

Decent Humans: What We Can Learn From Moral Exemplars
By
Sebastian Theodor Betzer

A thesis submitted in partial fulfillment
of the requirements for the degree of
Bachelor of Arts with Honors
Department of Philosophy
at the University of Michigan
2018

Advisor: Professor Sarah Buss

Second Reader: Assistant Professor Meena Krishnamurthy

Acknowledgements

Firstly, I want to express my profound gratitude to Professor Sarah Buss for advising this thesis. I am grateful that she expected nothing less than the best from me. My writing and thinking have vastly improved as a result of her guidance.

Secondly, I would like to thank the incredible professors and GSIs and my peers in the philosophy department at the University of Michigan who have always encouraged my love of philosophy and my subsequent desire to write this thesis.

I would also like to thank my Mum for her loving support and acknowledge her help in finding many of the opportunities that inspired me to ask the central questions of this thesis.

I would also like to acknowledge David Cadman for his encouragement, willingness to spend many hours discussing and challenging my ideas, and for introducing me to many inspirational people working on exemplary projects.

I am grateful to my father for the support which enabled me to focus as much time as possible on this thesis.

Finally, I must express my gratitude to my fabulous girlfriend Katrina and to my close friends who have all provided me with support and encouragement during the writing of this thesis. I am lucky to have friends that will put up with someone continually proposing thought experiments and testing their intuitions. They have encouraged me to look for heroes in unexpected places and occasionally provided me with some appropriately helpful skepticism. For that, I am forever grateful.

Chapter 1. Defining Moral Exemplars

In the course of our lives, we encounter many people who we might consider exemplary. We often refer to these people as our heroes. Heroes are a source of inspiration for many; they motivate some, and lead others to reimagine what is possible for human beings. We often feel a sense of wonder toward the achievements of our heroes. This wonder is manifest in the stories that we tell about them, and our reactions to observing what they do. “Hero” is a general term for a wide variety of exemplars. We might recognize exemplars in prestigious arenas such as sports, politics, in literature, and so on, but we also recognize heroes that are closer to home; community leaders, activists, service workers, all have the ability to become heroes too. The variety and abundance of heroes in the world suggests that there is much we can learn from analyzing intuitions about what it takes to be various kinds of hero. Given that we take inspiration from our heroes, it is worth thinking about what it might take for us to be like them.

Section I: Introduction to Exemplars

The athletic hero is one of the most readily identifiable, and publicized hero types. Athletic heroes are characterized by exceptional feats of athleticism. Among these athletes, success in crucial moments and any significant obstacles they overcome in the course of their training influence the degree of praise that they are due. For example, the most heroic athletes are those that perform best when it counts the most. The Olympic Games highlight the degrees of praiseworthiness befitting for an athletic hero; all of the qualifiers are exemplary in an athletic sense, but their successful performance under great pressure, differentiates them from one another. Athletic exemplars also become heroes by overcoming significant barriers to their success. We might think that the significance of the challenges an athlete overcomes is related to

their praiseworthiness. In sum, the athletic hero is set apart from ordinary people by her extraordinary athleticism, successful performance, and/or challenges and sacrifices she successfully faces.

When we encounter these heroes in the news, in literature, or on film, it is natural to wonder what exactly sets them apart from us. Generally, there are two ways of answering this question: talent or practice. One might think that an athlete is essentially just like the rest of us and is only set apart by her intense training. Had we chosen the same goal as the athlete, all that we would need to do is match her training intensity, and the rest of us could see very similar results regarding athletic ability. On the other hand, talent might set these agents apart. On the talent view, athletic exemplars are blessed with unique features that cannot be gained through training; no amount of training could enable an ordinary person to achieve the level of performance that an athletic hero does—we would always fall short. Of course, in reality it is never the case of an either or. All the talent in the world is not sufficient for exemplary athletic achievement in the absence of practice. The challenge here however, is determining whether talent is a required feature. When ordinary people consider this in the case of the athletic exemplar, it is difficult to say precisely what sets her apart because the two claims are counterfactual. Each requires an individual observer to imagine what would have happened had she dedicated her own life to a particular athletic endeavor. Nevertheless these heroes often inspire people to maintain an active lifestyle or be resilient in the face of personal adversity.

However, athletes are far from the only kind of hero we might turn to for inspiration. Saviors are a different sort of hero who are praiseworthy for very different reasons. This hero is an agent who intervenes in a dangerous situation to rescue another person from immediate

danger. In the US, The Carnegie Heroes Medal is an award that highlights these feats of heroism each year. This award is presented to civilians, who leave a condition of safety and risk injury or death to rescue another human being.¹ A winner of this award and paradigmatic example of the savior hero is Lora Shrake. During her drive home from a friend's house, she saw from her car window a woman in a nearby field injured, trapped and getting mauled by a 950-pound bull. Shrake stopped her car immediately, climbed through an electric fence, and proceeded to hit the bull repeatedly, distracting it until the victim could crawl to safety. Shrake, a 21-year old college student had no prior experience with farm animals, and no prior training that would make her especially strong, fast, or brave. She saw someone in a life-threatening situation, and according to her, "Here is a problem, here's what I need to do, and something needed to happen... I didn't really take the time to think about what else could happen."²

Describing how the savior hero differs from "ordinary" individuals is more challenging than in the case of athletes, because the relevant considerations are substantially different. For instance, this status frequently depends at least in part on situational factors beyond the agent's control. There are many cases, in which the agent may not have time to deliberate before they act. For example, when Anthony Sadler, Alek Skarlatos, and Spencer Stone were on the train to Paris that became the target of a terrorist attack, they could not have known beforehand that the horrifying situation would arise. However, the three of them sprang into action in the face of danger, successfully subdued the attacker, and saved over 500 lives.³ Since circumstances play a

¹ The Carnegie Heroes Fund Commission. (n.d.). Requirements for a Carnegie Medal. Retrieved from <http://www.carnegiehero.org/about-the-fund/mission/>

² Levy, L., & Howard, T. (2018, Jan 9). *How to Be a Hero*. Retrieved from <http://www.radiolab.org/story/how-be-hero/>

³ Chrisafis, A. (2015, Aug 22). France Train Attack: Americans overpower gunman on Paris express. *The Guardian*. Retrieved from <https://www.theguardian.com/world/2015/aug/21/amsterdam-paris-train-gunman-france>

significant role in this kind of exemplary action, it is especially difficult to ascribe particular characteristics to these agents. It might be the case that saviors also have exceptional character such as bravery, or possess special traits such as quick thinking, but these are not necessary conditions. Even psychologists who study these exemplars are unsure about what exactly motivates such heroic action.⁴ In the face of this uncertainty, the defining feature of savior heroes is the outcome of their action—saving lives in dangerous situations.

It is worth noting however, that we do have some intuitive sense that a savior who acts for moral reasons is more praiseworthy than one who does not. Contrast a savior who acts with an eye toward fame and headlines with a savior whose motivation for running into a burning building just is the fact that there are people inside who are in danger. While it is the case that both should be considered saviors on the definition that I presented, it seems as though the agent with apparently selfless motivations is worthy of a different kind of praise.

In order to explore this moral intuition more deeply, it is helpful to analyze yet another category of hero—the moral exemplar. The moral exemplar is characterized by exemplary moral dispositions connected with her extraordinary actions. The characteristic actions of a moral exemplar might overlap with other kinds of heroes. In fact, moral exemplars' range of exemplary actions can include some of those acts that place agents into the category of saviors—specifically those in which there is time to engage in moral deliberation prior to acting. However, performing these acts is neither necessary nor sufficient for being a moral exemplar. A moral exemplar must have a willingness to make substantial sacrifices in pursuit of moral objectives, where the willingness reflects a stable moral disposition. A moral exemplar is not a moral saint; it is not

⁴ Dr. Robert Sapolski on Radiolab

necessary that a moral exemplar be morally flawless. An agent might be a moral exemplar in regard to a particular praiseworthy act, but ordinary or worthy of reprimand in other respects. To illustrate this definition, I will recount two stories of paradigmatic moral exemplars; Paul Farmer and Bob Moses.

Paul Farmer dedicates himself to providing medical care to people living in rural and under-resourced areas in developing nations.⁵ The sacrifices that Paul Farmer makes in pursuit of this goal are undeniably great. He founded the social justice organization Partners In Health and in the early years of the organization, he spent his time away from his family, sleeping in the office in Boston or staying near his clinic in Haiti. He went long periods without seeing his wife or daughter, did not buy anything beyond his minimal needs, did not take any vacations, and deposited his salary directly back into the organization. While observers might think that these sacrifices are very significant and that accepting them would feel especially unbearable, Farmer displays a generally upbeat disposition and constantly feels as though he should be doing more.

Bob Moses is an activist who is extensively involved in the struggle for civil rights in the US.⁶ As one of the leaders of SNCC (Student Nonviolent Coordinating Committee) and an organizer of the Freedom Summer voter registration movement, Moses faced extreme danger in the face of his actions in pursuit of moral commitments. He gave up a relatively comfortable middle class life to move to Mississippi and endure constant physical and emotional violence. During the Freedom Summer, a movement to register black voters despite violent efforts to keep them disenfranchised, he had to make decisions that concerned the well-being of others and take responsibility for the violence that often befell them. Taking on the emotional labor of this

⁵ Carbonell, V. (2012). The Ratcheting-Up Effect. *Pacific Philosophical Quarterly*, 93, 228–254.

⁶ Neiman, S. (2009). Chapter 12: Theory and Practice. In *Moral clarity* (pp. 372–414). Princeton: Princeton Univ. Press.

responsibility is a sacrifice that most ordinary people could not imagine taking on. However, Moses was deeply influenced by his understanding of arguments about race and morality, and persisted in the face of these immense challenges.

