. Any potentially intended outcome is also a foreseen outcome, and it seems like there is nothing to stop an agent from categorizing the outcome of an act as a merely foreseen outcome rather than an intended one.

Looking back at the case of Mordred and King Arthur, anybody who is familiar with Malory's version of the story knows that Mordred told King Arthur about Lancelot and Guinevere not because he believed it was the right thing to do (in spite of the fact that it may have been), but because he believed that it would hurt King Arthur, and he wanted to hurt King Arthur. Sharing this information was a direct cause of King Arthur's death and the disintegration of his kingdom. That seems pretty bad. Mordred acted from purely malicious motivations, as fictional characters often do. Suddenly, the harm that Mordred causes King Arthur by sharing information with him seems much less okay. The cover that he uses, that he is merely doing the right thing, does not seem to be enough for us to think that his behavior is permissible. We do not want to let him say that his hurtfulness is a mere side-effect of his actions because we know that his aim is to hurt Arthur. So, we are left with the

[39] See Sir Thomas Malory, *Le Morte d'Arthur* (Digireads.com Publisher, 2009) for more on the story of Arthur and Mordred, especially Book VIII: "The Death of Arthur."

question of how we are meant to draw out the distinction between a merely foreseen negative outcome and an intended one.

Bennet argues against the use of the intend/foresee distinction in moral discourse precisely because it is too vague.[40] It seems like any agent could use the intend/foresee distinction to justify her harmful act if framed her intentions in the right way. If Mordred wants to avoid punishment for his genuinely hurtful behavior, he need only tell the world that he was aiming at being honest, and he would face no negative consequences for his act. This seems like a problem for the intend/foresee distinction.

In "The Intend/Foresee Distinction and the Problem of 'Closeness'", William J. Fitzpatrick suggests that there are three ways that we can resolve this apparent vagueness in the intend/foresee distinction. [41] First, he suggests that many cases of vagueness can be resolved simply by redescribing cases in the right way. He thinks that we often "mistake proximate means as side-effects" and that this is where we get confused when we talk about cases of intending versus merely foreseeing.[42] Second, he notes that some philosophers, like Foot and Anscombe, appeal to the idea that the acts are constitutive of the outcomes in cases where the outcomes are intended but not in cases where the outcomes are merely foreseen. This is an appeal to the *excessive closeness* of acts and their outcomes and is meant to resolve the apparent vagueness in the intend/foresee distinction. The main goal of his paper is to pick up on this argument and develop it further as the best method for accomplishing the task.

Third, he notes that some philosophers take the intend/foresee distinction in a different direction. Fitzpatrick cites Warren Quinn as the originator of this type of view in his paper. Quinn suggests that we can avoid appealing to the issue of closeness by looking not to whether an agent intended to cause some harm, but to whether an agent intended someone else's involvement in the harm she caused.[43] Fitzpatrick, and others, argue quite convincingly against Quinn, who I will only discuss briefly. Another philosopher who steps away from the appeal to closeness is Michael Bratman. His aim is to show us that intentions that are closely related to one's causing harm are not necessary for saying that one brought about that harm intentionally. I will take some time to discuss his view because I think that he does a good job

---

[40] J. Bennet, "Morality and Consequences," *The Tanner Lectures on Human Values II* (Salt Lake City: University of Utah Press, 1981) 111 – 113.
[41] William J. Fitzpatrick, "The Intend/Foresee Distinction and the Problem of "Closeness"," *Philosophical Studies*. 128 (2006): 585 – 617.
[42] Fitzpatrick 588.
[43] Warren Quinn, "Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing," *Morality and Action* (Cambridge: Cambridge University Press, 1993).

of resolving the vagueness in the intend/foresee distinction without appealing to any sort of need for specific intentions to cause harm. This is to say that an agent can *legitimately claim* that she did not intend to cause harm and still cause harm intentionally. This, ultimately, puts a very stringent requirement of moral responsibility on anybody who causes harm – even when they do not intend to do so.

Resolving Vagueness: Proximate Means

I will take the time here to talk more in depth about how Fitzpatrick suggests we resolve the vagueness in the intend/foresee distinction. The first method that he suggested was simply clarifying how harms are related to an agent's intentions when he acts. So, first I will look at his claim that we are often mistaken about whether a harm is intended or merely foreseen based on the idea that we misidentify "proximate means" as side-effects of acts.

According to Fitzpatrick, we know that an outcome was intended if it stands in relation to an act as a means to some end. This is just to say that if an agent brings about some outcome for the purpose of bringing about some other outcome, the original outcome serves as a proximate means to accomplishing the latter outcome; it is a proximate means to an end. Reminding ourselves of the case of the beguiling grocer, we know that profit is his aim. As he goes about pursuing profit he cheats his customers and causes them harm. He states that, while he could foresee that his pursuing profit would result in him causing harm to his customers, he did not intend it. Fitzpatrick would say that the beguiling grocer cannot get away with cashing out his harming his customers as merely foreseen rather than intended because he chose to cheat his customers, thereby causing them harm, for the purpose of gaining profit. Fitzpatrick thinks that when the beguiling grocer cheats his customers, he is doing it as a means to some end, namely acquiring profits. Thus, beguiling grocer did in fact intend to cause them harm. His causing harm to his customers was merely the penultimate step in his profit-gaining activities.

The problem with the view that Fitzpatrick puts forward, the idea that we often get the intend/foresee distinction wrong because we are merely defining the agent's activities incorrectly, is the same problem that I will pose against using the intend/foresee distinction as a way of resolving the inconsistencies in Praise-Blame Asymmetry. Fitzpatrick's view does not support Praise-Blame Asymmetry. He gets the case of the prudent grocer wrong. Reminding ourselves of the case of the prudent grocer, we know that profit is his aim. As he goes about pursuing profit he is fair to his customers and does them some good. He states

that, while he could foresee that his pursuing profit would result in him doing good for his customers, he did not intend it. Fitzpatrick would say that the prudent grocer cannot get away with cashing out his doing good for his customers as merely foreseen rather than intended because he chose to be fair to his customers, thereby doing them good, for the purpose of gaining profit. Fitzpatrick thinks that when the prudent grocer is fair to his customers, he is doing it as a means to some end, namely acquiring profits. Thus, the prudent grocer did in fact intend to do them good. His doing good for his customers was merely the penultimate step in his profit-gaining activities. This judgment that the prudent grocer intended to do good for his customers would mean, according to the folk theory of moral worth, that the prudent grocer was praiseworthy for his actions. This way of resolving the vagueness in the intend/foresee distinction is not consistent with Praise-Blame Asymmetry.

Resolving Vagueness: An Appeal to Closeness

This failure to accommodate Praise-Blame Asymmetry is not a real problem for Fitzpatrick. The major claim of his paper is that appeals to excessive closeness should be sufficient for resolving most cases of vagueness in the intend/foresee distinction. Fitzpatrick argues that an agent is unable to apply the intend/foresee distinction "when the relation between the relevant *states of affairs* is a *constitutive* one rather than a *merely causal* one."[44] He states that when an outcome has a merely causal relationship to the agent's act, it is not a part of the agent's plan; the outcome is incidental to the agent's activities. He cites the Trolley Problem as an archetypal case of a causal relationship between a harm and an act.

The Trolley Problem is a case where an agent diverts a wayward trolley from its original course, in order to avoid killing a group of people, only to wind up killing a person who was standing on the track that the trolley was diverted to.[45] It is often deemed permissible for the agent to divert the trolley, and it is said that the agent who diverts the trolley does not intend to harm the person on the second track. Fitzpatrick thinks that it is circumstantial that there was a person on the second track and that the agent did not intend to kill the person even though she very well may have foreseen it.

She did not go about killing the person on the second track as a means of avoiding killing the group on the first track. This is to say that she did not aim at killing the person on the second track as a way of avoiding killing the group on the first track; she clearly aimed at

[44] Fitzpatrick 593.
[45] Foot 22.

diverting the trolley from the first track (thereby securing the demise of the person on the second track) as a way of avoiding killing the group. If the person on the second track had found a way to escape, this would not have muddled with the trolley-diverting agent's means or ends at all. This is to say that the person's escaping from the trolley or being hit by it was incidental to the aims of the person who diverted the trolley onto the second track.

On the other hand, in cases where an outcome has a constitutive relationship to the agent's act, the outcome is essential to the agent's plans. The beguiling grocer cheats his customers thereby causing them harm. The grocer says that he is not at all interested in harming his customers, and that he is very interested in gaining profit. When he cheats his customers, his aim or what he intends to do is acquire profit, and causing them harm is a merely foreseen side-effect of his profit-gaining activities. We might, however look at what it means for the beguiling grocer to cheat his customers and determine that the relationship between his cheating his customers and harming them is not a merely causal one. This is to say that his cheating his customers is not merely an act that results in their being harmed with cheating and harming being two separate states of affairs. Cheating his customers makes up what is harmful about the act – they were cheated. Because the beguiling grocer's cheating his customers is constitutive of his harming them, and because he intended to cheat his customers, he can be said to have intended to harm his customers.[46]

Comparing the beguiling grocer and the Trolley problem, we can see that cheating one's customers (thereby causing them harm) and diverting a trolley to a new track (thereby securing the death of a person on the track) are two different types of acts. In the trolley problem, the agent chooses an action that directly causes harm to another person. However, we do not think that, by aiming at diverting a trolley to a different track, one has to also aim at harming someone on that track. People divert trolleys all the time without simultaneously seeking to kill with them. However, by aiming at cheating one's customers, that agent is pursuing an effective, and almost guaranteed, means of causing his customers harm. It is the act of cheating his customers that makes his act hurtful. The beguiling grocer cannot say that he was not aiming at harming his customers when he was aiming at a harmful act.

[46] Davidson defends a claim that sounds similar to this in his causal theory of action. He claims that if a person's death is brought about via a shooting, then the shooting and the killing are the same event (Davidson 105). Davidson is obviously discussing how we parse acts rather than how we parse intentions, but the issue here is similar. We cannot separate acts from their outcomes when we think that certain acts and their outcomes are the same event. Those who appeal to closeness are saying that we cannot aim at or intend some act and not its outcome when we think that the act and the outcome are the same event.

A similar example is the case of a terrorist whose aim is to demoralize his enemy and encourage an early end to the war. [47] His method for demoralizing the enemy is to make the population of an enemy's city appear to be dead. He seeks to make the population of the city appear dead by actually bringing about their deaths through the use of bombs. The terrorist wants to say that his intention was to make the people appear dead through the use of bombs, and that the bomb's actually killing them was a merely foreseen side-effect of his act. I think we can assume that the terror bomber is aware that bombs are often lethal and that the use of bombs around people often results in death. Thus, we can say that bombing people is an effective way of killing them. Making the population of an enemy city "appear dead through the use of bombs" is constitutive of killing the people in that city because the use of bombs on people is just a very effective way of killing people. When the terror bomber chooses to use bombs on people for any purpose, he is making an explicit choice to kill.[48]

By appealing to the closeness of the act and its outcome, or by determining whether an act (cheating or making people appear dead by dropping bombs on them) is constitutive of some outcome (people being harmed), it is suggested that we can determine whether some outcome was intended rather than merely foreseen. This breakdown of the intend/foresee distinction fails to accommodate Praise-Blame Asymmetry. To see this, we can see how appealing to closeness when assessing the intentions of agents who bring about positive outcomes results in problematic judgments about the intentions of those agents.

Reminding ourselves of the case of the prudent grocer, we will recall that the prudent grocer treats his customers fairly thereby causing them some good. The grocer says that he is not at all interested in doing good for his customers, and that he is very interested in gaining profit. When he is fair to his customers, his aim or what he intends to do is acquire profit, and doing good is a merely foreseen side-effect of his profit-gaining activities. We might, however look at what it means for the prudent grocer to be fair to his customers and determine that the relationship between his being fair to his customers and doing good is not a merely causal one. This is to say that his being fair to his customers is not merely an act that

---

[47] Bennett 111 – 113.

[48] By using these cases, it may seem that I am claiming that Fitzpatrick is committed to the idea that an act is constitutive of the outcome or merely related causally in cases of certain *types*. Cases of diverted trolleys are not cases where the act is constitutive of the outcome just because an agent's diverting a trolley does not guarantee running people over on side-tracks, and cases of grocers cheating are constitutive of the harm in question just because all cases of cheating result in harming others. Quinn is put off by appeals to closeness because he believes that this is what most philosophers are claiming when they make the appeal (Quinn 173 – 180). Fitzpatrick is concerned with the relationship between causal or constitutive acts and their outcomes only in particular cases, even though it is possible that some types of acts do reliably share a certain relationships to their outcomes (Fitzpatrick 601).

results in some good being done for them with being fair and doing good being two separate states of affairs; his being fair to his customers makes up what is good about the act – they were treated fairly. Because the prudent grocer's being fair to his customers is constitutive of his doing some good for them, and because he intended to be fair to his customers, he can be said to have intended to do some good for his customers.