Paul Farmer and Bob Moses are only two of countless examples of moral heroism. One key feature of their work is that it involves sacrifices that outside observers might think is too much to bear. However, the exemplars recognize the sacrifices which stem from their moral commitments and proceed despite these challenges. Stories about these heroes captivate the imagination in part because of our curiosity regarding the exemplars' ability to act in such morally praiseworthy ways. We analyze this kind of hero in film, literature, and on the news. In these constructed mediums, the author presents a number of facts about the exemplar and invites us to assess the quality of her moral disposition as it relates to the particular circumstances of her action. These literary representations are often designed to provide the rest of us with moral information, along with some inspiration for striving towards emulating these characteristics. In this thesis, I am specifically focused on moral exemplars. I will argue that moral exemplars are not fundamentally different from ordinary people; their morally praiseworthy acts are possible for us too. I go on to argue that the similarities between us and moral exemplars suggest that ordinary people could and should be doing more to fulfill obligations that we have to provide aid to others; our obligations exist even when they are not recognized by the moral status quo. Although ordinary people may have justified beliefs that they are not obligated to do more, we can learn that these beliefs are false by observing the actions of moral exemplars. Ultimately, I argue that even if we have an excuse for failing to meet obligations, our exposure to moral exemplars gives us reason to question justified moral beliefs. In fact, I argue that we have further

obligations to seek out and gain inspiration from moral exemplars in order to improve the moral status quo.

Section II: The Talent and Practice Views of Exemplary Agents

In what follows, I will consider the moral status of moral exemplars and their actions. In particular, I want to explore whether their moral status differs significantly from ordinary people; do their exemplary actions exceed what duty can require? I will consider some reasons to think that they do; but after raising objections to this view, I will then turn to considering what follows if we accept moral exemplars' claims that they are merely doing what they ought to do.

A promising approach to evaluating the moral status of moral exemplars is to compare them with ordinary people. It is helpful to recall the arguments about the athlete when thinking about how to make this comparison. Recall that with the athlete, we have competing intuitions about what sets them apart. There is a view on which they have innate qualities and talents that ordinary people do not (which I will call the "talent view"). There is also an argument that exemplars are very similar to ordinary people in relevant respects, and that training and dedication set them apart (I will call this the "practice view"). These two claims appear to have analogues in the case of moral exemplars.

The "talent view" for moral exemplars holds that a moral exemplar is deeply different from the rest of us in relevant moral respects. This particular view is similar to superhero narratives, or ancient myths, in which heroes are gifted by the gods with particular skills they need to overcome great challenges. In the case of moral exemplars, these special traits might include: unique psychology contributing to unusually high tolerance for sacrifice, special

capacities for empathetic feeling, and so on. If we assume that the talent view is accurate, what might the status of their moral actions be? There seem to be two possible options.

A proponent of the “talent view” might claim that moral exemplars exceed obligations. The argument that it is possible for praiseworthy actions to exceed obligation is attributed to J.O. Urmson in his article entitled “Saints and Heroes.” The traditional categorization of acts as impermissible, permissible, and obligatory do not, Urmson argues, correctly recognize acts that are morally praiseworthy but not required.⁷ It seems like heroic actions are too demanding to require ordinary people to do, but it would be inadequate to simply say that saints and heroes act merely permissibly. Surely, they are worthy of moral praise, which the broad category of permissibility does not imply. To account for these intuitions, he argues that we should add a further category of action called, “the supererogatory”. If we accept this view, then the moral exemplar just is someone who performs supererogatory acts. On this view, when ordinary people observe the moral exemplar, moral content is not conveyed; observers do not learn about their own obligations, but they might nonetheless be inspired by the apparent praiseworthiness of the exemplar’s acts.

Supererogation is not the only possibility for evaluating the status of moral exemplars’ actions available to someone endorsing the “talent view.” They might also claim that both moral exemplars and ordinary people are acting in accordance with obligation, but moral exemplars face more demanding obligations on account of their unique traits. This view is supported by the “ought implies can” principle of moral obligation. Given that obligation implies possibility, it follows that we cannot be obligated to do things that are impossible. If moral exemplars are

⁷ Urmson, J. O. (1969). Saints and Heroes. Joel Feinberg (Ed.), *Moral Concepts* (pp. 60–73). Oxford University Press.

fundamentally different from ordinary people in a way that makes it uniquely possible for them to do more good than others, then insofar as they ought to do good, they are obligated to perform more challenging feats than what others are capable of. Likewise, if ordinary people do lack the required psychology to perform the actions that the exemplar does, they cannot be obligated to perform these acts. A weaker conception of what it takes for an action to be “impossible” might suggest that ordinary people still cannot be obligated to do what moral exemplars do. The fundamental “talent” central to this weaker conception is the exemplars’ strong moral dispositions. If this difference is truly intractable, the argument goes as follows; Since ordinary people are not able to achieve the exceptional moral disposition of exemplars, our obligations depend on the quality of moral disposition we can reasonably expect ordinary people to attain.

However, the “practice view” is a persuasive alternative to the “talent view.” This particular view forms the basis of the iconic narrative of the “Hero’s Journey,” in which a hero begins a quest and as a result of various experiences learns to overcome her own fallibilities to achieve a good outcome.⁸ On this view, the moral exemplar is just like the rest of us in relevant respects. She is not marked by extraordinary psychology that makes her especially disposed to empathetic emotions, she does not have any special resilience to feelings of loss from sacrifices, and so on. She is set apart by practice and experience. Moral practice could be a process of reflection, an experience that makes a rational agent recognize that moral principles apply to more situations than she previously thought, and so on. If we hold that the “practice view” of moral exemplars is true, implying that moral exemplars are under the same constraints as the rest of us regarding moral conduct, what does this mean for the moral status of their actions? This

⁸ Campbell, J. (1993). The Hero Today. In *The hero with a thousand faces* (pp. 387–391). London: Fontana.

question is the same as the one raised for the “talent view,” and once again there seem to be two possibilities.

One option for evaluating the moral status of exemplary actions compatible with the “practice view” is that the exemplar’s actions are supererogatory. Recall that this means that her actions are morally praiseworthy but not required. The “practice view” gives an account of the exemplar that makes her relevantly similar to the rest of us, with the exception of some additional moral training. This training might result in the motivation for the exemplars’ actions, but it does not generate new obligations. If we endorse this account, we might think that the additional moral training that sets exemplars apart is praiseworthy but optional. Ordinary people might be inspired by the moral exemplar to pursue moral training in their own lives, but this inspiration would not require that they actually accept more demanding obligations.

The other possibility compatible with the “practice view”, is that the demanding actions that characterize exemplars are obligatory. On this view, the exemplar is doing exactly what duty requires of her. This means that ordinary people fail to meet obligations; each of us could, and should, be doing more. Moral exemplars are able to determine what they are obligated to do and successfully discharge these duties through practice. An ordinary person who observes the moral exemplar observes what obligation requires, and yet she fails to make the relevant choices with the proper disposition in her own life. This might be explained by the fact that a stable moral disposition takes experience and reflection to develop and most people do not use their time for these forms of moral practice. However, even if we are able to explain it, this view suggests that this failure to work on improving our moral sensibilities is morally wrong. In other words, the

failure to develop the disposition to make relevant sacrifices involves failing to do something one is obligated to do—namely, engage in moral development.

Section III: Evaluating the Options

Canvassing options that correspond to each theory about the distinguishing factor that explains the difference between moral exemplars and ordinary people reveals a range of possible moral evaluations. I will evaluate each of these in order to determine which of these conflicting views we should accept.

Classifying the moral exemplars' actions as supererogatory is an option that is available to both conceptions of what makes an exemplar and might seem promising for this reason. This classification supports the intuition that these agents are exemplary; whether or not we could be like them, their actions are especially praiseworthy.

One strong reason supporting the claim that moral exemplars' actions are supererogatory, is that moral exemplars make significant sacrifices that seem unreasonable to expect people to make. An unreasonable sacrifice involves incurring a particularly significant harm or a loss of particular goods. Paradigmatic examples of unreasonable sacrifice include loss of life and suffering severe injury. Without a moral imperative, it seems unreasonable to expect me to sacrifice my own life to save a drowning child. However, not all sacrifices are unreasonable; it would be unreasonable for me to think that I was not obligated to save the child if doing so would require some smaller sacrifice like the destruction of an expensive new suit I was wearing.

As a matter of fact, it is generally quite difficult to define how reasonable or unreasonable a sacrifice is because there are many factors that alter how they are perceived and what their actual impacts on agents are. Some moral theories require more demanding sacrifices. Act

consequentialism is one such moral theory. For an act consequentialist, an obligation to help others who are suffering requires taking on a more helpful career and relocating to a place in the world where one could bring about the most good possible. However, this would involve taking on a large sacrifice of the potential goods that a less helpful career, or living near one's family rather than relocating, might bring. Many people might reject act consequentialism simply because they might feel that the demands involved with accepting it are unreasonable to expect. If the sacrifice is really unreasonable to expect someone to accept, we might think that one would not be failing an obligation to aid by omitting those actions that involve such sacrifices.

In fact, we consider some sacrifices to be unreasonable even when accepting them would avoid harm to others. Applying to university can be construed as one such occurrence. The potential sacrifice of a benefit (getting into a good university, and the benefits that come from that) from not applying reasonably defeats the obligation not to harm the person who will not get in if you are accepted (not getting into a university they preferred, and missing out on the possible benefits that could have resulted from that). The same is true when accepting a scholarship even when you can reasonably afford tuition. This prevents someone with greater need from attaining this benefit, but it still might feel unreasonable to some to be expected to let the opportunity pass by. In fact, most systems that select how to allocate scarce resources generate this type of harm.

Consumption habits are another normalized system which generates harms. It is a fact that humans cause harm to others and the environment as a part of many contemporary consumption practices. Manufacturing the products we use, as well as the waste that we produce, has harmful impacts on humans and the environment that we could lessen by changing our

consumption habits. However changing these habits might also be considered too great a sacrifice to reasonably expect someone to make. The costs associated with changing to a less harmful lifestyle might be too demanding to expect people to take on, despite it being morally praiseworthy.

Accepting the claim that moral exemplars' actions are supererogatory suggests that moral exemplars take on sacrifices that ordinary people would take to be unreasonable. The claim implies that ordinary people have no moral imperative to try taking on greater sacrifices than they have reason to believe they can bear. I gave reasons to think that ordinary people often consider the costs of changing lifestyle habits or challenging problematic systems that benefit them to be particularly unreasonable. This might lead us to think that accepting the status quo involves accepting some obligations, but considering most efforts to change the status quo praiseworthy but optional.

Section IV: The Limits of Supererogation

We should abandon the supererogation view in the case of moral exemplars. One reason for this is that among morally praiseworthy actions, the difficulty of an action is not an appropriate consideration to determine what is obligatory and what is not. Many actions that are obligatory require some form of sacrifice, even if this sacrifice is just the time that might be spent doing something else. Recall that the kind of sacrifice I am referring to here are sacrifices of well-being rather than physical harms. What should the test be to determine how much sacrifice is too much? We might think that we should turn to the intuitions of ordinary people with moderate commitments to moral views to define such a test. Shelly Kagan characterizes the

“Moderate,” as an individual who thinks that moral obligation should not be so demanding. The Moderate might claim that obligations cannot require that we sacrifice our life projects.⁹

A project, as I am using it here, is a series of actions undertaken over some period of time that together serve a coherent goal motivated by an agent’s interests. Bernard Williams argues that a project is made up of those interests with which an agent is centrally concerned. Agents tend to live their lives pursuing various interests which take the form of careers, personal decisions, hobbies, familial relationships, and so on.¹⁰ Moral exemplars are set apart from ordinary people because their chosen project is directed at bringing about moral good (such as joining Doctors Without Borders) over less aid-focused projects (dedicating oneself to studying the philosophy of mathematics) when the options are mutually exclusive. Kagan’s Moderate appeals for what he calls “obligation with options.” The rough idea is that what an agent is obligated to do is dependent on her interests.¹¹ An example of this view in practice might be that I am obligated to donate a portion of my income to charity provided the amount that I am donating does not impede me in my pursuit of my personal projects. One challenge this view faces is laying out explicitly what counts as a reasonable project. For example, an agent might say that her interests motivate her goal to advance in her chosen career. The actions that her career requires constitute the project that generates the options for her to accept less significant moral demands. However, another person might take the well-being of her family to be her central project. One might challenge the “Obligation with options” view because it is unclear how the moderate would adjudicate what counts as an unreasonable sacrifice between diverse agents with very different projects. While this does not necessarily threaten the view altogether,

⁹ Kagan, S. (2002). Chapters 7-10. In *The Limits of Morality* (Reprinted, pp. 231–403). Oxford: Clarendon Press.

¹⁰ Williams, B. (1981). *Persons, Character, and Morality*. In *Moral Luck* (pp. 1–19). Cambridge University Press.

¹¹ Kagan 258-262

projects, according to this definition, seem to be a product of an agent's wants. If we ground important moral considerations in mere wants, obligations also seem to become a product of individual agents' wants. If it were the case that each person could modify the demands of obligation to suit their wants, morality would be a poor guide to action.

Consequently, if we accept the supererogation view of moral exemplars, we might think that the difference between the exemplar and an ordinary person is a mere choice of projects. However, this particular view is unacceptable if we are committed to the idea that morality ought to be a meaningful guide to action, because it provides no guidance concerning how to choose one of the most central aspects of our lives—our projects.

The other reason that we might be skeptical of the supererogation view, is the testimony of exemplars themselves. Moral exemplars frequently report that their actions are obligatory. That they had no other choice but to act the way they did. One might think that this is a mere expression of modesty, or have something to do with motivation, but I think that we should take this testimony more seriously. If we are to consider these agents moral exemplars and think that they impart valuable moral information and inspiration, it seems strange to suggest that they might be wrong about themselves. A moral exemplar is characterized by a stable moral disposition, meaning that she has consistent moral reasons for acting. These might be explicit reasons as in cases where an agent acts because she believes a particular action is obligatory. On the other hand, the disposition of an exemplar might also be based on implicit reasoning as in cases where an agent acts from general morally praiseworthy values like generosity. In either case, what matters is that the moral exemplar does not achieve this status by accident. If their

action is the result of luck, then it would be wrong to call them a moral exemplar. Such an agent could be wrong about what obligation requires, and would not be an appropriate example.

Lastly, another reason to be skeptical of supererogation is it seems to assert that moral exemplars are praiseworthy specifically because they are not required to do what they are doing. The argument is supported by an implicit claim that fulfilling obligations is not praiseworthy due to being required. This intuition is supported by the observation that in everyday life; we generally do not expect praise for fulfilling a requirement. It would be quite odd to expect praise after achieving the bare minimum requirements to secure a passing grade in a class, or to expect to be considered someone's best friend after only doing the minimum required for being considered a friend at all. These observations might lead someone to conclude that perhaps moral requirements are the same; if something is obligatory then it is not worthy of moral praise. Our moral praise of exemplars must mean that their actions are not required. However, we should consider cases of perseverance a counterexample to this claim. When situations make moral requirements more costly to carry out, it would be appropriate to recognize the cost an agent took on to achieve what was expected of them. For example, an agent might recognize an obligation to uphold a promise to a friend. Under normal circumstances, fulfilling this obligation is not especially praiseworthy. On the other hand, if the situation the agent finds herself in raises the cost of upholding the promise especially high, then staying on course to fulfill this obligation is much more difficult and therefore praiseworthy if she succeeds. Similarly, securing a passing grade in a class after only achieving the minimum requirements may well be praiseworthy for an agent with learning difficulties. In cases like this, it does not seem as though obligation changes, but rather, that further factors impacting an agent are important for being exemplary. When an

agent is faced with factors that require perseverance, we do not think that the obligation has changed. Rather, it seems as though the agent recognizes the importance of her obligation and uses this as motivation. This conclusion should make us reluctant to consider demanding actions supererogatory merely on the basis of their praiseworthiness.

Although these are not decisive reasons for rejecting the supererogation view altogether, they do constitute a good reason to explore what the moral status of moral exemplars and ordinary people is if we accept the hypothesis that exemplars' actions are not beyond the call of duty. Our task then, concerns the two remaining possibilities for evaluating the moral status of actions. Either ordinary people are failing to meet obligation most of the time, or ordinary people have a different set of obligations from moral exemplars. We have already established that this depends in large part on our belief about how moral exemplars differ from the rest of us. If they are fundamentally different then they should accordingly have different, more demanding obligations. If they are like the rest of us, then we ordinary people are failing to meet our obligations.

Section V: Arguments for the Practice View

In the remainder of this chapter, I will argue that there are good reasons to think that moral exemplars are not deeply different from the rest of us.

Firstly, a reason to take the position that moral exemplars are relevantly similar to the rest of us is simply to take their word for it. In addition to the assertions that their actions are not supererogatory, moral exemplars also reject the notion that they are very different from the rest of us. This should not be dismissed as mere modesty. Prior to their morally exemplary action, we would assume that these agents have hopes and dreams, spend time grappling with significant

and not-so-significant challenges, make mistakes, and the whole range of human experiences that we assume others have when we meet someone new. We should find it absurd to think that learning that an agent performed a morally exemplary action means we should drop these other assumptions about her being an ordinary person.

Secondly, moral exemplars change disposition with respect to moral issues over time, just like the rest of us. For ordinary people, the process of moral learning is one of continuous trial and error along with a certain degree of reflection. One way moral exemplars might differ from us, according to a supporter of the “talent view”, is that they might have a much stronger appreciation of moral facts than ordinary people. If this were true, we would not expect to see drastic changes in moral disposition on various issues over time. However, it is often a change of heart that motivates an agent to make costly sacrifices in pursuit of obligations that marks them as exemplary. Daniel Ellsberg, the publisher of the Pentagon Papers is an illustrative example.¹² Ellsberg began his career in service to the US government with full commitment to the ethics espoused by the political establishment. He even initially supported the military and espionage related policies that he would later come to take a strong stance against. For Ellsberg, the horrors of the Vietnam war, and the dawning realization about how little the American public were informed about, lead him to question the morality of the governments practices. This questioning alone was enough to invite ridicule from his government colleagues. However, Ellsberg accepted far greater sacrifice in doing the right thing; Ellsberg sacrificed his job, potential legacy, and risked being jailed for life when he decided to copy and release the top secret information contained within the Pentagon Papers to the public. Ellsberg is a moral

¹² Neiman 393-414

exemplar due to the motivation to carry out his obligations, but his process of seeking answers to contentious moral questions is one that is shared by ordinary people and moral exemplars alike. In sum, Ellsberg is not set apart from ordinary people as a result of some particular innate talent, but rather a commitment to moral practice that most people are capable of carrying out.

A proponent of the “talent view” might also contend that the moral exemplar could differ from us as a result of some psychological feature that makes her more resistant to sacrifice. Whether that is a result of unusually high tolerance for sacrifice or a failure to recognize sacrifices, the general idea is that she might differ from ordinary people with regard to what she can consider reasonable sacrifice. A failure to recognize sacrifice and an unusually high tolerance for sacrifice would both have implications for what the exemplar would accept as reasonable sacrifice relative to ordinary people. However, we should reject this view.

In fact, it seems necessary for being a moral exemplar that the relevant perception of sacrifice is accessible to the agent. Vanessa Carbonell provides an account of a woman named Susan Tom that illustrates that moral exemplars are able to recognize the significance of their sacrifices.¹³ Tom, a single mother who adopts terminally ill children in order to give them loving and supportive end of life care, is not numb to the pain of their eventual passing, or ignorant of the financial hardships that come with her moral project. Instead, she takes on these challenges because they arise from her belief that she has an obligation to give aid to these children who would likely suffer needlessly otherwise.

A supporter of the talent view might argue that while Tom recognizes the sacrifices she takes on, perhaps she gains a great deal of satisfaction from her projects which makes the

¹³ Carbonell 229-230

sacrifice bearable. Tom does not deny that she gains satisfaction and happiness from caring for these children. In fact she talks fondly of the many happy moments that she and the children share together. The proponent of the “talent view” is interested in whether these moments of positivity outweigh the sacrifices for Tom, due to her feelings of empathy, or some other feature that is unique to her. In other words, do the sacrifices and the benefits confer upon Tom a “net good”? The worry is that the significance of the sacrifice Tom takes on is compensated by benefits of her actions; these benefits might not have the same effect on another person’s well-being and therefore we might conclude that Tom is set apart by a special trait after all. Determining the subjective experience of sacrifice before and after experiencing it would likely require the work of psychologists. However, we should not think of sacrifice in terms of net effects on an individual. Altruism confers many benefits on those agents that exhibit these practices. Despite this, it seems as though what they give up is significant. In fact, it would undermine the very nature of altruism if an agent only gave at levels that were insignificant to her. Attempting to calculate the net change in well-being that accepting a sacrifice causes an agent fails to appreciate the significance of an agent's decision to accept sacrifice in the first place. Carbonell argues that we ought to think of sacrifices in terms of "gross losses of well-being." A "gross loss" is the immediate cost of making a sacrifice. I will analyze her argument more closely in the following chapter, but the intuition seems to provide evidence against the talent view. In essence, evaluating the significance of sacrifice in this way allows us to consider the relevant sacrifices in the same epistemic context that the exemplar does. Net evaluations involve knowledge of more factors than exemplars themselves perceive when they

initially make sacrifices. Accordingly, we should reject this argument in favor of the practice view.