Because of this assessment of the prudent grocer's intentions, we would have to say that the prudent grocer is praiseworthy when he does good for his customers because he intends the good that he brings about and has, therefore, brought about that good intentionally. This just results in backwards Praise-Blame Asymmetry. It seems right that we want to put restrictions on the intend/foresee distinction and resolve the vagueness in these concepts and the way that agents apply them while describing their activities. These restrictions cannot, however, be such that they parse the cases of intending versus foreseeing positive outcomes incorrectly.

Resolving Vagueness: Avoiding the Appeal to Closeness

Since it seems that appeals to closeness and attempts to redefine cases of intending versus foreseeing as cases of mere intending will get us the same confused result (backwards Praise-Blame Asymmetry), we must move on to another way of trying to get clear on the intend/foresee distinction in order to sort out the inconsistencies in judgments about praiseworthiness and blameworthiness. A third method for sorting out the vagueness in the intend/foresee distinction is stepping away from the appeal to closeness and examining other features of the act that might allow us to say that an agent intended some outcome versus merely foreseeing it. Quinn, a philosopher who is often cited for putting forward this sort of view, suggests that we can avoid appealing to the issue of closeness, by looking not to whether an agent intended to cause some harm, but to whether an agent intended someone else's *involvement* in the harm she caused.[49]

On this view, we know that an agent intended some harm if she sought out to include in her activities some person, and she knew that that person would be harmed in the course of her acting.[50] While the trolley-diverter, in the *Trolley Problem*, knew that the person on the track would be harmed as a result of diverting the trolley to the side-track, she did not set out to enlist the participation of the person who was stuck on the side-track in her trolley-

[49] Quinn 178 – 180.
[50] Fitzpatrick 610.

diverting activities. Once again, we can say that the person stuck on the track is involved in the trolley-diverter's trolley-diverting activities only incidentally. However, in the case of the beguiling grocer, he does have to choose some person to involve in his harmful activities, and, when he chooses to involve her, he intends to harm her whether he succeeds or no.

The merit of Quinn's view is that it takes the emphasis off of explicitly intending harm, completely avoiding issues of agents claiming that they did not intend to cause harm when they acted. All he needs to know is whether the agents intended to involve someone else in their harm-causing activities, and it has be the activities of the agent in question that are directly responsible for the harms caused to the person involved for the agent to have intended that person harm. The fact that Quinn takes the emphasis off of harm is also a demerit of his view because, after all, the intend/foresee distinction and the DDE are both about whether an agent intends some harm. He's trying to answer the question by avoiding it entirely. On Quinn's view, we would say that the beguiling grocer intended harm to his customers. It is clear that the beguiling grocer intended to involve his customers when cheating them; thus, we would say that he intended to cause them harm when cheating them.[51]

Once again, it seems that this view cannot support Praise-Blame Asymmetry. If Quinn means for this view to help us make the distinction between whether an outcome was intended versus merely foreseen, it gets the case of the prudent grocer wrong. In the case of the prudent grocer, he has to choose some person to involve in his good-doing activities, and, when he chooses to involve her, he intends to do good for her whether he succeeds or no. If Quinn wants to say that the beguiling grocer intended to harm his customers, he has to say that the prudent grocer intended to help his customers. Quinn's view is not successful either.

I will now set Quinn's view aside to talk about another view that that steps away from the appeal to closeness in order to resolve the vagueness in the intend/foresee distinction. Michael Bratman offers a view intention and intentional that might allow me to clarify some of the ideas associated with intention, intentional action, and the intend/foresee distinction. To start, Bratman notes that there are at least two ways to conceive of what it means to intentionally perform some act. One view is the Belief-Desire Model, which states that intentions to act are reducible to belief-desire pairs and that intentional actions are actions

[51] While Quinn's view is more sophisticated than what I am describing currently, this question of whether an agent intended the involvement of some person in her harmful activities should be enough to catch an agent like the beguiling grocer in his bad behavior. I do not, after all, take the case of the beguiling grocer to be a particularly sensitive one. We think he intended the harm to his customers when he cheated them, Thus, any view about intending harm, at its most basic, ought to be able to say that he intended harm.

that are based on a complex set of belief-desire pairs, and the other view is the Simple View, which states that, to act intentionally, one must have a specific intention so to act. This is to say that when an agent performs an act intentionally, that act and its outcomes must be supported by intentions that correlate directly to that act and its outcomes. On the Simple View, intentions are primary mental states apart from beliefs and desires. Bratman believes that both of these commonly held views are incorrect. However, because laypeople do not seem to clearly subscribe to either view, it is worth quickly investigating both views to determine whether Bratman can show us where laypeople go wrong with their understanding of intentional action.

It is possible that some laypeople conceive of intentions to act as secondary mental states that are composed of belief-desire pairs and that intentional actions are acts that are based on a complex set of belief-desire pairs. These people would believe that both a desire for some outcome and a belief that some act will bring about that outcome are necessary criteria for an agent's having an intention to act and for an agent's acting intentionally. The Belief-Desire Model appears to be consistent with the folk theory of moral worth that I already put forward.

Using the Belief-Desire Model to evaluate whether an agent like the prudent grocer or the beguiling grocer acts intentionally, we would look only to their relevant beliefs and desires. The prudent grocer desires to acquire profits and he has the belief that being fair to his customers will help him acquire profits. This means that, when the prudent grocer is fair to his customers, he has the intention to acquire profits. On the other hand, the prudent grocer has no such desire to help or be good to his customers even though he does have the belief that being fair to his customers will result in some good for them. This means that, when the prudent grocer is fair to his customers, he has no such intention to do them good. Similarly, the beguiling grocer desires to acquire profit and he has the belief that cheating his customers will help him acquire profits. This means that, when the beguiling grocer cheats his customers, he has the intention to acquire profits. Further, we know that the beguiling grocer has no such desire to harm his customers even though he does have the belief that that cheating his customers will cause them harm. This means that, when the beguiling grocer cheats his customers, he has no such intention to harm his customers.

This breakdown of intentional action results in the odd view that the prudent grocer and the beguiling grocer both performed their actions intentionally and performed them unintentionally – depending on which belief-desire pair we appeal to when describing the act.

The prudent grocer intentionally treats his customers fairly, but he does not intentionally do them any good. This brings us back to the vagueness in the intend/foresee distinction.

The other commonly held view of intentional action is the Simple View, according to which one must have a specific intention to act in order to act intentionally. This view is not obviously at odds with the folk theory of moral worth because it conceives of intentions to act as necessary criteria for intentional action and it is likely that some laypeople conceive of intentions as mental states that are not reducible to beliefs and desires (though they do think that intentions to act are possibly dependent upon them).

Bratman is opposed to the Simple View because it fails to meet an important rationality requirement that we place on intentional action. He demonstrates this by appealing to three examples. In the first two examples, he shows us that we often think that an agent has an intention to act just in case he has the relevant skills to perform the act and has some relevant goal in mind related to the outcome of the act as well as a decision to pursue that goal. In order to act intentionally, an agent simply requires an intention so to act while successfully acting. This break down of intentional action, however, means that agents who try to perform acts with potentially contradicting outcomes have to have the intention that they both succeed and fail at their acts.

In the third example, an agent has to aim (in a video game) two different missiles at two different targets. If he hits one target, he wins, but, if he hits both targets, he loses. If the player had successfully hit both targets, the game would have shut down. The player did not want the game to shut down while he was playing. If the player intended to hit the first target and he intended to hit the second target, this would mean that he intended to perform two acts that were at odds with each other. This amounts to the player both intending to play the game successfully and intending to lose the game while playing it. Bratman says that if this is what the player is doing, then he is guilty of a criticizable irrationality.

According to Bratman, the view that the player held two contradicting intentions is based on the assumption of tight fit. The assumption of tight fit is the assumption that an agent must have a specific intention for some outcome in order for that agent to have brought about that outcome intentionally. The intentional action and the intentions for that action must fit together neatly. According to Bratman, the assumption of tight fit is not useful for ascriptions of intentional action. It can result in contradicting intentions, which are irrational, and it also allows for the vagueness in intending versus foreseeing.

Instead, Bratman suggests that we abandon the assumption of tight fit and conceive of intentional action differently. Imagine, once again, the beguiling grocer. The beguiling grocer

intends to acquire profit when he cheats his customers. When he cheats his customers, he causes them harm in spite of having no intention to cause them harm. Bratman believes that the beguiling grocer intentionally causes harm to his customers even though he only intends to acquire profit while interacting with them. This is to say that it is possible for the beguiling grocer to intentionally harm his customers without intending to harm his customers and while only intending to acquire profit. This is because harming one's customers is within the *motivational potential* of the beguiling grocer's intention to acquire profit from his customers. An outcome is within the motivational potential of some intention just in case it is possible for an agent to intentionally bring about that outcome while executing that intention. We take it for granted that an outcome, like harming one's customers, is within the motivational potential of the specific intention to harm one's customers, but Bratman posits that this same outcome is within the motivational potential of the beguiling grocer's intention to acquire profit when he chooses to do that by means of cheating his customers.

His idea is that we do not often have specific intentions to do a lot of things when we perform an intentional act, and that an agent does not have to have intended a specific outcome in order to have brought that outcome about intentionally. Thus, an agent can be said to have intentionally brought about all of the outcomes of an act, even though many of the outcomes were merely foreseen, just in case she intended to bring about any outcome when she acted. The act was intentional, and ascriptions of intentional action apply when an agent has any intention when she acts. According to the folk theory of moral worth, we are morally responsible for all of the outcomes of an intentional act. If we look again at the folk theory of moral worth, we will note that an agent is morally responsible for her act when she performs a morally relevant act and she does it intentionally. It does not matter whether or not the agent intended a specific outcome. She is morally responsible for bringing about that outcome just in case she intentionally performed the act that it was a consequence of.

Once again, the problem with Bratman's understanding of intentional action is that it fails to accommodate Praise-Blame Asymmetry. Like many other views on what it means to intend a harm, Bratman's gets the case of the prudent grocer wrong. Imagine, once again, the prudent grocer. The prudent grocer intends to acquire profit when he is fair to his customers. When he is fair to his customers, he brings about a good for them in spite of having no intention to do them any good. On Bratman's view, the prudent grocer intends to be fair to his customers when he acts, and he does not intend to do good for his customers. Just in case the prudent grocer intends some outcome when he acts, he acts intentionally. As long as the prudent grocer meets the requirements for intentional action when he acts, he can be

considered morally responsible for his act and all of its outcomes. This is to say that the prudent grocer is morally responsible for his doing good for his customers, and he is thereby praiseworthy. It is possible for the prudent grocer to intentionally do good for his customers without intending to do good for his customers and while only intending to acquire profit. This is because doing good for one's customers was within the motivational potential of the prudent grocer's intention to acquire profit from his customers. Bratman's view accommodates the layperson's desire to say that agents like the beguiling grocer harm their customers intentionally, but he fails to accommodate the view that agents like the prudent grocer do not do good for their customers intentionally.

Moving Away from the Intend/Foresee Distinction

Having explored the various ways that philosophers have tried to resolve the vagueness in the intend/foresee distinction, it does not look like laypeople are misapplying the folk criteria for intentional action because they are confused by the vagueness between intending versus foreseeing. Becoming clearer on this distinction leaves us with the same sort of problem over and over again; we seem to get Praise-Blame Asymmetry wrong. There is one other obvious method that philosophers could use to approach the inconsistencies in Praise-Blame Asymmetry that I have discussed so far. One place in which philosophers might criticize the folk theory of moral worth is by denying some of Malle and Knobe's folk criteria for intentional action, namely the criterion for having a specific intention so to act and a desire for some outcome.[52]

I have already shown how Bratman makes a very good case for denying that a specific intention for some outcome is a necessary criterion for someone's performing an intentional act. Though, I have also shown that this view is not successful in accommodating Praise-Blame Asymmetry. This leaves philosophers only with the opportunity to deny that having a desire for some outcome is a necessary criterion for intentional action. It is highly debatable in philosophy whether this specific criterion is required for someone's intending to do something. Pointing once again to the Chairman in the HELP condition, he is not believed to have helped the environment intentionally, and he is therefore not praiseworthy because he lacks such a relevant outcome-oriented desire. He did not desire to help the environment when he acted.

---

[52] Malle & Knobe 48.