Still another reason to think that moral exemplars are like the rest of us is that exemplars' actions are broadly intelligible. They are intelligible in the sense that ordinary people can observe a moral exemplar and comprehend her motivations and moral praiseworthiness. If an exemplar were fundamentally different, this comprehension would break down. In fact, were the difference stark enough, it would be difficult for ordinary people to make relevant moral judgements. We could determine whether or not the outcomes of her actions had a positive effect, but we could not judge her moral disposition. In order to judge her disposition, we must understand the considerations that an agent makes, and the challenges she takes on. When we understand these two features, we can consider the praiseworthiness of an agent's action. Moral exemplars give us intelligible narratives about their moral development over time in a way that makes sense to us. We should take this as a consideration in support of the claim that they are relevantly similar.

Taken together, these reasons support the view that we are not fundamentally different from moral exemplars. Instead, experience and practice builds up their moral character and stable dispositions that produce their willingness to make significant sacrifices towards moral ends. This would be possible for the rest of us were we to dedicate ourselves to moral projects. We claim that moral exemplars are morally praiseworthy, and we have good reasons to believe that they are relevantly similar to us. Therefore, it is possible for us to behave like moral exemplars, and we are morally obligated to do so.

This conclusion receives a significant amount of resistance from ordinary people because it involves changing a moderate conception of morality to one that is more demanding. The conclusion that ordinary people fall short of obligation requires ordinary people to change their personal projects or accept a greater degree of sacrifice to act according to the obligations that moral exemplars recognize.

Even if we accept the conclusion that ordinary people fall short of obligation, there might be a separate claim to be made about our moral blameworthiness. Our social environment makes it especially difficult to gain relevant moral knowledge that might in turn impact our moral culpability. We will return to this argument later. I will first consider a further argument about moral exemplars that suggests that ordinary people might not be falling short of obligation after all. Vanessa Carbonell argues that moral exemplars are not different from the rest of us, but that their obligations do, in fact, differ from our own.

Chapter 2. Vanessa Carbonell's Moral Ratcheting Argument

To this point, I have given reasons to think that moral exemplars are like the rest of us in relevant ways; they are not extraordinarily talented or blessed with rare traits. I showed that this leads to the conflicting suggestions that either moral exemplars are acting beyond the call of duty and ordinary people are doing what duty requires, or moral exemplars are the ones fulfilling obligation while ordinary people fall short. Vanessa Carbonell agrees that moral exemplars are like the rest of us, but she argues for the surprising conclusion that both moral exemplars and ordinary people are doing what duty requires. On this view, ordinary people are not necessarily falling short when they fail to act as moral exemplars do.

Section I: Defining Reasonable Obligations

No one denies that different agents have different obligations from one another. Throughout the course of our lives we take on new obligations as we age, take on professional responsibilities, families, and so on. This can create variability in the obligations that individuals face. For example, parents have a special responsibility for the well-being and conduct of their own children that others do not. However, in the cases of moral exemplars that we are examining specifically, we are considering obligations that apply to all moral agents. Obligations regarding the way that we ought to live our lives, conduct ourselves in interpersonal relationships, act altruistically and so on, apply to all moral agents. Despite this, Carbonell argues that both the moral exemplar and the ordinary person can be doing what is required of them. On her view, what someone is required to do depends on several factors which might apply differently to similar moral agents. These factors arise from the particular premises of her moral theory. She

endorses an account of morality on which moral agents face obligations that are the product of the legitimate demands of others.

“Others” refers to members of the moral community that the appraiser is also a part of. This fits with our intuitions about our duties to others quite well. In the absence of extreme circumstances, Eleanor has a moral obligation to keep her promise to Leila because it is reasonable for the moral community to demand that people keep the promises that they make. “Others” could also refer to oneself. If, for example, James realizes that he has unhealthy habits, he might come to see this as a violation of his own reasonable demand that he maintain good health, and recognize a moral failing on these grounds. This example suggests that what is truly important to Carbonell’s moral theory is the distinct character of moral demands.

“Demands” refers to the moral claims that agents make of one another. Carbonell argues that demands are legitimate when the demands originate from agents with the proper authority. Her claim is supported by a view that defines obligations as relations of accountability. The accountability relationship is detailed further by Stephen Darwall who argues that obligations are second-personal claims. He explains that moral demands without accountability do not have the appropriate authority to ground obligation.¹⁴ Consider the demand that one does not steal. In the first-person case, one agent demands of another that they do not steal from them. Darwall points out that in this case, the strength of the demand is dependent on the authority of the person making the demand relative to the potential thief. Compare this to the authority of a demand of a moral community. Here, the authority to make the demand comes from the agreement between moral agents in a community to enforce it.

¹⁴ Darwall, S. (2010). Review: *Precis: The Second-Person Standpoint*. *Philosophy and Phenomenological Research*, 81(1), 216–228.

Enforcement motivates the argument that second-personal claims have moral authority. Darwall follows J.S. Mill in arguing that there is a conceptual connection between second personal accountability and obligations.¹⁵ X demands ϕ from/of Y, where X has the authority to make such a demand and ϕ is reasonable. In this case, Y is accountable to X, meaning that if Y fails to perform ϕ , Y fails to do what is required and is worthy of reproach, blame, and so on. Evidence for the view that obligation and accountability are linked is that we do not claim that an action is morally wrong without implying that an agent who so acts ought to be held responsible provided there is no excuse for the wrong action. For example, when the moral community asserts that we have an obligation to keep promises, failing to keep promises without a reasonable excuse is worthy of reproach. If we did not think that their action was worthy of reproach, then we could not say that it was morally impermissible.

Merely specifying that obligations are generated from the second personal perspective with authority is not sufficient. It is also necessary that the demands of others be “legitimate.” What makes a demand legitimate? One clear constraint is set by ability; it seems quite uncontroversial to say that we can only be required to do things that are possible for us to do. This constraint has a strong and weak interpretation.

On the strong interpretation, only our abilities are considered in determining what we are obligated to do. It is however, particularly difficult to identify what counts as an ability. Are these abilities one is born with? Abilities one develops or ought to develop? Abilities of the average member of a moral community? These questions are particularly complicated, but I will set them aside because this strong interpretation seems sufficiently challenged by the intuition

¹⁵ Darwall 220-222

that it is good for moral agents to lead diverse lives. If the scope of obligations were only limited by ability we might think that obligations demand that we use the extent of our abilities to pursue moral ends. This seems to lead to an outcome in which we might worry that being required to respond to the moral demands of others would not adequately recognize the value of non moral projects, experiences, and relationships. It seems that our intuitions about obligation are better explained with a weaker interpretation of what is “possible” for us.

On weaker interpretations, there might be other relevant factors that are reasonable to consider when determining what we are obligated to do. In the first chapter, I pointed out that projects of central importance to our lives might seem to be one such factor. It certainly seems to many of us that our careers, families, and other similarly important features of our identity limit what others can reasonably demand of us. However, I gave some reasons to think that accepting projects as the limiting factor on obligation seems to suggest that mere wants are enough to avoid particular obligations. This is an unsatisfactory conclusion because it seems as though obligation should be functionally independent of our desires. We may desire that an action not be obligatory, but that should not determine whether it is. Obligations, following this rule, would break down because there would be too much variability due to individuals' preferences. Duties to others, and the norms of interpersonal conduct, would be especially nonsensical because there would be no expectation of accountability behind any demand. No one could rely on promises for example, because if keeping the promise conflicted with an agent's mere wants, there would be no basis on which to reprimand her for breaking the promise.

Not all considerations can be reduced to mere wants as projects seem to. Carbonell considers the significance of sacrifice to be an important consideration that determines what we

are obligated to do. On her view, accepting a sacrifice involves capacity and reasonable belief which are relevant for what an agent can motivate themselves to do.¹⁶ This is different than mere wants which we can be expected to control when they conflict with the reasonable demands of others.

As mentioned in the first chapter, there are two distinct types of sacrifice an agent might experience: direct physical harm and non-physical losses of well-being. It is especially challenging to consider whether or not suffering a physical harm is ever reasonable to demand from an agent. It seems that in many cases an agent would have a right against physical harm that would render the demand illegitimate. For my purposes, I am focusing on losses of well-being that do not involve physical injury. I am not considering our intuitions about the acceptance of sacrifices that are physical injuries because I am focusing on the projects we choose and the considerations that keep us from doing more to help others. We have more to learn from examining intuitions towards accepting non-physical sacrifices because these considerations are more often reasonable and are pertinent to our choices in many more cases.

It is appropriate to consider wellbeing loss when determining the moral status of ordinary people because this consideration is typically relevant to our choice of projects. In fact, if we think of well-being loss in broad terms, we can recognize it as the cost of our choices. When an agent chooses X over Y, we can think that they lose any potential well-being that would have come from Y. For example, when faced with two job offers, one might sacrifice a higher salary upon learning that one of the companies is involved in shady business practices. It is important to note that this definition of sacrifice is not the one that is frequently deployed in everyday usage.

¹⁶ Carbonell 232-234

“Sacrifice” is commonly associated with negative emotions and feelings of loss or sadness. However, “sacrifice,” as it is used here, is roughly equivalent to “cost;” sacrifice just is the cost, in terms of well-being, that accompanies a given action or choice. According to Carbonell, the moral exemplar chooses a morally praiseworthy project, and accepts some sacrifice of well-being in order to do so. An ordinary person however, sees deviation from her current project as a loss of well-being that is unreasonable to demand. This raises the further question, how should we accurately determine what sacrifices are reasonable to demand? What seems reasonable to an agent, and what is really reasonable, often comes apart.

Individual preferences vary widely across individuals and impact how reasonable a sacrifice seems to a particular agent. Sacrifices that seem reasonable to one person may seem wholly unreasonable to another. Taking a pay-cut may seem reasonable to someone who is wealthy, but to anyone facing challenging economic circumstances, the same pay-cut might be unthinkable. Fanatic collectors also might value certain items far beyond what ordinary individuals would. A car collector would find sacrificing the ability to drive certain cars because of their negative environmental impact especially unbearable. Her strong individual preferences for driving these cars would make a demand that she sacrifice the ability to drive them feel unreasonable. The average consumer without these strong preferences likely would not consider restrictions on driving these cars to be unreasonable. Furthermore she might think that it is reasonable for the collector to make such a sacrifice. This example illustrates that the same action types will represent varying degrees of sacrifice for different agents.