Some philosophers however would deny that such a relevant outcome-oriented desire is a necessary criterion for ascriptions of intentional action. Henry Sidgwick makes this claim explicitly.[53] He notes that when we talk about an agent's motivations, we are referring to the consequences of an act that are foreseeable and desired, and that when we talk about having intentions to act we also have these desired foreseeable consequences in mind. However, he thinks that, when we use the term intention, we are also referring to all of the foreseeable consequences of the act and not just the ones that are desired. This is to say that we are referring to the outcomes that motivate an agent to act and the outcomes that do not motivate an agent to act. This view is compatible with the folk theory of moral worth regarding ascriptions of blameworthiness to morally indifferent agents like the beguiling grocer and the Chairmen in the HARM condition. He states that "we cannot evade responsibility for any foreseen bad consequences of our acts just because we felt no desire for them."[54] He states that once we choose to perform an act with foreseeable bad consequences that these foreseeable bad consequences "are clearly chosen or willed by us."[55]

Sidgwick thinks that if an outcome was a foreseen consequence of an act, then it was also an intended outcome. To him, a belief about an outcome, when one acts, is a sufficient criterion for having intended that outcome. However, this does not accommodate our intuitions. We think that people sometimes knowingly perform acts with negative consequences, and we do not think that they're blameworthy for it. Additionally, Sidgwick's conception of having an intention (and, through this, his conception of intentional action) cannot accommodate Praise-Blame Asymmetry. This view also gets the case of the prudent grocer and the case of the Chairman in the HELP condition wrong. If, when we use the term intention, we are referring to all of the foreseeable consequences of an act and not just the ones that are desired, then the prudent grocer's doing some good for his customers (by treating them fairly) and the Chairman's helping the environment has to be thought of as having been done intentionally. This is because, in both cases, the good consequences of their acts were foreseeable, even if they were not desired.

Laypeople, however, do not think that the prudent grocer and the Chairman in the HELP condition brought about the positive foreseeable consequences of their acts

---

[53] H. Sidgwick, *The Methods of Ethics* (Indianapolis: Hacket Publishing Company, 1981) 202 – 203. Sidgwick is one of the few philosophers that I noted who takes up specific issue with desires as necessary criteria for intentions. Other philosophers bypass the problem by putting forward different criteria for intentions, harp on the fact that beliefs are necessary criteria for intentions, or talk about the role that desires play in reasons, which is not the topic of this section.

[54] Sidgwick 202.

[55] Sidgwick 202

intentionally. Laypeople do not think that the mere fact that the consequence of an act is foreseeable when an act is chosen by an agent is enough to say that that consequence was brought about intentionally. This is something along the lines of what I was saying earlier when I noted that beliefs alone do not appear to be sufficient grounds for generating intentions to act.

Sidgwick criticized the idea that desires are necessary criteria for ascriptions of intentions to act and, ultimately, the idea that they are necessary criteria for ascriptions of intentional action to an agent. I have shown that Sidgwick's argument does not work because it does not accommodate the asymmetry in ascriptions of intentional action to the prudent grocer and the beguiling grocer; his view gets the case of the prudent grocer wrong.

Both Sidgwick and Bratman develop conceptions of intentional action that deny certain criteria set forth by Malle and Knobe's folk criteria for intentional action. Bratman's view serves to help us clarify the intend/foresee distinction, while Sidgwick's serves, in some way, to deny it, or, at least, to deny certain characterizations of it. These conceptions of intentional action, and the various breakdowns of the intend/foresee distinction that I have offered so far are problematic because they fail to match up with ascriptions of intentional action that are made by laypeople. This is important. If we take it for granted that Praise-Blame Asymmetry and the Knobe Effect are phenomena that philosophers want to accommodate, then any view on intentional action that includes all outcomes of an act, intended, desired, or not, as intentional is not a successful conception of intentional action because it cannot accommodate these phenomena.

Ultimately, these philosophers fail to consider the importance of moral motivations when making judgments about moral responsibility. Bratman and Sidgwick want to say that it does not matter whether we are motivated to act by specific outcomes, and that we are morally responsible for all of the outcomes of our intentional acts. However, it is clear that laypeople do not want to say this. They want to say that the prudent grocer was not morally responsible for the good that he brought about because he was not motivated to bring that good about, and laypeople want to say that the beguiling grocer was blameworthy for causing harm because he failed to be properly motivated to avoid causing harm. Agents are not morally responsible for the good or bad outcomes of their acts just because they brought about good or bad outcomes; they are also morally responsible for what motivated them to act to begin with. In other words, they are morally responsible for why they performed that act.

I propose that it may not be conceptions of intentional action that are problematic within the folk theory of moral worth. While laypeople may be getting their criteria for intentional action wrong, I think there is an additional criterion that is not addressed in the empirical model of the folk theory of moral worth. This additional criterion is the reason for which the agent in question acts. While some researchers do in fact refer to this criterion in their papers, none of the researchers who talk about the Knobe Effect or Praise-Blame Asymmetry provide an empirical basis for its inclusion in their folk models. In the next section I will talk about reasons for acting and how they might be included in the folk theory of moral worth. I will eventually follow this by talking about one philosophical view that successfully accommodates Praise-Blame Asymmetry.

Four

I want to return once again to the folk theory of moral worth as I previously presented it. On this view, an agent is praiseworthy or blameworthy just in case she performed a good or bad act and performed that act intentionally. I noted that laypeople tend to ascribe these criteria bi-directionally, meaning that they sometimes ascribe praiseworthiness and blameworthiness prior to ascribing intentional action and just fill in the blanks. I also noted that this method is not philosophically defensible, so I will treat the above as a uni-directional model. So far I have demonstrated that this view, as described from an empirical perspective, is not capable of accommodating Praise-Blame Asymmetry. Laypeople get the case of the beguiling grocer and the Chairman in the HARM condition wrong, since, according to the folk theory, having a specific desire to cause harm is a necessary criterion for determining that these agents intentionally caused harm. These agents have no such desire, and, as a result, they did not intentionally cause harm nor are they blameworthy for their actions.

In the previous section, I took a look at some of the ways philosophers might try to resolve the inconsistencies that I discovered in Praise-Blame Asymmetry. After all, we would all like to think that agents like the beguiling grocer and the Chairman in the HARM condition are blameworthy when they cause harm. I demonstrated two strategies for remedying the problem. First, I tried to determine whether laypeople were simply failing to appropriately make the intend/foresee distinction. I think that this distinction helps redefine the criteria for intentions in the folk theory of intentional action. I clarified how this distinction is made and then applied it to the cases of the prudent grocer and the beguiling grocer. I also examined Bratman, who offers a different set of criteria for what it means to perform intentional actions; he denies the need for specific outcome-oriented intentions. I believe that his criteria offer further clarity on the intend/foresee distinction. Second, I talked a little bit about Sidgwick's view and how specific outcome-oriented desires might not be necessary criteria for intentional action. He takes up issue with the desire criteria in the folk theory of intentional action.

The merit of these views is that they do allow us to say that the beguiling grocer intentionally harms his customers, but, for the most part, it seems that philosophers would deny that the beguiling grocer's motivations matter when we try to ascribe blame to him for the harm that he brings about when he cheats his customers. However, denying the import of an agent's motivations when he brings about morally relevant outcomes means that we would

also have to deny that the motivations of agents like the prudent grocer matter, meaning that he too would be morally responsible and thereby praiseworthy for the good that he brings about when he acts. This claim is inconsistent with Praise-Blame Asymmetry.

So, by denying the role that motivations play in judgments about praise and blame, we get unwanted results, yet we get similarly problematic answers when we do attend to an agent's motivations when we ascribe praise and blame. Since it seems like we really do care about motivations, as demonstrated by the fact that we attend to motivations in the cases of the good-willed grocer, the malevolent merchant, and the prudent grocer, I think that there must be some motivational factor that we are neglecting to talk about. So far, I have touched on issues with desires and intentions, and taking up issue with either of these concepts has led us nowhere. What I have failed to talk about so far is *reasons for acting*, and, more importantly, what it means to *act for a reason*.

In this section, I plan on talking about acting for a reason. I will first take some time to talk about the relationship between reasons and intentions, and why, ultimately, I think we can focus on questions about reasons rather than questions about intentions when we talk about the moral responsibility of agents. Next, I will explore what it means to have a reason to do something versus what it means to have most reason to do something. I will follow this by explaining what we mean when we say that someone has acted for a reason. Finally, I will talk about Nomy Arpaly's view on reason responsiveness, which includes her view that it is possible to act rationally without deliberation. Arpaly's views will be fundamental to my re-formulation of the folk theory at the end of this section, and will be used in further exploration of the problems with Praise-Blame Asymmetry.

Intentions and Reasons

Questions about intentions are questions about what one means to do. As Bratman puts it, intentions to act involve our aims and goals.[56] Similarly, questions about desires are about what one wants to do. This is one way of talking about what motivates a person to act. A different question, however, is to ask for an explanation of one's activities to show that acts and aims match up. When we ask the 'why?' question, we are asking about reasons. Simply put, reasons answer to the 'why?' of actions.[57] They can be cited in the event there is any confusion about why an action could, should, or would be done.

---

[56] Bratman 376.
[57] Anscombe 9.

Philosophers make distinctions between a variety of different uses for the word 'reason;' currently, I am appealing to motivational and explanatory reasons. Motivational reasons are the reasons for which a person actually acts when she acts, and explanatory reasons are reasons that are cited in order to account for an agent's behavior. These two types of reasons play very similar roles in describing why a person has acted the way that she has, but they are sometimes at odds when an agent explains her action by citing a reason that was not actually the reason for which she acted for at all (this is apparent in cases of confabulation or acting on unconscious motivations).

Reasons play a role in intentional action. When an agent acts intentionally, it is important for him to be able to explain the purpose for which his act is performed. Oftentimes, the question "why?" is satisfied by appealing to an agent's aim or what he was trying to accomplish or what he wanted, but this is really a shorthand explanation for one's activities. What we really want to know is that an agent has successfully engaged in means-end reasoning, and that the means that the agent has taken (the act in question) was a legitimate method for accomplishing his end. It is not enough to have some goal in mind when one acts; we also need to believe that the act in question is an appropriate way of reaching that goal. Many philosophers, Williams and Korsgaard for example, believe that the acceptance of these means-end pairs constitutes a reason to act. This is to say that when we have some end in mind and some appropriate means to that end, then we have a reason to act on that means to accomplish that end.

Intentions probably do play a very important role in intentional action, but it seems like it is not the criteria for intentions that have to make or break our ascriptions of moral responsibility. It is possible that the motivations that we are really attending to when we ascribe praise and blame are the specific reasons for which people act (their means-end pairings), rather than just the ends themselves. This could mean that either Bratman's or Sidgwick's claims about the make-up of intentions are true without posing much of a problem for the application of Praise-Blame Asymmetry.[58]

[58] This means that Bratman could be right about the prudent grocer's having, in some sense, intentionally brought about good for his customers, but this is not sufficient for ascriptions of praiseworthiness. If we follow Bratman, the prudent grocer was in fact fair to his customers, and he did do this intentionally. This means that he both acquired profit intentionally and he brought about good for his customers intentionally. This is useful because it reduces the issue of the vagueness in the doing/allowing distinction. However, the prudent grocer was not acting on any sort of morally relevant reason in favor of being fair to one's customer's that is relevant to one's receiving praise. He was acting on reasons; his desire was to pursue profit, he believed (correctly) that being fair to one's customers will acquire profit, Thus, he acted on these beliefs and desires, but he was not acting on the reasons that matter to morality, such as some sort of consideration for the well-being of his custom

Reasons for Acting

I am now going to offer further discussion on what we think it means to have a reason for acting. It is important for me to note that there is a difference between having *a* reason for acting and having reason for acting all-things-considered. On some views, a person has *a* reason to perform an act insofar as she has some end that would be satisfied by her performing that act. For example, according to Williams, an internal reason statement (the only real reason statement on Williams' view) is based, conditionally, on an agent's goals.[59] Suppose the beguiling grocer has a desire for profit and that profit could be made by being honest with one's customers. We would say that the beguiling grocer has an internal reason to be honest with his customers because being honest with one's customers is a means to satisfying his desire for profit. Our beguiling grocer, in this instance, looks a lot like the prudent grocer.

However, we can also suppose that even greater profits can be acquired by cheating one's customers, and, because the beguiling grocer wants to maximize his profits, we might then say that the beguiling grocer has an internal reason to cheat his customers because cheating one's customers is an even better means to satisfying his desire for maximizing profit.[60] On Williams' view, we must show the relationship between the claim that the beguiling grocer has reason to be honest with or to cheat his customers and the fact that the beguiling grocer has some goal, which would be satisfied by his treating his customers in a certain way. In other words, for an agent, A, to have a reason to perform an act, it must be true that performing the act is related to some element, D, in the agent's subjective motivational set, S.[61]

A person can have a reason for performing an act, but still be more rational for not performing that act. For example, I might have a reason to skip class this morning. I am tired, my glasses are broken, and I have not had enough coffee (frankly, I think that any one of these facts provide me with a good reason for not going to class). However, skipping class

---

[59] Williams 365.

[60] When I say that the beguiling grocer has *an* internal reason to cheat his customers, I mean only that he has *a* reason to act in this way but not necessarily reason all-things-considered. The question of whether the beguiling grocer has reason all-things-considered for cheating his customers will be addressed later on.

[61] According to Williams, the subjective motivational set, S, of an agent, A, contains all of her beliefs and desires, and other motivational elements, D. The D's of an S are largely acquired by non-rational means. The S is filled with acculturated beliefs and values and is permeable in that D's can be added to or subtracted from the S through experience and through rational deliberation. According to Williams, an agent has *a* reason to PHI just in case PHI-ing coheres with an agent's subjective motivational set (Williams 365).

When I say that the beguiling grocer has *an* internal reason to cheat his customers, I mean only that he has *a* reason to act in this way but not necessarily reason all-things-considered. The question of whether the beguiling grocer has reason all-things-considered for cheating his customers will be addressed later on.

does not cohere with my other ends. I want to be a good student, I want to attend all of my classes in order to avoid hardship on the exam, and I really like the course material. When comparing the reasons I take myself to have in favor of skipping class (reasons that I am aware of and am responsive to, to the extent they serve as reasons for me) to the reasons I have in favor of attending class, it looks like I have more reason, all-things-considered, to attend class.

It is possible to have conflicting ends, and, as a result, conflicting reasons for acting. In one situation, a person might have one reason for performing act X and a whole lot of reasons for performing act Y, Thus, this person would have more reason, all-things-considered, for performing act Y. I will illustrate the distinction between having a reason and having reason all-things-considered by offering the case of Eleanor. Suppose Eleanor has some goal, namely to get to the hot dog vendor on the other side of the street as soon as possible. Then, so long as she has this goal, we would say of Eleanor that she has a reason to cross the street as soon as possible. Eleanor has some end, getting to the hot dog vendor, and crossing the street as soon as possible would be the best means for reaching that end. A person is thought to engage in successful means-end rationality if she successfully takes the appropriate means to her ends.

However, taking the swiftest or fastest route to one's ends is not always the most rational way to go about getting what one wants. This means that always doing what one has reason to do may not be the most rational thing to do. Eleanor may have a reason to quickly cross the street, but it is important to note that there is a lot of traffic, that the drivers on this road are inattentive, and that the cars on this street have hit many people. These facts, coupled with Eleanor's strong desire to live, means that she has greater reason, reason all-things-considered, to wait until a lull in the traffic or find a cross walk and not cross the street right now, even though she really wants a hot dog.

In fact, if Eleanor did just run across the street to get to that hot dog vendor, we would probably think of her as impulsive – not quite rational in some way. This is because, while it is one thing to take the successful means to one's ends, it is another to perform the act that is best supported by all of the reasons, all of the means-end pairs that one has, and these reasons include more than goals related to individual instances. This is a coherence concept of rationality. An ideally rational person would never have these kinds of conflicting reasons, but, since none of us is ideally rational, we are more rational when we perform the acts that we have the most reason to do or that we have reason to do all-things-considered. Eleanor

may have some reason to cross the street right now, but she seems to have greater reason to wait a bit.

A person is means-end rational insofar as she takes the successful means to her ends.[62] A person is rational insofar as she does what she has reasons to do. These two ideas are clearly closely related. However, a person is more rational when she does what she has most reason to do; that is a person is more rational when she does what she has reason to do all-things-considered. This view is explicitly endorsed by Arpaly, but will also pose a problem for her view on the blameworthiness of morally indifferent individuals. I will discuss this more later on.[63]

Acting for a Reason

Typical cases of acting for a reason involve an agent choosing to perform an act for which she has a reason, and, more importantly, for which she has a reason that she can cite. If this agent has done a good job, then she has performed an act for which she has reason all-things-considered. To come to the point where this agent could make a decision about what she had reason to do all-things-considered, she probably engaged in some form of rational deliberation. Rational deliberation is the activity we engage in when we are reflecting on and evaluating reasons for acting and making judgments about which acts are the best to perform. Making choices often involves deliberation; often we use deliberation to put our instincts and intuitions in check, and deliberated-upon decisions tend to hold some kind of special rational standing compared to other kinds of decisions. Arpaly contrasts her view on reason responsiveness with the standard view of rational deliberation. She argues that, while deliberation is helpful when making decisions and commonly utilized in our decision making processes, it is unnecessary to grant it any kind of special rational status.

Let us suppose a rational agent, Harriet, whose circumstances require of her that she make a choice. In cases of what we take to be rational deliberation, Harriet is trying to decide whether she wants to eat a snack before dinner. To make this decision, Harriet will consider all or most of the reasons that she has for eating or not eating the snack before dinner. She will think about the fact that the snack will ruin her appetite for a healthy meal later on, as well as the fact that she frequently becomes irritable when she is hungry. Since Harriet will

[62] Korsgaard 379.
[63] Arpaly 37.

be interacting with people before dinner and does not want her irritability to get the better of her, she decides to eat the snack.

We would say that Harriet rationally deliberated while determining whether she should eat the snack before dinner. This case is different from a case where Harriet just walks into the kitchen and unconsciously grabs a snack. Yet, in both of these cases we would claim that Harriet did in fact perform a rational act. If we were to ask Harriet, after she unconsciously grabbed the snack, why she had done it, she would probably be able to list all of the above reasons as reasons in favor for her doing what she did; it would simply be the case that she had failed to sort through these reasons consciously before actually performing the act. Both Harriets are moved by the same reasons to eat the snack, but only the first Harriet deliberated and explicitly reflected on her reasons. This is the distinction between being consciously responsive to reasons and being responsive to reasons without deliberation.

On Arpaly's view, an agent does not have to engage in rational deliberation in order to be considered rational when acting.[64] The latter Harriet, above, seems perfectly rational when she seeks out her snack; it is simply the case that she does not engage in deliberation while walking to the fridge and gathering up the food. When I call Harriet rational, I am saying that she has taken the successful means to her ends – she has acted for a reason. To support her claim that rational action does not require deliberation, Arpaly presents examples of actions we generally think are rational but that are rarely deliberated upon. We regard actions as rational or irrational even when we know that the individual whom we are evaluating could not possibly have engaged in deliberative reasoning before the act was performed. Arpaly uses the example of good and bad plays made by expert tennis players. Expert tennis player most certainly do not have time to engage in deliberation while on the tennis court, but we feel perfectly legitimized in applauding for particularly graceful demonstrations and we feel similarly justified in yelling angrily when the tennis player has made a poor move.

Arpaly argues that we are able to make these claims about the rationality of the individual whom we are evaluating because we take this person to be reason responsive.[65] This is to say, we believe that the expert tennis player is acting for a reason when she makes a good or bad play, and we are judging the tennis player's rationality in making a particular play on the reasons for which she is acting. As a matter of psychological fact, Arpaly notes, people frequently act for reasons of which they are not particularly aware. Anyone who has

[64] Arpaly 51.
[65] Arpaly 52.

heard a popular culture reference to Freudian psychoanalysis should be able to imagine covert motivators hidden in the murky unconscious of humans, but Arpaly does not simply mean to state that we act for reasons that are always veiled in this way from the light of consciousness.

I, for example, occasionally forget the deadlines of class assignments when I am under a lot of stress, but can suddenly find myself poring through the syllabus expectantly only to realize that I had a paper due the next day. I start looking through my syllabus on the basis of an inkling or an intuition that I am forgetting something and then I realize that I am. My decision to look through my syllabus is based on a piece of knowledge that I have, but which I did not have immediate access to. I must have noticed the deadline at some point in the past and probably would have readily recalled it if I had not had three other assignments due and ample quantities of caffeine already in my system. In looking through my syllabus, I was responding to a reason that I had for looking through it, even though that reason was not apparent to me at the time.

There are a couple of questions that could be raised regarding Arpaly's understanding of acting rationally without deliberation. From a psychological, rather than philosophical, perspective, it is very hard to say whether a person would ever be able to report the actual reasons for which they were acting. Harriet, above, might never be able to tell us the real reasons she grabbed food out of the refrigerator, even though she is capable of listing off some plausible reasons for acting post hoc. There is ample empirical evidence that demonstrates how we frequently have no idea what reasons we are actually responding to when we act.[66]

In response to this suggestion that it is difficult to tell what reasons an agent actually acted for, it is important to note that this is not a problem for Arpaly. If a person was acting for reasons, period, it is those reasons that the individual is acting on that we would regard with judgments about praise or blame. In fact, this reality about human psychology poses much more of a problem for any view of moral worth that requires that an agent be acting upon deliberated decisions in order to be regarded as praiseworthy because we would not know whether or not the reasons the agent in question reported were the actual reasons for which she acted.

In which case, one might wonder whether it is true that these decisions are actually made without deliberation. The deliberation was done at one point or another. One might

[66] See Kelley J.D. Burns & Antoine Bechara, "Decision Making and Free Will: A Neuroscience Perspective," *Behavioral Sciences and the Law*, 25 (2007): 263–280.

suggest that the nearly reflexive decisions made by a tennis player on the tennis court are the result of years of conscious decision-making in the past. This means that the player is able to make fast rational decisions about what shots to take because she has already spent years practicing and deliberating upon her moves.

In fact, it is true that a beginning tennis player deliberates about what plays she should make, and, as she advances, she comes to require less deliberation when she plays. This example poses no threat, however, to Arpaly's account. On Arpaly's view, it only needs to be true that the person does not consciously deliberate when the act is performed. She would not deny that the tennis player's good playing comes about as a result of prior deliberation, but prior deliberation does not entail that deliberation occurs when the most recent act is performed. It is important to note, once again, that Arpaly's view is partially in response to the long held belief that only actions performed from one's most recent deliberations are rational actions. Clearly, the tennis player is playing rationally, even if she is not actively deliberating. This fact says nothing about prior deliberation or prior commitments to acting rationally. We want to assess the agent on the basis of her rational processes when she performs the act because we can really only assess her potential candidacy for praise or blame on the basis of the reason for which she was acting when she performed the act.

Reasons and Praise-Blame Asymmetry

In terms of the role of reasons in Praise-Blame Asymmetry, it looks like, in the case of the grocers, laypeople are attending, largely, not to whether they performed their acts intentionally, but instead to the reasons for which they were acting when they interacted with their customers. I want to take another look at Praise-Blame Asymmetry and try to further explain this phenomenon. In cases where Praise-Blame Asymmetry has been observed, we are presented with something like the following example: There is some act A and its unintended side-effect X. Under condition 1, X is a desirable state of affairs. We would say that an agent, S, is not praiseworthy for bringing about side-effect X while performing act A. Under condition 2, X is an undesirable state of affairs. We believe that agent S is blameworthy for bringing about side-effect X while performing act A. The judgment that an agent is blameworthy for bringing about some unintended negative side-effect but not praiseworthy for bringing about some unintended positive side-effect demonstrates an asymmetry in our actual judgments about praise and blame. Neither of the side-effects brought about by the agent were produced by her intentionally (or both of them were,

depending on which conception of intentional action we adhere to), but this is irrelevant to our actually praising or blaming the agent in question. Hindriks suggests that we make these asymmetrical judgments about praise and blame because we are attending to the agent's motivational versus normative reasons for performing or not performing act A.[67]

Taking another look at the previous example, under condition 1, if agent S had pursued act A with side-effect X in mind, then agent S would have been praiseworthy. However, S did not pursue act A with side-effect X in mind, and so we think that S is just not praiseworthy. When making these judgments, we attend to S's motivational reasons, the morally relevant reasons for which S was actually acting, when we determine whether she is praiseworthy. In this case we determine that S is not praiseworthy for bringing about side-effect X because we believe that the reasons for which the agent was actually acting were not enough to merit her praise.

Under condition 2, we believe that the agent is blameworthy for performing an act with a negative side-effect because we determine that the agent in question should have been more attentive to the morally relevant reasons that exist in favor of avoiding act A. Reasons that we think should be responded to, independent of what the agent actually does, are normative reasons. In other words, we have the belief that reasons exist in favor of agent S's not performing act A, and we think that his failure to attend to these reasons when he brings about negative side-effect X makes him blameworthy. Because the agent performed act A while failing to acknowledge normative reasons that existed for not performing act A (namely that it would bring about side-effect X), the agent is considered blameworthy for bringing about side-effect X.

The asymmetry in judgments about acts occurs when the agent is perceived as intentionally performing some right or wrong act, and when the agent fails to perform the act for its rightness or wrongness. An agent who intentionally performs a right act without caring that the act is right is not praiseworthy for performing that act because the reasons for which she is acting do not merit praise. An agent who intentionally performs a wrong act without caring that the act is wrong is blameworthy for performing the act because she should have attended to the reasons in favor of her not performing the act.