One might take this as evidence that what is reasonable to demand of one agent is not reasonable to demand of another. Applying this line of reasoning to obligation faces serious

challenges however because it risks trivializing obligation by returning to a case in which mere wants determine what an agent is required to do. In order to avoid individual biases, we need an objective standpoint to evaluate what sacrifices are reasonable to demand.

An agent-neutral approach to determining what sacrifices are reasonable is useful because it renders individual preferences mutually intelligible between agents. On Carbonell's view that obligations are formed by the demands of a moral community, individuals' preferences must be intelligible to others. Agents with unusual preferences, such as a desire to grow and study slime mold, may be overlooked by a moral theory that bases obligation on what appears reasonable to a community of moral agents. Respecting diverse preferences requires the moral community to appreciate the value of those preferences relative to the individual agent. We do not need to have the preference for cultivating slime mold to recognize the importance this might have for someone with this preference. Carbonell argues that an agent-neutral approach is the most effective way to respect the diversity of preferences in a moral community. She specifically endorses Darwall's rational care account of welfare. The "Rational Care" theory defines welfare as, "the concept of what we would rationally desire for someone insofar as we care for her, or equivalently, what is rational to desire for her for her sake."¹⁷

This account of well-being addresses the problems that arise when individuals evaluate well-being from their own point of view; people tend to overvalue their own perceived well-being, undervalue others', and often fail to be concerned with things that are objectively good for them. Applying the rational care theory of well-being is useful because it links well-being to concern, but not to the preferences of the person whose well-being is at stake. This

¹⁷ Carbonell 233-234

does not entail a cold lack of empathy for the individual interests and preferences of agents however. Instead, the specific preferences an agent has fall within a larger category of well-being. In other words, preference satisfaction may be a part of well-being insofar as we should want an agent that we care about to have some degree of preference satisfaction. It seems harmful for an agent to never have any of her preferences satisfied; some degree of preference satisfaction is important for people's well-being. If Lucy prefers learning to play the trumpet to taking painting lessons there is some sense in which she would be better off learning to play the trumpet. However, a comprehensive theory of well-being ought to take many more factors than preference satisfaction into account. An agent might have preferences that are ultimately bad for her. She might prefer to smoke despite this directly leading to a loss in her overall well-being. An agent that cared about her would not want this particular preference to be satisfied. She might also have conflicting preferences in which satisfaction of one necessarily violates the other. For example, she may prefer to live in one particular neighborhood but also prefer to take a job that is ten hours away. As an agent who cares about her, it does not make sense to want both preferences to be satisfied. Rather, we would consider further factors about each option relative to her well-being and desire that she choose one preference over the other. And insofar as there is positive value in learning from mistakes or disappointment, there is reason to think that our well-being is partially determined by having our preferences denied from time to time. For example, we refer to people who have overly high levels of preference satisfaction in one of two ways. They are either spoiled, meaning their preferences were constantly satisfied without regard for other considerations like costs or well-being. Or we might think that an agent achieves this by setting their preferences to accord with what is truly good for their well-being. That said, it seems

as though humans begin their lives desiring the first option and must learn to embrace the second option. This process of learning can happen through a process of being denied one's preferences and reflecting on one's well-being after the preference was frustrated. If this is true, we should want people that we care about to have some amount of preferences going unsatisfied so that they might experience frustration, desire, determination, and perhaps even learn to prefer what is really good for them. The rational care account gives us a way of understanding what it is for a sacrifice to be more or less significant from a view in which our preferences do not get in the way of determining our well-being.

If we accept the rational care conception of well-being, we still need a compatible theory to define what specifically constitutes well-being. Carbonell argues that it is compatible with objective list theories. These are theories that attempt to define metrics of well-being based on a list of objective criteria. For example, Martha Nussbaum proposes a list of "capabilities" that are relevant to assessing the well-being of agents. On this view, the capacities of agents determine what comprises their well-being. Nussbaum's conception of well-being specifically includes items such as: control over one's environment, life, bodily integrity, and affiliation, among others.¹⁸ Without arguing for one objective list over another, it is clear that the objectivity condition is important for impartially assessing whether a sacrifice is reasonable and supporting Carbonell's view of obligation.

Section II: Carbonell's Knowledge Conditions

Determining that a sacrifice is objectively reasonable, and is demanded by others with moral authority is not sufficient to ground obligation on Carbonell's view. There is one further

¹⁸ Carbonell 236

condition that must obtain for an agent to have an obligation to ϕ —the knowledge condition.

Carbonell argues that what an agent knows, or reasonably believes, is relevant for what she can be required to do. We cannot be motivated to act by considerations that we do not know about.

On the weaker interpretation of ability we were considering, this inability to motivate action suggests that an obligation is illegitimate when an agent lacks relevant knowledge because she can not be expected to fulfill the obligation.

There are two ways of formulating the knowledge condition. One formulation is that an agent's obligations depend on what she knows about her obligations. This formulation seems far too strong to be taken seriously as it suggests that all it takes for an agent to not be obligated to act is ignorance. In extreme cases, the actions of agents lacking the capacity for moral consideration, such as sociopaths would have to be considered morally permissible because the concept of an obligation would be beyond their knowledge. Even in less extreme cases, we would lose the moral justification for reprimanding willfully ignorant or negligent agents because they might not know that they have an obligation to apply moral rules to their actions. This formulation of the knowledge condition trivializes obligation to such an extent that I will leave it aside in favor of the second formulation.

The second formulation of the knowledge condition can be stated in terms of what it is reasonable for an agent to believe. This formulation is significantly more tenable because it often seems appropriate to make moral judgements about ignorant moral agents. In practice, we base these judgements on a standard of what it would be reasonable to expect an average moral agent to believe. I will use this formulation of the knowledge condition

This knowledge condition covers two kinds of relevant knowledge; on Carbonell's view, an agent must reasonably be expected to believe both moral and non-moral facts about sacrifice in order for her to be obligated to ϕ .¹⁹

The necessary non-moral knowledge an agent needs about sacrifice is whether or not making the sacrifice will be bearable for her. This seems like an intuitive extension of the principle that we can not be required to do things that are impossible for us. An objective theory of well-being might suggest that a sacrifice is bearable. However, according to the knowledge condition, an agent's justified belief that an act requires unbearable sacrifice is sufficient to make it unreasonable to expect the agent to carry out that act. On this view, the difference between moral exemplars and ordinary people that leads to the differences in their obligations has something to do with non-moral knowledge of sacrifice.

The moral exemplars we are considering know that the sacrifices associated with their praiseworthy acts are bearable. Perhaps by experiencing the sacrifice, they might learn that despite what people commonly believe, the sacrifice is not so substantial. Furthermore, they might find that there are other benefits that result from their actions that they value as much or more than what they gave up.

In response to the testimony from exemplars that a given sacrifice is not really so bad, one might be tempted to think that the exemplars have particular benefits that are conferred on them by their good actions. An agent might be particularly attuned to the subtle benefits that come with a very good act of altruism. We might worry that this agent is not actually behaving in a morally exemplary way, if she is compensated for her sacrifice by resulting benefits. This way

¹⁹ Carbonell 241

of thinking defines sacrifices as net losses of well-being, in which some benefits conferred by acting compensate for losses from the sacrifice. However, it would be a mistake to consider the benefits and the costs of actions in this way. Consider the following case; Joe is an Olympic ice skater, and diverts large amounts of time from school, family life, etc. to training. Joe eventually competes in the olympics and feels the satisfaction of representing his country, and feels extremely positive emotions that he would not have felt had he decided not to dedicate himself to figure skating. In fact, Joe feels that, given these positive outcomes, the sacrifices were worth it. In this case we still recognize that Joe made sacrifices, he was not able to pursue a life that many young people find fulfilling and this might have caused him some hardships at the time. Carbonell refers to the sacrifices here as gross losses of well-being.²⁰

What is necessary for an obligation then, according to the knowledge condition, is that agents know, or reasonably believe, that a gross loss of well-being that might result from their morally praiseworthy action is not too much for them to bear. When taking on an action, we generally do not know what the net effect on our well-being will be. For example: we typically do not know ahead of time what the outcome of our actions, what we will feel at the conclusion of our action, or what the knock-on effects of the good that we do might be. Knowledge of gross losses of well-being on the other hand is more feasible because we can learn or reasonably estimate what the costs for our actions will be when making a decision. Of course, there might be cases in which there are unforeseeable gross losses of well-being that result from a particular decision, but given that they are unforeseeable we cannot make our decisions on the basis of these factors. I take it, then, that Carbonell's non-moral knowledge condition excludes

²⁰ Carbonell 237

considerations of these possibilities as well. An agent must know or reasonably believe that a gross loss of well-being is bearable in order for that sacrifice to be legitimately required.

Section III: Carbonell's Moral Knowledge Condition

The non-moral knowledge that a sacrifice is bearable, is necessary but not sufficient for it to be obligatory on Carbonell's view. Carbonell's knowledge condition also states that for an agent to be obligated to \emptyset , she must know or reasonably believe that she is obligated to \emptyset .²¹ Recall that for Carbonell obligations are established by the reasonable demands of others. This moral knowledge condition requires an agent know two moral facts: knowledge of who has the authority to make legitimate moral demands, and knowledge of which demands require action.

In order for an agent to be obligated to \emptyset , Carbonell argues that an agent must know that it would be reasonable to expect other moral agents to accept the gross losses of well-being that are demanded of her. If an agent knows that others within her moral community would accept some sacrifice in pursuit of a moral commitment, then she should think that it is appropriate for her to accept this sacrifice as well.

The moral knowledge condition also requires that an agent knows, or can be reasonably expected to believe, the demands of others. A way in which a moral agent might come to know what others demand is simply by growing up within a moral community. She might ask herself what moral norms she is governed by, and what objectively defined sacrifices she finds appropriate to demand from her peers.

This requires knowledge of which agents have authority to make moral demands.

Carbonell and Darwall seem to agree that any moral agent has equal authority to make demands

²¹ Carbonell 241

as a member of the moral community. This argument forms the basis for many contractualist moral theories; T.M. Scanlon's argument that obligation arises from principles that no one could reasonably reject, and John Rawls' argument that obligation ought to be determined from behind a "veil of ignorance" are two excellent examples.²² In both cases, moral agents are considered to be in relationships of mutual accountability. Each agent is given equal consideration when determining the obligations that the others face. If it is reasonable to demand that people have freedom of expression, then it must be true for each member of the moral community. Each individual agent has the authority to reproach someone who violates the reasonable demands of the moral community.

Given the claim that each moral agent has the ability to reproach any other, Darwall claims that in the absence of a legitimate excuse, only a lack of ability to recognize or acknowledge the demands of others changes what an agent is obligated to do. He argues that the demands are generated by principles that every moral agent has epistemic access to. As a result, failing to recognize these demands, entails being unable to make or assess second personal moral judgements—a literal lack of moral agency. Darwall takes it to be a merit of his account that moral obligations are "inescapable" in this way; neither our individual interests, aims, nor other idiosyncratic tendencies, can absolve individuals of accountability to the second personal demands of their moral community.²³ Accepting this claim would suggest that most obligations are the same for each member of the moral community. However, Carbonell argues for the opposing view that ordinary individuals and moral exemplars face different obligations on the basis of differences in what they can reasonably be expected to know.

²²Darwall 226-227

²³ Darwall 228

One possible explanation for moral exemplars' different knowledge of obligation could be that moral exemplars constitute a separate moral community. On this view, when moral exemplars' claim that they are obligated to make more substantial sacrifices to help others than what ordinary people reasonably believe is required, they are separated from the demands of ordinary moral agents. The notion of mutual accountability breaks down on this view because it suggests that moral exemplars have greater capacities than ordinary people. On this view, moral exemplars respond to demands that would be illegitimate for an agent to make of members of their own moral community. If moral exemplars constituted a second moral community, their actions would not be intelligible to ordinary people. Ordinary observers could not consider the exemplar's action morally praiseworthy because they would not be able to comprehend the moral basis for the action. The observers might think that the action is praiseworthy in a nonmoral sense because it lead to an increase in well-being. However, their praise would lack the moral component because it could not make reference to the demands of others. Carbonell rejects this explanation, because she argues that moral exemplars are fundamentally the same as the rest of us; they are moral agents with the same capacities as the rest of the moral community.

Instead of suggesting that capacities set moral exemplars apart from ordinary people, Carbonell argues that agents recognize differences in the level of sacrifice it is reasonable to expect from others. Recall that Carbonell claims that it is illegitimate to demand that an agent make sacrifices that she believes would be too much for her to bear. In a case of mutual accountability, each moral agent recognizes that each agent they might make a demand of could have a variety of considerations that alter their non-moral knowledge about what is bearable. Furthermore, each agent would respect this non-moral knowledge as a limitation on what they

can reasonably demand, because they would also expect to be treated in this way. Moral exemplars recognize this fact and claim that others are not obligated to do what they are doing. Likewise, observers recognize that moral exemplars have acted in a morally praiseworthy way because they recognize that demands that are too costly for them to bear must not be unbearable for moral exemplars.

Carbonell argues that the knowledge that moral exemplars have regarding sacrifice ratchets up their obligations.²⁴ “Ratcheting up” is a comparative term describing the process in which the exemplar faces more demanding obligations relative to ordinary people when they gain knowledge that sacrifice is bearable. They also face more demanding obligations upon learning that others legitimately demand some sacrifice. Carbonell’s ratcheting up effect relies on the knowledge condition; it is a process by which an exemplar provides important knowledge to observers which changes what it is reasonable to believe can be demanded of them. Carbonell argues that observers face more demanding obligations when exemplars demonstrate that the belief that some sacrifice is unreasonable is based on limited evidence. The more that agents observe other people taking on substantial sacrifices in response to a demand that they recognize as an obligation, the less reasonable it is for the observer to believe that the sacrifice is too demanding.

Carbonell suggests that this effect can be observed in the case of agents who perform similarly praiseworthy acts. For example, if David is a philanthropist and attends a conference of other philanthropists, he might learn from the testimony of his peers that donating an additional 5% of his income would not represent an unreasonable sacrifice and would answer the needs of

²⁴ Carbonell 245

many more people. In this case, Carbonell would say that upon gaining this knowledge, David is obligated to give more. The same is true for ordinary people as well. Imagine a case in which Mary spends nearly every Saturday volunteering at a local homeless shelter even when it means passing up other enjoyable or important events. She takes her volunteer work to be an obligation that stems from the legitimate demands of moral agents who rely on the shelter for food, protection from the weather, and other rights. Her friend Tom thinks that the volunteer work is morally praiseworthy but requires too much sacrifice to constitute an obligation. By observing Mary's actions, and hearing her testimony however, Tom should gain evidence that the sacrifices involved are bearable and generally reasonable to expect him to make. In turn, this would suggest that other things being equal, Tom should recognize an obligation to volunteer that he did not prior to observing Mary's actions.

Carbonell also claims that moral ratcheting applies to agents with very different levels of knowledge and, on her view, different obligations.²⁵ This occurs in the case of the paradigmatic moral exemplar and ordinary people. When we read about, or perhaps witness the work of exemplary people we see that something far beyond what we consider bearable sacrifice is in fact possible. We might further learn that these exemplars who have experienced the gross well-being losses that accompany their moral projects do not find the subjective experience of these losses to be so bad. In these cases, while it is unlikely that we will suddenly be obligated to behave like moral exemplars, we have good evidence to suggest that we could be doing at least a little bit more. Our new evidence suggests that even if we struggle to believe that we could take on the impressive sacrifices that exemplars do, we should update our beliefs that we cannot bear

²⁵ Carbonell 245-246

to spend more time volunteering, increasing the amount we donate to effective charities, and so on. According to Carbonell, our obligations change to take this new evidence that exemplars provide into account.

While Carbonell's argument explains the disparity between the testimony of exemplars and observers, it seems as though there are further problems that arise if we accept that what we are obligated to do is determined by the reasonable demands of others limited by our knowledge about sacrifice. These problems involve: injustices that occur to agents who are kept separate from the moral community, collective action problems, and motivated ignorance. In the following chapter, I will argue that we should abandon the notion that what we are obligated to do depends on our reasonable beliefs.

Chapter 3. Separating Obligation from Reasonable Belief

Having laid out Carbonell's account of the relationship between moral exemplars and the rest of us, I now want to focus on the moral knowledge condition. In the previous chapter, I explained that there are two possible formulations of this condition and argued that the "reasonable belief" formulation is the strongest.

In this chapter, I will focus on the relationship between reasonable belief, reasonable demand, and what we are obligated to do. I take Carbonell's knowledge condition to be an internalist view of moral obligation. I will raise concerns that the internalist view fails to capture what we are obligated to do in contentious moral situations. I will argue instead for an externalist view of moral obligation; that obligation does not depend on agents' ability to know that certain actions are required of them; what matters is what agents are actually obligated to do.

Section I: Internalism and Externalism about Moral Obligation

On the internalist view, an agent is required to know, or reasonably be expected to believe, that she is obligated to perform action x . On the externalist view it is not necessary that an agent knows that she has an obligation to x for this obligation to exist. It is worth noting that internal and external views about moral obligation ought to agree about what actions are required of moral agents. Agents should reasonably be expected to believe the truth about what they are obligated to do. However, there are cases in which an obligation does not conform to an agent's moral norms. In these cases, it seems unreasonable to expect that an agent knows that she has a particular moral obligation.

Before analyzing these cases, it is important to consider what constitutes reasonable moral ignorance. I am ascribing to Carbonell a formulation of the internalist view which holds that when it is reasonable for an agent to believe that she does not have an obligation to \emptyset , she is not obligated to \emptyset . There are plenty of cases in which we might have inadequate evidence about the needs of others, or the harms that our current status quo causes that could be avoided. On this view, without sufficient contrary evidence about moral facts, no agent violates an obligation to continue acting in accordance with their moral community's moral norms.

I reject the claim that what is reasonable for us to know about moral obligation impacts what we are actually obligated to do. Our obligations, especially those concerning others, persist even when we do not have reason to think we have such an obligation. Additionally, the internalist view does not adequately explain moral education or allow the possibility of agents gaining a better view of what they were obligated to do all along.

Moral education gets straight to the heart of the tension between internalism and externalism because it involves an agent gaining knowledge that she is obligated to \emptyset that she had not formerly recognized. The knowledge dependent view of obligation claims that prior to learning that others reasonably demand \emptyset of her, she was not obligated to \emptyset . Although this view might apply to children who do not possess the relevant capacity for moral understanding, it does not apply to moral agents with the capacities to make moral judgements.

Imagine a case in which a child grows up in particularly challenging circumstances and is lead to believe that using people as mere means to her own ends is a natural way of life. In such a case, it is difficult to say whether and to what extent her moral capacities have developed given that she was raised in a particularly challenging context. There is a contentious point of debate

here about whether agents with particularly challenging upbringings lack moral capacity. Susan Wolf argues that this agent lacks the capacity to correctly make moral judgements. The intuition behind her view is that the agent would enter the wider moral community and be unable to make correct moral judgements.²⁶ However, I do not think that a capacity is lost when an agent has an upbringing that inculcates her with false moral beliefs. I follow Sarah Buss in arguing that while it may be reasonable to expect this agent to make incorrect moral judgements, she does not lack a capacity to learn that she is mistaken.²⁷ Furthermore, it is appropriate for others to correct the agent whose upbringing gives her good reason to have false moral beliefs. Correcting this agent respects her capacity for moral reasoning. It is quite difficult to imagine what one might say to someone who has been raised in this way to get them to recognize that they are wrong, but rejecting her capacity as a moral agent on the basis of her reasonable false beliefs is wrong. For my purposes, it will suffice to say that if it is possible to show the misguided agent that her view is false, we must also say that she violated an obligation to treat people as ends even while she could not reasonably be expected to believe that this was obligatory. The fact that she has a good reason not to know that she is wrong, does not change the fact that it is morally wrong for an agent with the capacity for moral reasoning to treat others in this way.