If we recall the cases of the good-willed grocer and the malevolent merchant, we can note the symmetry in our judgments about their praise and blame. The good-willed grocer performs the right act for the right reasons and we think that she is praiseworthy. The

[67] F. Hindriks, "Intentional Action and Praise-Blame Asymmetry," *The Philosophical Quarterly*, 58.233 (2008): 630 – 641.

malevolent merchant performs the wrong act for sinister reasons and we think that she is blameworthy. This is because we think that they performed their acts with the reasons for which they were right or wrong in mind.

However, when we consider the cases of the prudent grocer and the beguiling grocer, we judge them asymmetrically. In discussing the potential praiseworthiness of the prudent grocer, we attend to the motivational reasons for which the prudent grocer is acting. Because the prudent grocer was honest with his customers only because he wanted to acquire profit, he is not praiseworthy for performing the right act.

In discussing the potential blame of the beguiling grocer, we appear to be attending to the normative reasons for which the beguiling grocer is failing to act while regarding the reasons for which he actually acted, his motivational reasons, as irrelevant. It does not seem to matter to us that the beguiling grocer cheated his customers only because he wanted to acquire profit. He failed to respond to the relevant moral reasons in favor of not cheating his customers. For this, he is blameworthy. While both the beguiling grocer and the prudent grocer act for morally irrelevant (motivational) reasons, only the beguiling grocer is determined to have performed an act that has moral worth, and this is because we attend to the normative reasons for which the beguiling grocer is failing to act.

Here I will note that, in the case of Praise-Blame Asymmetry, in order for laypeople to determine that the beguiling grocer, or the morally indifferent individual, is morally blameworthy, they have to adopt an alternative strategy. By this I mean to say that laypeople fairly consistently attend to the reasons for which the agent actually acted or the motivational reasons when they make judgments about praiseworthiness and blameworthiness, but they have to make an exception to this rule in order to accommodate the view that the morally indifferent individual is blameworthy.

When they appeal to normative reasons to justify blaming the morally indifferent agent, they run the risk of appealing to reasons that the agent does not actually have in favor of acting. While making judgments about moral worth of the acts of the good-willed grocer, the malevolent merchant, and the prudent grocer, laypeople appeal to the morally relevant reasons for which these agents were acting. If they did the same in the case of the beguiling grocer, they would determine that he acted for the same morally irrelevant reasons as the prudent grocer and consider him not blameworthy. This problem parallels the original issue that I brought up with the folk theory of moral worth in that there appears to be an asymmetrical application of criteria for praiseworthiness and blameworthiness that is not obviously justified.

Reformulating the Folk Theory of Moral Worth

Wright and Bengson do include in their explanation of Praise-Blame Asymmetry reasons for acting, though they do not offer any empirical evidence for including it.[68] As I noted earlier, these authors suggested a bi-directional folk model, which I later demonstrated was not philosophically defensible. I will return to their model and demonstrate how it can be used uni-directionally by ultimately ignoring ascriptions of intentional action as a necessary requirement for ascriptions of praiseworthiness and blameworthiness.

According to Wright and Bengson, under normal circumstances, as I have noted, laypeople first determine the evaluative status of an act (good/bad), then determine whether that act was done intentionally, and then ascribe moral responsibility to the agent who performed the act (praiseworthiness/blameworthiness). Depending on which conception of intentional action we go with, this way of assigning moral worth to an act gets the case of either the beguiling grocer or the prudent grocer wrong. To accommodate instances where laypeople get cases wrong (which would only be in the case of morally indifferent agents who act), the authors suggest that laypeople sometimes apply these rules bi-directionally. So, if the negative outcome of an act is foreseen and the act is bad, then when that act is performed the agent who performed that act is blameworthy (this is the case of the beguiling grocer). In their bi-directional model, the judgment that the outcome of an act was foreseen, that the act was wrong, and that the agent was blameworthy for performing the act, allows us to determine that the act was performed intentionally. I already showed why making this leap does not work from a philosophical standpoint.

What I am suggesting in this section is that we can look to an agent's acting for a reason to fulfill the requirement of acting intentionally that I put forward in Section Two. More importantly, as Hindriks suggests, we can look to an agent's acting or failing to act for a certain set of reasons for making judgments about moral responsibility. When an agent acts or fails to act for these reasons, when she performs a right or wrong act, she is appropriately called praiseworthy or blameworthy. This is to say that when an agent performs a right act for the right reasons, we think she is praiseworthy, and that when an agent performs a wrong act for the wrong reasons or when she fails to act for the right reasons, we think she is blameworthy.

---

[68] Wright and Bengson 29 – 30.

In Section Five, I am going to suggest a philosophical view, that of Nomy Arpaly, which seems to accommodate this lay view of moral worth. I will flesh out her view, and determine whether this type of view is capable of supporting the Praise-Blame Asymmetry that arises when it is employed. Ultimately, I think that even a more sophisticated theory of moral worth, like Arpaly's, fails to clearly justify Praise-Blame Asymmetry. Though I will not be able to offer a comprehensive critique of Arpaly's view, I will suggest a couple of places where pushing on her theory will lead us to confusion, as well as suggest a couple of ways to resolve this confusion.

Five

So far, I have made an effort to describe a folk theory of moral worth that successfully accommodates Praise-Blame Asymmetry. I am now going to undertake to describe a more well-developed theory of moral worth that successfully accommodates Praise-Blame Asymmetry and parallels the criteria for ascriptions of moral worth in the folk model I have presented, and I am going to take the time to question whether a view like this can be justified.

In her book Unprincipled Virtue, Nomy Arpaly creates a theory of moral worth that accommodates both our asymmetrical judgments about praiseworthiness and blameworthiness as well as our intuitions that people are potential candidates for praise and blame on the basis of the reasons for which they act. I think her view also does a good job of including information from the field of psychology to better explain when we think people can be potential candidates for praise and blame.

Ultimately, I want to use Arpaly's view on moral worth as one example of how we make asymmetrical judgments about praise and blameworthiness and I am going to criticize this view, showing where our judgments, in this asymmetry, go wrong. Many philosophers come to similar conclusions about praiseworthiness and blameworthiness in their theories, but I think that Arpaly's view expands potential candidacy for praise and blame in a way that makes escaping it more difficult. Since I want to offer a sincere challenge to this Praise-Blame Asymmetry, I wanted to put my views up against a theory of moral worth that successfully appeals to both our intuitions and actual human psychology.

Arpaly's project consists of two important parts. First, she shows that an agent can be more rational by acting against her best judgment than she would have been if she had been acting in accordance with her best judgment. Second, she shows that an agent can be praiseworthy or blameworthy just in case she is acting for certain reasons. The connection between these two parts is that if we can claim that a person who is acting against her best judgment is in fact acting for reasons, then she is a potential candidate for praise and blame when she acts. This view expands the realm of potential candidacy for praise and blame to anybody who can act for reasons, no matter how irrational she may take herself to be. On this view, the act of anybody who is acting for the relevant moral reasons has some moral worth, and the agent who performs the act is an appropriate candidate for praise or blame.

In this section, it is going to be important for me to explore a number of concepts in order to better explain what Arpaly's view of moral worth actually is. First, I must distinguish between different types of acts. Some acts are morally required and some acts are morally irrelevant, and this distinction is going to play an important role throughout this paper. Second, I will follow Arpaly and make a distinction between moral desirability and moral worth. Where these two concepts come apart is a central point that I will get back to later.

Finally, I will address the question of negative moral worth and candidacy for blameworthiness. My description of Arpaly's view on negative moral worth and blameworthiness is partly an interpretive project. This is to say that her view is unclear in some places, and I will set out to determine what I think her view actually is, or, at least, what would be the strongest position for her to hold.

Types of Acts

As I have already noted, I am first going to address the topic of types of acts in order to get a better idea of when we are going to think of an agent as a potential candidate for praise or blame. Typically philosophers make a distinction between right acts and wrong acts. Right acts are acts that are in accordance with morality and wrong acts are acts that are not in accordance with morality. Furthermore, right acts are morally desirable and wrong acts are not. Right acts are morally required and wrong acts are morally prohibited. Acts that are neither morally required nor morally prohibited are morally permissible. The good-willed grocer acts in accordance with morality and is performing the right act by treating her customers fairly, and the malevolent merchant acts contrary to morality and is performing the wrong act by cheating her customers.

Moral Desirability versus Moral Worth

Certain acts are more morally desirable than others are. Following Nomy Arpaly, I will make a distinction between an act's moral desirability and its moral worth or moral value.[69] The extent to which an act is right or wrong determines the act's moral desirability. By moral desirability, I also mean the extent to which an act is in accordance with the requirements of morality. In other words, moral desirability admits of degrees. An act that involves more right-making features than wrong-making features is more morally desirable

[69] Arpaly 69.

than an act with more wrong-making features than right-making features. The good-willed grocer's act of honesty is morally desirable and the malevolent merchant's act of cheating is not morally desirable. In popular discourse, telling the truth is more morally desirable than telling a white lie, but telling a white lie is more morally desirable than telling a lie that is hurtful.

Two acts of comparable moral desirability can, however, differ in terms of their moral worth, so two individuals might perform the same morally desirable act, and yet these acts might have different moral worth. Suppose two different women provide their school-age children with lunches. One woman provides her children with lunches because she loves them and deeply cares about their well-being. The second woman provides her children with lunches because she wants the other mothers to think well of her. While both women perform the right and equally morally desirable acts of providing their children with lunches, the first mother's act has greater moral worth than that of the second mother.

Questions about moral worth have been addressed by a number of philosophers, and many philosophers come to the same conclusions about the potential praiseworthiness or blameworthiness of certain agents. On the accounts of Kant and Arpaly, for example, the good-willed grocer is praiseworthy for being honest with her customers, but the prudent grocer is not. Even so, Kant's and Arpaly's views are importantly different. Kant suggests that an act that has moral worth is a right act done from the motive of duty.[70] This means that an act that has moral worth is a right act performed by an agent for precisely the reason that it is believed by the agent to be the right act. On this view, the motive of duty is a moral reason for acting. Hence, for Kant, the good-willed grocer's act has moral worth because she treats her customers honestly as a result of believing treating her customers honestly is the right thing to do. By contrast, though the prudent grocer treats his customers honestly, Kant says his act lacks moral worth, and the prudent grocer is not praiseworthy.

An agent who chooses her act because it is the dutiful act is different from an agent who performs the right act out of an inclination. Only the agent who acts from the motive of duty is a potential candidate for praise. This distinction is similar to the one between the agent who acts from the motive of duty and the agent who acts from the motive of prudence. While visiting one's mother in the hospital is a very nice thing to do, on Kant's view, the act only has moral worth if the act is chosen from the motive of duty. This is not to say that the dutiful agent cannot derive pleasure from performing the right act, so long as the motivation

[70] Kant 10.

for an agent's choosing the pleasant act is foremost the fact that it is the act that accords with one's duty.

Like Kant, Arpaly thinks the good-willed grocer is praiseworthy and thinks that the prudent grocer is not praiseworthy. Even so, her account is different from Kant's. Arpaly believes that acts that have moral worth are right acts that are performed by an agent for the relevant moral reasons. Agents who perform morally right acts for the relevant moral reasons are praiseworthy.[71] She takes responsiveness to relevant moral reasons to be an agent's desire to perform actions with the appropriate right-making features and to perform actions without wrong-making features. Arpaly calls this *Praiseworthiness as Responsiveness to Moral Reasons* (PRMR), and she calls the motive from which praiseworthy actions arise *good will*. On both Kant's and Arpaly's views, acts that have moral worth have it in virtue of the agent's having the right motive when she acts.

As noted, Kant's and Arpaly's views differ in important ways. Arpaly's view, but not Kant's, allows for degrees of moral praise. Arpaly offers a revised version of the PRMR, which includes the importance of degrees of moral concern in our judgment about moral praise.[72] Her ideas on depth of moral concern are consistent with our intuitions that people can be more or less motivated by the moral reasons that count in favor of or against their performing certain acts. The *merely morally-inclined individual* and the *moral enthusiast* might both be responsive to relevant moral reasons that count in favor of their visiting their mothers in the hospital, but only the moral enthusiast is motivated to perform the right act by these considerations. The merely morally-inclined individual might visit her mother for some moral reasons (compassion for suffering people) and some not-so-moral reasons (the expectation of an inheritance). The merely morally-inclined individual might not show up to the hospital if she feels stressed out about work, or if the traffic is heavy. The moral enthusiast is very motivated to act in accordance with morality, she is acutely responsive to moral reasons, and she always shows up to the hospital regardless of inheritances, stress, or heavy traffic.

The merely morally-inclined individual does act for relevant moral reasons; she does care about her mother, but these considerations can be outweighed by other kinds of considerations, meaning her concern for moral considerations are not so very deep that they successfully motivate her whenever she is faced with a moral decision. She is still

[71] Arpaly 72.
[72] Arpaly 84.

praiseworthy for visiting her mother in the hospital. She performs the right act for relevant moral reasons. It is just the case that she is not as praiseworthy as the moral enthusiast.

On Arpaly's view, agents are more praiseworthy the more concerned they are with morally relevant reasons.[73] She takes moral concern to be the extent to which someone is concerned with actual moral reasons rather than reasons he believes to be moral reasons. An actual moral reason for acting is a reason that corresponds to the morally relevant features of an act, whereas a reason that is believed to be a moral reason may or may not correspond to the morally relevant features of the act. We can be mistaken about the truth of our beliefs. Arpaly's understanding of degrees of moral concern will be very important when we discuss her conception of the blameworthiness of moral indifference.

The Moral Worth of Wrong Acts

So far, I have explained that Arpaly thinks that a person is praiseworthy when she performs the right act for the relevant moral reasons, and she thinks that a person is more praiseworthy the more she is responsive to relevant moral considerations in a given situation; this person demonstrates a greater depth of moral concern. Acting for the relevant moral reasons does not amount to acting from the motive of duty or acting on one's all-things-considered deliberated-upon decisions. Acts have moral worth insofar as an agent is acting for the relevant moral features of a situation. This brings us to our judgments about the moral worth of wrong acts and judgments about the blameworthiness of morally indifferent agents like the beguiling grocer. According to Arpaly, an agent can be considered blameworthy for performing an act when she has been insufficiently responsive to moral reasons, or when she has been responsive to malicious reasons, when she has a *deficiency of good-will* or is *ill-willed*.[74] To be unresponsive or insufficiently responsive to moral reasons is to lack the desire to perform actions for their appropriate right-making features, and to be responsive to malicious reasons is to perform wrong acts for the reasons for which they are wrong. An act with negative moral worth is a wrong act performed by an agent who has a deficiency of good will or who is ill-willed. I will call a deficiency of good will, or insufficient responsiveness to moral reasons, moral indifference.

It is difficult to discern Arpaly's view on blameworthiness. In some places, it seems like she is saying that performing a wrong act is a sufficient condition for ascribing blame,

[73] Arpaly 84.
[74] Arpaly 79.

and in other places it seems as though performing a wrong act is only a necessary condition, and in others places in her writing performing a wrong act seems like neither a necessary nor sufficient condition for being blameworthy. In some places in her writing, Arpaly seems to indicate that all wrong acts have negative moral value and that moral agents who perform wrong acts are blameworthy. For example, she does not think that individuals who act out of ignorance are excused from blame when they perform wrong acts; she just thinks that individuals who act out of ignorance are less blameworthy than individuals who act for malevolent or morally indifferent reasons. Hence, it seems like Arpaly thinks that it is a sufficient condition for ascribing blame that an individual performs a wrong act. However, Arpaly also notes that individuals who are irrational in certain ways can be excused from blame because the reasons for which they act say nothing about their regard for moral reasons. This suggests that performing a wrong act is not a sufficient but a necessary condition for being blameworthy.

Arpaly's view here makes a difference because it has implications for how successful her claims about candidacy for blameworthiness will be, and Arpaly probably does not want there to be a lot of ambiguity about what her view actually is. That said, she also seems to indicate that being responsive to sinister reasons at all is something for which a person can be considered blameworthy, suggesting yet a third interpretation of her criteria for blameworthiness. An individual who is responsive to sinister reasons, but who never acts on them and, thus, never performs a wrong act, might be blameworthy.

It seems a little strange to say that a person who did not act on his sinister motivations might be blameworthy because this would make actually performing a wrong act neither a necessary nor a sufficient reason for ascribing blame. We typically blame people for the acts that they actually perform. However, if this is the view that Arpaly has, then it also provides us with an interesting way of interpreting her view. Arpaly might be more interested in assessing people on the reasons that they are responsive to than on the acts they perform. This is to say that Arpaly might be more interested in the reasons that motivate an agent than the rightness or wrongness of any act he might perform. I have a suspicion that this is, in fact, what Arpaly is doing in some places in her book, even though I do not think that this is the interpretation of her view that she actually endorses. I will explore the ambiguity in Arpaly's view in greater depth later.

To demonstrate how we attribute blameworthiness, Arpaly suggests that we should imagine the interactions between Jeanne and Joseph.[75] Jeanne tells Joseph to "shut up," and her doing this hurts Joseph's feelings. While we always think that telling people to "shut up," thereby hurting their feelings, is wrong, there are some circumstances where doing so is less blameworthy than others. In the first case, Jeanne comes from a country where telling people to "shut up" is a common practice, Thus, it is not typically something that people respond to with hurt feelings. When Jeanne comes to the United States and tells her friend Joseph to "shut up," her doing so says nothing about how concerned she is with morality. She makes what we call an honest mistake. If she had known that telling Joseph to "shut up" would be hurtful, she would have avoided doing so. While Jeanne is probably blameworthy for performing the wrong act with negative moral worth, she is not *particularly* blameworthy for her offense.

In the second case, Jeanne knows perfectly well that telling her friend Joseph to "shut up" would be hurtful, but she is just so frustrated from her day at work that she does not particularly care about how her behavior will be received. When Jeanne acts in this case, she is indifferent to the moral reasons that exist in favor of her not offending her friend Joseph, and she is, therefore, blameworthy. However, Jeanne's behavior in the second case is still less blameworthy that her behavior in the third case where she tells Joseph to "shut up" precisely because she knows that her doing so will be hurtful to him. Joseph is a sensitive person, prone to having strong, negative responses to rudeness, and, on this day, Jeanne is taking advantage of the opportunity to get a rise out of Joseph. It makes her feel powerful when she can hurt people. Jeanne is responsive to sinister reasons when she tells Joseph to "shut up" in this case, and this makes her blameworthy.

In case one, Jeanne is acting out of ignorance of moral reasons, in case two, Jeanne is acting out of indifference to moral reasons, and, in case three, Jeanne is acting from sinister motivations. While Arpaly seems to think that all three cases are cases of blameworthiness, she thinks that some of these blameworthy acts are worse than others. What is interesting to me about Arpaly's view is that she does not think that all cases of moral indifference are instances where the agent in question is morally blameworthy. Recall the comparison I drew between the prudent grocer and the beguiling grocer. Neither agent has any desire to perform the right act, and neither agent responds to moral reasons when he performs the act, but one

[75] Arpaly 79.

agent is thought to be blameworthy and the other praiseworthy on the basis that one performs the wrong act and one performs the right act.

Six

In the previous section, I attempted to provide the reader with a brief overview of Arpaly's position. Particularly important to my paper is Arpaly's view on how we make judgments about blameworthiness. However, as I mentioned, Arpaly's written work leaves me with some interpretive issues regarding her view on blameworthiness. I will attempt to resolve these issues now. In this section I am going to try to determine the necessary and sufficient conditions for ascribing blameworthiness to a morally indifferent agent on Arpaly's view. First, I try to determine whether the actual performance of a wrong act is a sufficient or necessary condition for ascribing blame to an agent. In fact, I suggest at one point that performing the wrong act is neither, though I will ultimately determine that Arpaly takes it to be a necessary condition for ascribing blame to an agent. Second, I will try to determine whether being morally indifferent when one acts is a sufficient or necessary condition for ascribing blame to an agent. The formulation of her view that I establish in this section will be the criteria for ascribing blameworthiness to a morally indifferent agent that I will use throughout the rest of the paper.

First, I am going to ask whether it is a sufficient condition for ascribing blame to an agent that he performed a wrong act. Is the beguiling grocer blameworthy just because he cheated his customers? If we look at her work, we can see that Arpaly cannot get away with stating that the rightness or wrongness of an act alone is enough to assign praise or blame to an agent. This is to say that acts do not have intrinsic moral worth. This is why Arpaly makes a distinction between the moral desirability of an act and its moral worth.

The distinction between the moral desirability of an act and its moral worth is the reason why the prudent grocer's honesty with his customers is considered not praiseworthy. The prudent grocer is not praiseworthy for being fair to his customers, even though being fair to one's customers is, in fact, morally desirable. This is because the reasons for which he acted were not the relevant moral reasons. In the case of the beguiling grocer, he performs an act that is morally undesirable, but not necessarily an act that has negative moral worth (an act for which he can be blamed). Arpaly seems to think that something else is required for assigning moral worth to an act and praise or blame to an agent, namely the fact that the agent acted for some relevant moral reasons. This fact is the centerpiece of her work and is contained explicitly in the PRMR. Arpaly takes a person to be a potential candidate for praise

or blame when she has performed some act with certain moral reasons on her cognitive horizons.[76]

Given what I have already discussed, it looks like one would not be able to take Arpaly as saying that performing a wrong act is a sufficient condition for ascribing blameworthiness. However, Arpaly's view becomes unclear when we take a look at some of her examples. She offers the example of Boko, an individual who captures and roughly ties up a man who is sneaking around in his father-in-law's backyard in the middle of the night.[77] Boko mistakes this man for a thief; it was an honest mistake. He was unaware of some of the facts of the situation. Boko harms the man, but he was not intending to cause harm to some non-thief, and probably was not really intending to cause harm at all.

Arpaly notes that Boko is less blameworthy than people who intend to cause harm, but he is blameworthy nonetheless.[78] However, Boko acted for neither sinister reasons, nor was he particularly morally indifferent, so in this case it seems that his performance of a wrong act is sufficient to make him blameworthy. Arpaly makes similar judgments about other "ignorant" agents, taking them to be blameworthy when they perform wrong acts. However, in later sections of her book, she does note that people who are crazy in certain ways might be excused from blame when they perform wrong acts. The example she looks to is the man who mistook his wife for a hat.[79] He mistreated his wife, but was not the type of being who could handle reasons in the right ways to be considered a candidate for blame.

I think Arpaly believes that Boko was in fact insufficiently responsive to moral reasons, that he should have known better than to cause harm to anyone, regardless of his thief status. I am just left wondering how acting from ignorance and acting from irrationality are relevantly different when Arpaly is making claims about acting for reasons. If the reasons that Arpaly wants agents to act on are not apparent or available to the agent in question, it is hard to see how the agent could have acted on them. Yet, even if the agent does not see the reasons that he is supposed to act on, Arpaly would still think of this agent as blameworthy. I can accept that Arpaly, ultimately, does not want to make the claim that the rightness or wrongness of an act is not enough to ascribe praise or blameworthiness, but I think that it is worthwhile to note where this interpretation of Arpaly's view might be met with resistance.

Let us suppose that Arpaly does not want to say that an agent's performing a wrong act is a sufficient condition for her being blameworthy. Does Arpaly take it to be a necessary

[76] Arpaly 72.
[77] Arpaly 102.
[78] Arpaly 103.
[79] Arpaly 150.

condition for assigning blameworthiness to an agent that he performed a wrong act? Given the information I provided earlier in this section, this looks like it is going to be the best interpretation of her view, however, there is some room to contest this interpretation. I will discuss this inconsistency in her view and then get back to the question I just posed.

In two different chapters in her book Unprincipled Virtue, Arpaly notes that being responsive to sinister reasons, at all, is bad, indicating that being motivated by sinister reasons at all might make one blameworthy regardless of whether or not the agent in question actually acts on those reasons. In the first instance, Arpaly is explaining that having certain motivations is not inherently problematic. Desiring money or love, for example, is not a moral problem, even if being overly susceptible to these concerns leads one to act in morally problematic ways. She states, however, that being susceptible at all to racist or hurtful motivations is always bad, even if they are never actually acted on.[80] The second time she brought up the idea that motivations themselves can make a person blameworthy she notes that acting on any sinister motivations makes one blameworthy, even when it does not result in a wrong act. Further, she states that "one can be condemned for having racist or sexist desires in the first place."[81] This interpretation of Arpaly's view might lead us to believe that actually performing a wrong act is neither a necessary nor a sufficient condition for ascribing blame to an agent.

To further support this interpretation, Arpaly notes, in a personal communication,[82] that she takes some thoughts to be acts, so having racist thoughts, even if one does not act on them, might make one blameworthy. The suggestion that one need not actually perform an act to be blameworthy justifies our intuitions about people who we think are blameworthy for having bad intentions, even if they are not efficacious (cases of trying and failing). A terrorist might set out to plant a bomb in a NYC subway and get distracted along the way. We still might think he is blameworthy even though he never actually followed through with his plans.