Section II: The Case of Problematic Moral Norms

Setting individual instances of false but justified moral beliefs aside, we should consider obligations that exist even if a moral community does not recognize an obligation to \emptyset . This can be seen in the case of moral paradigm shifts over time. The expansion of human rights is one

²⁶ Wolf, S. (1982). Moral Saints. *The Journal of Philosophy*, 79(8), 419. <https://doi.org/10.2307/2026228>

²⁷ Buss, S. (2017). *Some Musings about the Limits of an Ethics That Can Be Applied*. Draft.

such example. When the status quo involves the violation of human rights, the apparent normalcy of the practice might interfere with an agent's ability to recognize her failure to act in accordance with these rights. Over time, people continually learn, gather evidence, and debate contentious moral questions, which leads to breakthroughs at the frontiers of moral knowledge.²⁸ This new knowledge is similar to the knowledge that the child with a morally challenged upbringing gains when she comes into contact with other moral agents who critique her problematic views and convince her that she is wrong. Unlike in the case of a particular abnormal upbringing, the frontiers of moral knowledge implicate an entire moral community. Instead of an isolated case where an agent emerges from her specific circumstances into a moral community where here false (but justified) moral beliefs are abnormal, social norms can be cases of widespread problematic moral beliefs. The difference is significant because while an abnormally troubled upbringing might be corrected by the reasons and demands of others within the broader moral community, there is no such corrective force for people who take the very normalcy of their moral beliefs as justification that they are correct. Even when new information arises at the "frontiers of moral knowledge" that implicates the moral status quo, it is a difficult process for this fact to update the moral norm. The process involves a confluence of social factors that reinforce the normalcy of an updated moral view.

However, we might wonder what the status of the demand was before it was adopted as part of the moral community's set of norms. Given that moral agents eventually gain sufficient reasons to recognize the moral authority of the source of the demand, the subject issuing the demand must have been such that the evidence was always there. This is true despite the moral

²⁸ Calhoun, C. (1989). Responsibility and Reproach. *Ethics*, 99(2), 389–406.

community failing to notice or appreciate it. As a result, we should conclude that the moral community made a mistake. The externalist view recognizes that the obligation to the newly recognized moral subject was ever-present. This means that ordinary people reasonably believed false moral information; they were obligated even though they could not reasonably know it.

There are many examples that highlight the consequences of the debate between internalism and externalism regarding societies' moral developments. One need only look at the news for five minutes or open a history book to recognize the ongoing struggles for civil rights and self-determination faced by marginalized people the world over. Often, these struggles are the direct result of a moral community failing to give equal authority to marginalized moral agents within it. The internalist picture of moral obligation is able to recognize that it is wrong for this omission to occur. For example, many contractarian principles arise from thought experiments in which agents choose moral principles behind a "Veil of Ignorance" or in an "Original Position" that is idealized. Even Scanlon's account that moral obligation arises from principles that no one could reasonably reject is an idealized case in which the moral community is sufficiently homogenous for moral evaluations to be intelligible to its members. It is a failure on any view of obligation when moral communities in practice fail to give moral agents equal moral authority. The internalist view is able to claim that these situations are wrong and morally problematic. Furthermore, they also want to acknowledge that moral progress can be achieved. However, they also seem to insist that until the progress is sufficiently widespread, the specific actions of people who formed reasonable beliefs as a result of problematic norms were not themselves wrong. It is hard to see how we can reconcile admitting that wrongdoing occurred with the claim that individual actors were not obligated to act differently.

This is a problematic outcome because it upholds the privilege of those who in practice exercise control over the moral status quo. This view does assume that the moral status quo has some basis in sociological effects; what we come to demand as a moral community has some basis in what we are encouraged to think about, what views are given a platform, and so on. Privilege in this context refers to a condition in which a particular group is afforded greater moral consideration and inflated authority to make demands of other moral agents. We can identify agents who occupy this privileged position because there are normalized practices that disproportionately protect the rights of this group beyond rights of others. In America, wealthy, straight, white, males have, and continue to be given disproportionate consideration as moral agents. Movements like “Black Lives Matter” and “#MeToo” seek to highlight the ways in which this inequality of consideration within the moral community constitutes significant wrongdoing. The difference in treatment of black people under the law, and the gaslighting that women frequently face when they try to report sexual violence are only two examples of conditions that are caused by a lack of equal moral consideration. Control over the moral status quo, that reinforces the normalization of moral privilege is enabled by a variety of processes: monopolization of the power to establish and control institutions of learning, economic and political domination, outright violence, and so on. The efforts of activists aimed at making privileged agents recognize that they are falling short of obligations when they fail to respect the moral status of others. This is the case even if they are justified in believing the norms of their moral community over which they exercise illegitimate control.

The internalist picture of obligation is also problematic because it seems to allow cases of motivated ignorance. In most cases of moral development, discoveries or innovative moral

arguments contribute to a new moral paradigm. However, this new paradigm entails that actions justified by earlier norms were problematic. In this situation, the privileged beneficiaries of a flawed moral status quo have the ability to appeal to the widespread problematic norm as a justification for ignoring mounting evidence that they are wrong about their obligations to others. In practice, this can be observed when male-dominated professions systematically undervalue the work that women produce. Accepting the value of women's intellectual labor carries the implication that it was and continues to be wrong to demean women's accomplishments. It is unlikely that any individual male explicitly thinks this, but institutional design and the normalization of practices that are demeaning to women might seem to justify the problematic actions that keep these views in place. Cases of motivated ignorance are epistemically justified, but morally unjustified. The internalist account, which combines these methods of justification overlooks this possibility, and fails to identify the morally objectionable actions of individuals who fail to challenge the status quo.

My argument might be challenged by a concern that it seems as though privileged agents would need to form moral beliefs in epistemically unjustified ways in order to avoid violating obligations. This could be cause for concern on the grounds that privileged agents might not recognize the justification behind why they face more demanding obligations than it was once reasonable for them to believe. If it is true that social cues justify beliefs that ending oppressive behaviors is not obligatory, then it seems as though privileged agents are doomed to believe the wrong things about moral obligations to marginalized agents. However, it should be noted that there are two distinct ways out of this worry. The first is to simply accept that our justified moral beliefs can be incorrect. This might be considered a form of moral agnosticism whereby we

ought to operate under the assumption that reasonable beliefs about obligation are true, while giving an appropriately high credence to the possibility that they are false. This would entail a sensitivity to the testimony of marginalized agents and suggest that we should be willing to update our moral beliefs accordingly.

The second is to reject the idea that social norms can ground justified beliefs about obligation. Liz Harman argues that if an agent knows the relevant non-moral facts that pertain to a morally bad action, there is sufficient evidence to discover that the moral belief justifying the act is wrong.²⁹ The thought behind this is that the capacity to discover what is morally wrong exists whether we know the moral status of our actions or not. For example, at some point in history an agent may have claimed that women should not enter the workforce. This agent might think that this view is justified because it is a widespread perspective, there are no sanctions for expressing it, people that he considers moral authorities have told him that this is the case, and so on. Suppose the justification for this belief is based on a false idea that women lack various capacities to work. Harman claims that the non-moral evidence that undermines this claim is present and accessible. All it would take to undermine this view is to allow women to work and see whether they have the capacity to do so. Any false belief formed by accepting social norms is not justified on Harman's account. According to Harman, anyone with the capacity to make moral distinctions never lacks the potential to realize that their justified moral beliefs are false.

In either case, relying on social norms does not insulate us from the ability to be wrong about what obligation requires of us. As a result, I reject Carbonell's claim that what we are obligated to do is simply the product of what we can be reasonably expected to know about the

²⁹ Harman, L. (2011). Does moral ignorance exculpate? *Ratio*, 24, 443–468.

demands of others. The knowledge condition also concerns the non-moral knowledge that is relevant to accepting sacrifice. I will turn now to addressing this condition.

Section III: Challenging the Non-moral Knowledge Condition

Carbonell also argues that reasonable ignorance of non-moral facts changes what an agent is obligated to do. Differences in what agents are obligated to do depends on apprehension of the relevant non-moral facts—namely, facts about the bearability of sacrifice. An agent is only obligated to \emptyset if she knows or can reasonably be expected to believe that she can bear the accompanying sacrifices. On her view, the difference between a moral exemplar and an ordinary person is that it is reasonable for the exemplar to believe that a required sacrifice is bearable.

However, if we accept the hypothesis that moral exemplars are just like ordinary people, it is not clear how these agents have this knowledge on Carbonell's own view. She argues that most of us do not have adequate information about whether a sacrifice is bearable to take on. Given the fact that moral exemplars are not fundamentally different from ordinary people, this seems to suggest that they did not have adequate information about sacrifice when they first acted. What should we think about this agent, and how reasonable is it for us to accept that it would be reasonable for us to accept more demanding sacrifices on the basis of observing exemplars.

The testimony of exemplars suggests that they tend to find that sacrifices they made in order to fulfill obligations were not as arduous as they initially expected. They also frequently report that they feel as though they could reasonably sacrifice even more. If we accept their testimony, it does not seem as though the exemplars perceive their own sacrifices to be unbearable. However, this testimony typically comes after the exemplar has taken on a particular

sacrifice for the first time and learned that it is bearable. The knowledge of non-moral sacrifice condition does not hold the first time an exemplar decides to act; while exemplars must overcome not knowing whether a sacrifice will be bearable, it is not clear how they are able to do so. This is a significant case to consider because prior to making the decision to make the morally praiseworthy sacrifice we might think that the exemplar is just another moral agent. Is it reasonable for them to believe that they can bear the relevant sacrifices sufficiently for the associated action to be obligatory? It might initially seem that it is not reasonable given that this is the first time they are acting in an exemplary way, and they do not have superhuman powers of introspection or imagination, it is hard to see how they would know that they could bear it before taking on the action.

This leads to a challenge for Carbonell's ratcheting view of obligation. If the exemplar does not know that she could bear the sacrifice involved the first time that she acted, the non-moral knowledge condition holds that it was not reasonable for her to think that she was obligated to act as she did. She might have a justified false belief that the sacrifice is reasonable. In a case like this, ordinary people seem to gain knowledge relevant to obligation from someone that was not obligated to \emptyset . In other words, observers gain evidence that it is reasonable to make particular sacrifices from an agent that did not know that it was reasonable before she acted. This is incompatible with the claim that the ratcheting up of obligation is a process that respects what is reasonable for moral agents to believe, because it depends on seemingly unreasonable leaps of faith on the part of moral exemplars at the frontiers of moral knowledge. We might think that the process of learning about what sacrifice is bearable occurs through taking some risks, or

accepting we do not know the net effect of accepting a sacrifice before we take it on. Perhaps the observer learns that leaps of faith are necessary to achieve certain morally good ends.