It is possible that Arpaly assumes that having certain motivations always results in one's performing certain kinds of acts, or, rather, it is possible that she assumes that having certain motivations when one actually acts always results in the performance of certain kinds of acts. A person, like the good-willed grocer, who is responsive to the relevant moral reasons, will always treat her customers fairly. A person, like the malevolent merchant, who

[80] Arpaly 81.
[81] Arpaly 143.
[82] Nomy Arpaly. Wheaton College. Austin House, 26 East Main Street, Norton MA. 03 May 2010. Guest Lecture.

is responsive to sinister reasons, will always cheat her customers. Assigning praise or blame on the basis of motivating reasons for acting would successfully ascribe praise or blameworthiness to both of these agents when they are both efficacious and not efficacious.

However, attending to motivating reasons, or the reasons for which an individual acts, will not allow us to successfully ascribe praise or blameworthiness to all agents, and this is because of the Praise-Blame Asymmetry I described at the beginning of this paper. Both the prudent grocer and the beguiling grocer act for the same reasons, but one performs the right act and the other performs the wrong act. If only the reasons for which the agent actually acts matter when assigning praise and blame, then both the prudent grocer and the beguiling grocer would be equally not praiseworthy or blameworthy. In other words, if we only care about the reasons for which these two agents act, then we would have to regard them with the same moral attitudes because they act for the same reasons.

If Arpaly believes that the prudent grocer and the beguiling grocer should be regarded with different moral attitudes, then she must think that the wrongness of the act in question matters to some extent when assigning blameworthiness. In fact, in a personal communication, Arpaly notes that praise and blame can only be assigned to an agent for actually performing an act.[83] Agents who have good or bad intentions, but who fail to act on them, can be called virtuous or vicious but not praiseworthy or blameworthy. Thus, the racist or the sexist, who has sinister motivations, is considered vicious on Arpaly's view but not blameworthy for his sinister motivations. Getting back to my original question then, Arpaly takes it to be a necessary condition for assigning blameworthiness to an agent that she performed a wrong act.

The next interpretive problem I have with Arpaly's view is whether being insufficiently responsive to moral reasons is a sufficient or necessary condition for assigning blameworthiness to an agent for acting. I pose the question of whether it is a sufficient condition for an agent's being blameworthy that he is morally indifferent when he acts for two reasons. The first is that this question highlights, once again, the fact that the prudent grocer and the beguiling grocer act for the same reasons, which is a topic that will get more discussion in the next section on Praise-Blame Asymmetry. The second reason is that rejecting the sufficiency claim lends support to the interpretation of Arpaly's view that suggests that performing a wrong act is a necessary condition for ascribing blame.

[83] Arpaly, Guest Lecture at Wheaton College.

In Section Five, where I discussed the moral worth of wrong acts, I noted that it is Arpaly's view that an agent is blameworthy when he is insufficiently responsive to the relevant moral reasons when he acts. I wonder whether Arpaly thinks that being insufficiently responsive to the relevant moral reasons for acting is a sufficient condition for an agent's being blameworthy. Arpaly's view requires that being insufficiently responsive to the relevant moral reasons for acting is not a sufficient condition for assigning blameworthiness because the prudent grocer and the beguiling grocer, as I have already noted, act for the same reasons. If merely being insufficiently responsive to moral reasons, when one acts, is sufficient for ascribing blame, then we would have to think that the prudent grocer is blameworthy or that the beguiling grocer is just not praiseworthy. This is very much not what Arpaly wants to say, and it provides support for the interpretation that actually performing a wrong act is a necessary condition for ascribing blame on Arpaly's view. This is because, while Arpaly does not think that the rightness or wrongness of the act alone is enough for an agent's being praiseworthy or blameworthy, she does think that it will play a role in determining praise or blame.

The next question is whether it is a necessary condition for an agent's being blameworthy that he is morally indifferent when he acts. The answer to this question is no. This is because Arpaly very clearly thinks that agents who act on sinister reasons are blameworthy. In fact, it is a sufficient condition for an agent's being blameworthy, on Arpaly's view, that an agent is responsive to sinister reasons when she acts. Thus,, while Arpaly thinks that an agent's performing a wrong act is a necessary condition for ascribing blame, she does not think that being morally indifferent is either a necessary or a sufficient condition for an agent's being blameworthy when he acts.

Now there is one question left. Does Arpaly think that it is a sufficient condition for ascribing blameworthiness to an agent that he was insufficiently responsive to the relevant moral reasons when he performed a wrong act? This looks like the right formulation of her view. First, the individual in question is assessed on whether a right or wrong act was performed, and then the individual in question is assessed on the reasons for which he actually acted. The beguiling grocer, for example, would be blameworthy on Arpaly's account. She would determine this by pointing to the fact that the beguiling grocer cheated his customers (he performed a wrong act) and then noting that when he cheated his customers he did not act with regard for the moral reasons in favor of not cheating his customers.

Support for this final interpretation of Arpaly's view can actually be found in the PRMR. Recall that Arpaly claims that an agent is praiseworthy for performing a right act, if

she has performed the right act for the relevant moral reasons. Similarly, Arpaly should think that an agent is blameworthy for performing a wrong act if she performs a wrong act with a deficiency of good will. I took the time to look at other interpretations of her view because I did not want to assume that Arpaly's view on blameworthiness was parallel to Arpaly's view on praiseworthiness.

If this is last interpretation of Arpaly's criteria for blameworthiness is the right one, then I have one further problem with her view. Depending on what insufficient responsiveness to relevant moral reasons amounts to, Arpaly will never be able to say that a morally indifferent agent is blameworthy when he performs a wrong act. This is the claim that I ultimately want to make, showing that our judgments about the blameworthiness of morally indifferent agents, on Arpaly's view and in the case of Praise-Blame Asymmetry, are frequently wrong.

If morally indifferent agents really do have moral reasons out in the world or in their heads that they should act on but choose not to act on, then it is possible that Arpaly is justified in assigning blame to these individuals when they perform wrong acts. I just think that this characterization of the morally indifferent agent is inadequate.

Seven

At this point, I have fleshed out several ways of trying to accommodate Praise-Blame Asymmetry. As far as I can tell, Arpaly's theory of moral worth is the closest I have come to identifying a view that can successfully accomplish this task. On her view, for an agent to be praiseworthy, she must perform the right act for the relevant moral reasons, and she is more praiseworthy the more attentive she is to the reasons in favor of acting morally. For an agent to be blameworthy, she must perform the morally wrong act and she must perform the wrong act for the reasons for which it is wrong or perform the wrong act while being insufficiently responsive to the relevant moral reasons. This view is consistent with Praise-Blame Asymmetry. Arpaly would claim that the prudent grocer is not praiseworthy and she would claim that the beguiling grocer is blameworthy. This is the response that laypeople are after.

Arpaly's conception of moral worth, with its promises of Praise-Blame Asymmetry, appears to have problems of its own. I am still left wondering whether we are justified in making judgments about the blameworthiness of agents like the beguiling grocer on her view. Arpaly's theory successfully accounts for what judgments people actually make about praiseworthiness and blameworthiness. It can distinguish, for example, on an intuitive level between each of the examples using Jeanne and Joseph. Her view does a good job of showing that we make these kinds of judgments, and even works as a good principle for how people make these judgments, but her view does not show how we are justified in making these judgments. While I am not certain that Arpaly's view will fail, I will take some time to express my concerns with her view. I do not aim for this section to be a definitive criticism of Arpaly's view, but rather a list of places to look for potential criticisms of Arpaly's view.

What leaves me questioning Arpaly's theory of moral worth, and the theory of moral worth that I have hitherto attributed to laypeople, is the change in strategy in the kinds of reasons that we attend to when we make judgments about agents like the beguiling grocer versus the prudent grocer, the good-willed grocer and the malevolent merchant. As Hindriks puts it, in the cases of the latter three grocers, we attend to the motivational reasons for which the agents act when they perform their right or wrong acts. When we ascribe praiseworthiness and blameworthiness to them, we care about the reasons for which they were actually acting. In the case of the beguiling grocer, we attend to the normative reasons for which we believe he is failing to act when he performs his wrong act. When we ascribe blameworthiness to him, we care about the reasons that we take to exist that we think the beguiling grocer is

failing to act on when he acts. We do not care about the reasons for which he actually acts. I think this change in strategy seems a little odd. I do not know that there is anything problematic with attending to normative versus motivational reasons, but I am not sure how it is justified, and it leaves me with the same question that I started off with. Why do we sometimes attend to an agent's motivations when we make judgments about praiseworthiness and blameworthiness and sometimes not?

Leftover Concerns for Arpaly

One question I have to ask Arpaly at this point is what she means when she claims that someone is insufficiently responsive to the relevant moral reasons. This is because her claim that an agent like the beguiling grocer is blameworthy hinges on this concept. When Arpaly, or anybody else, makes the claim that the beguiling grocer has reasons in favor of not cheating his customers, it is possible that she really means that the beguiling grocer has other goals, goals that are not profit-oriented, that are satisfied by his being fair to his customers. These are goals that he is ignoring or neglecting when he decides to cheat his customers. Arpaly would suggest that the beguiling grocer has morally relevant goals, like a desire not to hurt people or the desire to do what is right. This is to say that, if the beguiling grocer were aware of all of the reason providing motivational states in his mind, he would learn that he does, in fact have reasons for acting morally that he is not successfully responsive to. It is these moral reasons that the beguiling grocer is not acting on, these reasons that he is unresponsive to. In virtue of being unresponsive to these reasons, the beguiling grocer is blameworthy for neglecting them when he cheats his customers.

Here, I want to offer my first response to Arpaly's normative reason claim. It is possible that we are just wrong when we suggest that the beguiling grocer has reasons in favor of his not cheating his customers. His psychology could be structured such that treating his customers honestly is not a means to any end of his. We could be mistaken about what goals the beguiling grocer has. I have made the general stipulation that no matter how long the beguiling grocer deliberates or no matter how many times he acts, the beguiling grocer would never come to the realization that he has these reasons. If this is the case, then he just might not have these reasons, and he would be justified in questioning our claims.

If we know that no matter how long the beguiling grocer deliberated he would never come to the conclusion that he had a reason to be honest with his customers, then we are granting that it is possible that there is just nothing there to motivate him. If I were to insist

that the beguiling grocer had a reason to jump into a creek, and there was no goal of his that would be satisfied by his doing so, then there would be nothing to lead him to believe that this was true. It is possible that the beguiling grocer really just has no reason to jump into a creek.

Taking another look at Arpaly's view, an agent can be considered blameworthy for performing an act when she has been insufficiently responsive to moral reasons. Previously, we saw that the beguiling grocer was blameworthy because he was perceived as being insufficiently responsive to moral reasons that counted against his cheating his customers. If there just are no reasons for him to be fair to his customers, then he cannot possibly be insufficiently responsive to them. It is just as likely that the beguiling grocer is responsive to all of the reasons that exist that count for or against his acting. In this case, the beguiling grocer is a reason responsive agent who performs a wrong act and fails to meet the criteria for blameworthiness.

My second response to Arpaly's normative reason claim is that I can accept that the beguiling grocer has moral reasons in favor of his not cheating his customers, but I will demonstrate that he is not actually insufficiently responsive to these reasons. The beguiling grocer, in this case is not actually a morally indifferent agent. He has reasons for acting morally, and he is actually successfully responsive to them, but, from the third person perspective, he looks like he is indifferent to moral reasons because he does not perform the right act. I think that we would generally call this type of agent morally indifferent because of what his act looks like, but he is not indifferent in the sense that he does not take moral reasons under consideration when he acts; he just does not act for the moral reasons that he has.

I will remind the reader that being reason responsive entails responding to the reasons that one actually has. The beguiling grocer has reasons in favor of being honest with his customers and he has reasons in favor of his cheating his customers. This means that the beguiling grocer can have reasons for being honest with his customers and still have reason all things considered in favor of his cheating his customers. In fact, he can be responsive to all of the relevant moral reasons in favor of his not cheating his customers; he can give his reasons all of their due consideration and take them as seriously as the next person and still have more reason in favor of his cheating his customers.

This individual is not insufficiently responsive to moral reasons; he is responsive to moral reasons insofar as they function as reasons for him. If the beguiling grocer were to be honest with his customers, even though he has reason all things considered to cheat his

customers, we would have to think that he was acting irrationally. Even Arpaly thinks that a person acts rationally when he performs the act that he has greatest reason to do, and she thinks that it is possible for a person to have more reason to perform a wrong act than to perform a right act.[84] This means that Arpaly does not think that it is necessarily irrational for an agent to perform a wrong act.

Arpaly cannot expect a person to act irrationally in order to act morally – this seems incoherent. Further, if Arpaly is an internalist about reasons (as she claims), she cannot assume that a moral reason should motivate a person to act out of proportion to the amount that it is supported by the goals that that person has in his subjective motivational set. If the person in question is perfectly responsive to the moral reason that he has in favor of acting morally, but those reasons are not enough to motivate him to act morally, then he is not blameworthy because he is not insufficiently responsive to the moral reasons in favor of his performing the moral act.

In some ways this response to Arpaly is much like my first response, the suggestion that the beguiling grocer has no reasons in favor of his acting morally. I could change this to say that the beguiling grocer has no additional reasons in favor of his acting morally. The point here is that there are no relevant moral reasons that the beguiling grocer is failing to be responsive to when he cheats his customers. Thus, he cannot be considered insufficiently responsive to the relevant moral reasons.

Arpaly offers a case that might be similar to the two that I have described above. One of the agents that she talks about who acts out of moral indifference is called *Mary the half-moralist*, a young woman who is not sufficiently motivated to keep her promise.[85] Mary borrowed a book from her friend and promised to return it. However, as a result of some mix of her desire to stay home and write her book and bad weather, she does not return the book. Mary is aware of the fact that she promised to return the book, knows that promising to return the book has certain moral implications, and is still not motivated to return the book. I think the case of Mary the half-moralist could be easily explained by the two responses that I have already offered regarding the types of normative reason claims that Arpaly seems to want to be making. However, I might have to agree with Arpaly that, sometimes, this type of agent would have to be blameworthy. It seems possible that an agent could have means-end pairs that he was just not particularly aware of and fails to respond to the reasons that he has in favor of acting morally.

---

[84] Arpaly 37.
[85] Arpaly 82.

When asked how an agent, much like this last version of the beguiling grocer, could still be considered blameworthy for being insufficiently responsive to moral reasons even though it seems like he is as responsive to moral reasons as he could be, Arpaly responded by saying "from the moral point of view, he does not care about morality enough."[86] This claim has three possible interpretations. First, it is possible that Arpaly was saying that he should have found a way to be the kind of person who had more reasons for acting morally in his subjective motivational set. Second, it is possible that Arpaly was really making a depth of moral concern claim. This means that it is possible that Arpaly was suggesting not that the individual in question was insufficiently responsive to moral reasons, but that the agent was not motivated enough out of moral concern. I will make this distinction clearer in a moment. Third, it is possible that Arpaly really was making a reason claim, but that she was really making a different kind of reason claim than what I have been discussing so far.

The first interpretation of Arpaly's view seems unlikely, given that she denies the possibility of this kind of activity when she argues against autonomy theorists in her book. She states that people are less likely to undertake character-building activities than they are to keep a New Year's resolution, and she states that they should not be considered praise or blameworthy for having failed to cultivate the right moral psychology.[87]

The second interpretation of Arpaly's claim is more plausible. It is possible that Arpaly was saying that maybe the beguiling grocer did take all of his reasons into account in the right way, but he still did not demonstrate enough moral concern. There is a difference between responding to reasons and the extent to which those reasons actually motivate us on Arpaly's view. Recall the merely morally-inclined individual and the moral enthusiast. Both agents are responsive to moral reasons, but one agent has a greater amount of moral concern than the other. However, if Arpaly was making a depth of moral concern claim, then she cannot establish that the beguiling grocer is blameworthy. On Arpaly's view, depth of moral concern allows us to determine *how* praise or blameworthy and agent is, not *when* the agent is actually praiseworthy or blameworthy. Depth of moral concern allows us to determine the degree to which an agent is praiseworthy or blameworthy; it is not a necessary condition for ascribing praise or blame on Arpaly's view.[88]

The third interpretation of Arpaly's normative reason claim is that she is making claims about reasons that are not based in the beguiling grocer's psychology, meaning that

[86] Arpaly, Guest Lecture at Wheaton College.
[87] Arpaly 142.
[88] Arpaly 85.

she thinks that there are reasons out in the world that are currently not included among the beguiling grocer's aims and goals, and that she thinks that the beguiling grocer is failing to pick up and act for these reasons. On this view, things that serve as morally relevant reasons are the morally relevant features of a situation, like the fact that lying to my friend will cause him harm, or that taking the last cookie in the cookie jar will make my little sister cry. In Arpaly's example, an agent who is blameworthy for being morally indifferent suffers from a deficit in moral perceptions. This agent is blind to certain moral considerations and is, thus, unable to respond to them appropriately. This is to say that there are situations with morally relevant features that this agent simply fails to see.

Arpaly uses the example of an over-bearing parent to illustrate this type of morally indifferent agent.[89] Some parents, for example, are unable to respect the autonomy of their children because they simply cannot understand how the child could make decisions if someone were not there to tell the child what to decide. Further, no matter how many times we try to explain to the parent that the child can make decisions on her own, and no matter how many times we show the parent the problems with paternalism, the parent will just be unresponsive to these kinds of considerations and will continue to be bewildered by the request that she stay out of her child's business.

What confuses me about this kind of claim is that it simply seems that these kinds of reason claims would be inaccessible to the agent in question. If he cannot see the reason that he is failing to act on, then in what sense is he supposed to act on it? It may be true that he is not sufficiently responsive to the relevant moral reason, but I am wondering how Arpaly justifies saying that the agent is blameworthy. This might just be a case of moral luck, where the agent in question happens to have the bad lot of being an unperceptive person and, as a result, is an appropriate candidate for blame. I would be interested in hearing more on this.

I have dedicated this section to listing my concerns about Arpaly's view, but I have not been able to demonstrate that her view cannot support Praise-Blame Asymmetry, so one part of my task has been accomplished. I have been able to discover a theory of moral worth that successfully accommodates Praise-Blame Asymmetry, and one that is not so very different from what I think laypeople are aiming at. I am still left questioning how Arpaly justifies her engaging in Praise-Blame Asymmetry, and I have suggested some places to look for criticizing denying that her view is justified, but I have not raised any concerns that currently put her view at risk.

---

[89] Arpaly 82.

Conclusion

By appealing to reasons for acting and adopting Arpaly's theory of moral worth as a theory that can successfully accommodate Praise-Blame Asymmetry, I was hoping to resolve the questions that I posed at the beginning of this paper. As a reminder, these questions were about the lay tendency to sometimes assess an agent's praiseworthiness or blameworthiness based on an appeal to her motivations and sometimes not and how these judgments can be justified. Arpaly's view did a wonderful job of escaping the very obvious inconsistencies that I discovered in the initial iteration of the folk theory of moral worth. She also escaped the problems associated with the appeal to intentions and intentional action that I discovered while trying to clarify the vagueness associated with the intend/foresee distinction. I will remind the reader that the views that I touched on while trying to resolve the vagueness associated with the intend/foresee distinction resulted in backwards Praise-Blame Asymmetry; they got the case of the prudent grocer wrong.

Arpaly's view has none of these problems, but it does have one outstanding issue of its own. She does not show her reader how she is justified in engaging in Praise-Blame Asymmetry. Her view, just like the folk theory, sometimes appeals to the reasons for which an agent actually acts when she ascribes praiseworthiness and blameworthiness and sometimes appeals to reasons that she takes an agent to be failing to act for. Sometimes she appeals to the agent's motivations and sometimes she does not. It is not clear how this switch in what makes a person morally responsible for his act is justified.

Her view is not obviously wrong, or inconsistent, and it is a merit of her view that she accommodates the lay tendency to engage in Praise-Blame Asymmetry. These are all good signs for resolving the questions that I posed at the beginning of the paper. I think that Arpaly's view (and views like Arpaly's) is a good place to continue looking, while investigating the justifications behind Praise-Blame Asymmetry. I do, of course, have concerns about whether Arpaly, or anybody, can really justify Praise-Blame Asymmetry. I made note of these concerns throughout Section Seven, but I do not intend for them to be taken as my definitive criticisms of the view. I am uncertain about how to answer my original questions at this point.

I have determined that the folk theory of moral worth is too wrought with inconsistencies to appropriately allow laypeople to engage in Praise-Blame Asymmetry, and I have discovered that the inconsistencies in the folk theory of moral worth do not lay in the

vagueness of the intend/foresee distinction. I have determined that laypeople and the researchers who record their views have neglected a key concept in what we attend to when we make ascriptions of moral responsibility: reasons for acting. Reasons play a very prominent role in moral philosophy, and it seem that the appeal to reasons, as Arpaly has done, is the best bet for creating a view of moral worth that can support Praise-Blame Asymmetry. I have already noted that Arpaly's view had none of the obvious problems that the other views that I appealed to contained.

In the end, I am still left with the question of why laypeople sometimes assess agents for praiseworthiness and blameworthiness based on their motivations and why they sometimes do not, but I am much clearer on how they make the judgments that they do, how they go about trying to justify them, and, ultimately, where we can look in philosophy to remedy the inconsistencies in the folk theory of moral worth. While looking to Arpaly's view does not completely answer my question, she does go a long way towards creating a theory of moral worth that accommodates Praise-Blame Asymmetry without leading us to any obviously inconsistent judgments about the praiseworthiness or blameworthiness of the agents in question.

Works Cited

Alicke, M. “Culpable Causation.” *Journal of Personality and social Psychology*, 63 (1992): 368 – 378.

– – –. “Culpable Control and the Psychology of Blame.” *Psychological Bulletin*, 126 (2000): 556 – 574.

– – –. "Blaming Badly." *Journal of Cognition and Culture*, 8 (2008): 179 – 186

Anscombe, G. E. M. *Intention*. USA: Harvard University Press, 2000.

Arpaly, N. *Unprincipled Virtue*. Oxford: Oxford University Press, 2003.

– – –. Wheaton College. Austin House, 26 East Main Street, Norton MA. 03 May 2010. Guest Lecture.

Bennett, J. “Morality and Consequences,” *The Tanner Lectures on Human Values II*, Salt Lake City: University of Utah Press, 1981: 45 – 116.

Bratman, M. "Two Faces of Intention." *The Philosophical Review*, 93.3 (1984): 375 – 405.

Burns, K. J. D., & Bechara, A. “Decision Making and Free Will: A Neuroscience Perspective,” *Behavioral Sciences and the Law*, 25 (2007): 263–280.

Cushman, F. “Crime and punishment: Distinguishing the roles of causal and intentional analysis in moral judgment.” *Cognition*, 180 (2008): 353 – 380.

Davidson, D. “Problems in the Explanation of Action.” *Problems of Rationality*. Oxford: Oxford University Press, 2004.

Fitzpatrick, W. J. “The Intend/Foresee Distinction and the Problem of “Closeness”.” *Philosophical Studies*. 128 (2006): 585 – 617.

Foot, P. “The Problem of Abortion and the Doctrine of Double Effect.” *Virtues and Vices, and Other Essays in Moral Philosophy*. Berkeley: University of California Press, 1978.

Harman, G. H. “Inference to the Best Explanation.” *Philosophical Review*, 74.1 (1965): 88 – 95.

Hindriks, F. "Intentional Action and Praise-Blame Asymmetry." *The Philosophical Quarterly*, 58.233 (2008): 630 – 641.

Kant, I. *Grounding for the Metaphysics of Morals*. Indianapolis: Hackett Publishing Company, 1993.

Knobe, J. “Intentional Action and Side Effects in Ordinary Language.” *Analysis*, 64 (2003): 81 – 87.

Korsgaard, C. “Skepticism about Practical Reason.” *Journal of Philosophy*, 83 (1986): 5-25.

Malle, B. F., & Knobe, J. "The Folk Concept of Intentionality." *Journal of Experimental Social Psychology*, 33 (1997): 101 – 121.

Malory, T. *Le Morte d’Arthur*. Digireads.com Publishing, 2009.

Nadelhofer, T. "Blame, Badness, and Intentional Action: A Reply to Knobe and Mendlow." *Journal of Theoretical and Philosophical Psychology*, 24.2 (2004): 259 – 268.

Ohtsubo, Y. “Perceived Intentionality Intensifies Blameworthiness of Negative Behaviors: Blame Praise Asymmetry in Intensification Effect.” *Japanese Psychological Research*, 49.2 (2007): 100 – 110.

Quinn, W. “Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing.” *Morality and Action*. Cambridge: Cambridge University Press, 1993.

Sidgwick, H. *The Methods of Ethics*. Indianapolis: Hacket Publishing Company, 1981.

Mark, T. *Adventures of Huckleberry Finn*. USA: Quill Pen Classics, 2008.

Williams, B. “Internal and External Reasons.” Reprinted in *Moral Luck*. Cambridge: Cambridge University Press, 1981.

Williams, B. & Nagel, T. “Moral Luck.” *Proceedings of the Aristotelian Society, Supplementary Volumes*, 50 (1976): 115 – 135 & 137 – 151. .

Wolf, S. *Freedom Within Reason*. New York, NY: Oxford University Press, 1990.

Wright, J., & Bengson, J. "Asymmetries in Judgments of Responsibility and Intentional Action." *Mind and Language*, 24.1 (2009): 24 – 50.