Consider the case of the exemplary activist, who works tirelessly to make her society recognize the rights of its marginalized members. Before she made the decision to dedicate herself to this moral project, she could reasonably believe that she would incur some amount of sacrifice. However, with normal foresight and a recognition that the resulting challenges she would face as a result are unpredictable, she can not know that resulting losses of net well-being will be bearable. Rather, she accepts the risk due to her reasonable beliefs that she has duties to the people harmed by the status quo. She acts because she recognizes that an obligation to others supersedes some degree of consideration for her own well-being and is willing to accept a sacrifice on this basis. We are impressed by the selflessness of people like this, not because they think that sacrifices are insignificant. When we are inspired by moral exemplars, it may be true that we are learning something about what sacrifice it is possible for us to bear. However, it seems far more likely that we are learning and admiring the selflessness that is required to challenge problematic moral norms. After all, I have given reasons to think that we are part of the same moral community, and have the same abilities as the moral exemplar. We too should and can recognize that the moral exemplar is exemplary for facing challenges for morally praiseworthy ends. Both the challenges and the moral goodness of the ends are comprehensible to the ordinary observer; which means that what the moral exemplar actually teaches us, are values of altruism and self-sacrifice in the face of sacrifices which may initially seem unreasonable to undergo.

Section IV: Conclusions about the Knowledge Condition on Obligation

Carbonell admits that the knowledge condition on obligation is controversial. I have shown that both the moral and non-moral knowledge conditions are not sufficient to impact what it is that we are obligated to do. However, Carbonell argues that even if the knowledge condition on obligation were rejected, her argument could be applied to considerations for moral blameworthiness; Our exposure to moral exemplars is relevant to what ordinary people can be blamed for.³⁰ We have the additional benefit of observing the actions of moral exemplars and hearing their testimony about what sacrifice is bearable. As a result, Carbonell argues that our blameworthiness for falling short of our obligations depends on what we can reasonably be expected to believe. Exemplars add to this pool of available evidence and raise the bar of what we can be blamed for. On this modification of her view, failure to meet our obligations seems more like a case of “motivated ignorance,” when we have praiseworthy moral exemplars among us frequently providing us with evidence that our current beliefs are not accurate—we are obligated to do more. This leads to the further question of when our ignorance is excusable. I will now turn to the question of whether there are conditions that might comprise a reasonable excuse for ignorance and subsequent failure to recognize obligations. Is our exposure to moral exemplars relevant to what we can be blamed for?

³⁰ Carbonell 245

Chapter 4. Appropriate Responses to Falling Short of our Obligations

In my analysis of moral exemplars to this point, I have argued that obligation is not variable between agents on the basis of their knowledge. This means that it is possible for us to do wrong without knowing that what we are doing is wrong. Intuitively we might worry that our typical method of adjudicating when agents are responsible for their actions might not be sufficient to address injustices that occur when agents reasonably believe that they are not acting wrongfully. In this chapter, I will argue that normalized injustice leads to conflicts of intuitions about blameworthiness. Ultimately, I suggest that in such contexts ordinary people might be accountable despite having an excuse for their wrongdoing.

Section I: Reactive Attitudes and Ill-Will

In many circumstances, when we conceive that someone has done something wrong, by falling short of an obligation, a further question remains. Is the person to blame for their wrongdoing? One might think that the agent that committed the wrong action is guilty of faulty moral reasoning, a reprehensible disposition, lack of important knowledge, or some other causal condition. We might think that these conditions should be appropriately addressed with various sanctions so that the agent does not continue to behave this way. This kind of corrective action is generally only thought to be appropriate when the agent is responsible for her bad action.

The responsibility I take as my central concern is moral responsibility specifically. This is distinct from responsibility that is part of any particular role obligation (“Judges have certain special responsibilities”) and different than causal responsibility (“The rain is responsible for the flooding”). Moral responsibility determines who is at fault for falling short of obligations and

what constitutes an appropriate sanction when an agent is culpable for wrongdoing. In most cases, moral responsibility also has implications for censure; an agent who is culpable for her wrongdoing is worthy of strong disapproval. Censure often takes the form of reproach directed at qualities of the agent. For example, imagine a case where Tim, an ordinary moral agent with normal capacities, is caught stealing and it is determined that he was stealing because of the thrill that he got from it. It seems appropriate to express disapproval directed at Tim directly because he acted in accordance with his will in the right way. He was not coerced to steal, and there were not additional circumstances that made him feel as though his life was in danger if he did not steal. It seems appropriate to express disapproval towards Tim for failing to manage his desires and motivations in accordance with appropriate conduct. It seems that in simple cases, censure is closely associated with the will of the acting agent at the time of the action; censure is appropriate when the agent's act was the product of ill will. However, the cases in which ordinary people are engaged in wrongdoing for falling short of obligations that they did not have reason to believe that they had seems much more complex. Is censure appropriate in these cases?

P.F. Strawson argues that it is reasonable to identify judgments of moral responsibility with judgments regarding which reactive attitudes would be appropriate.³¹ For example, consider the difference between sadness and resentment. If John learns that his dog ran out into the street and was struck by a car, John will certainly react with sadness. However, if John also learns that the driver of the vehicle was negligent and could have avoided hitting his dog had he been more attentive, John would likely react with feelings of resentment. Strawson argues that this difference in emotional response indicates a difference in responsibility. In the event that the

³¹ Brink, D. O., Nelkin, D. K., & et. al. (2013). Fairness and the Architecture of Responsibility. *Oxford Studies in Agency and Responsibility*, 1, 24.

driver genuinely was not blameworthy for what happened, John might have reacted with sadness toward the driver, but the emotional response could not legitimately be targeted towards the driver who was not negligent. Negligence is something that the driver can reasonably be expected to avoid, and therefore the driver is responsible for the harms caused as a result of this action; it would justify John's resentment. When an action is characterized by ill will, or a lack of good will, resentment against the acting agent is justified. The cases I am considering are those in which agents could not reasonably be expected to believe that they have a particular obligation. These are especially complex because it generally does not seem as though reactive attitudes are appropriate; the agent's action may cause harm, but she is not acting with an ill will. If it is true that the actions and projects of ordinary people fall short of obligations, but blame is not appropriate, then there must be some exculpating consideration.

It is important to note however that some reactive attitudes do seem appropriate. For example, marginalized people whose rights are constantly violated have a right to exhibit reactive attitudes towards those individuals with the privilege and ability to question and impact a problematic moral status quo but do not. For my purposes, I am considering agents who do not have an ill will that drives their complacency or motivates their ignorance. However, once I evaluate the moral status of these agents I will offer a suggestion for sanctions that may be appropriate for complacency.

Section II: Reasonable Excuses

Excuses and exemptions are two ways in which an agent might not be blameworthy. Both take an agent's lack of ill will into account. Exemptions concern cases in which an agent is never

responsible for particular wrongdoing. These can be contrasted with excuses, which are case specific and require that an agent would have been able to act well in normal circumstances.³²

Clear cases of exemption are situations in which an agent is not the kind that has the capacity for self-government. For example, we might imagine an agent who is a mere creature of instinct. Such an agent cannot reliably express her will, whether it be morally good, bad, or neutral. For this reason, her actions are exempt from moral judgement. Determining the intricacies of what counts as a sufficient capacity for moral judgement and determining the responsibility that differently-abled agents have to their actions is very complicated. For my purposes, I am stipulating that the “ordinary people” who are the subjects of the moral judgment in this thesis are self-governing and have the capacity to do more. As a result, they will not meet the general criteria for exemptions, so I will only have to consider whether ordinary people have an excuse for failing to recognize their obligations.

Excuses concern whether an agent is responsible for her wrongdoing. In cases where an agent’s actions do not conform to her will she might have an excuse. The agents who are exempt are excused in all cases because they are not able to act according to their will. I am only considering agents who have the capacity to do more, so I will set aside these particular agents in determining what constitutes a reasonable excuse for failing to act on obligations that it was not reasonable to know that we had.

Considering these ordinary agents, Carbonell’s knowledge conditions might provide reasons to excuse. Recall that while Carbonell admits that the knowledge condition is controversial for determining what we are obligated to do, she claims that the knowledge

³² Brink et. al. 5-6

conditions still hold in the case of moral judgements. On this view, an agent is not blameworthy when she could not know or reasonably believe that she has an obligation. Changes in what is reasonable for an agent to know changes what the agent can be blamed for. While I have argued that the knowledge conditions do not determine what is actually required, perhaps they are sufficient to exculpate agents engaged in wrongdoing.

Ignorance of non-moral facts is exculpating provided the agent has a good reason not to know. Mistakes of perception are simple examples where this might occur. John might make Susan a sandwich, but unknowingly use spoiled ingredients. As a result of his ignorance, he is not blameworthy for making Susan ill provided it was unreasonable to think that he could have noticed his mistake.

Ignorance of moral facts can be exculpating provided the agent has a good reason not to know. This seems to be the case in idealized moral contexts. This context is one in which moral agents are able to arrive at true conclusions about what they are obligated to do as a result of their normal capacities for moral learning. The idealized context is the imagined context that contractual theories of moral obligation use to determine what is obligatory. In these idealized contexts, failure to recognize the demands of other moral agents is indicative of a failing within the agent.

Part of what makes exculpating ignorance reasonable in the real world cases that we are considering is that most ordinary people formulate their beliefs based upon training and education that reflects the social norms that they were exposed to from birth. Generally speaking, these norms form the basis of what is considered reasonable for moral agents to know. In the previous chapter I argued that this does not amount to agents not acting wrongfully, but it is not

obvious that it is not relevant for blameworthiness. Cheshire Calhoun refers to cases in which the social norms that an agent grows up in are problematic, as “an abnormal moral context.” In contexts where wrongdoing is only identified by agents denied proper moral consideration, others within the moral community may continue to form justified beliefs based on evidence from their peers, social interactions, parents, etc and fail to recognize their wrongdoing. In normal contexts, we take these to be an acceptable means of coming to know moral truths.³³ An example of a normal moral context is one in which agent’s mutually recognize the authority and demands of others. In practice, many moral evaluations do work this way. We know that stealing is wrong because we learn from our interactions with others and our society that it is. We are justified in believing many common sense moral facts because we are attuned to moral argumentation, defer to relevant authorities, and test our intuitions in socially accepted ways.

Our reliance on these accepted epistemic processes for forming justified moral beliefs can produce some troubling conclusions. Consider again the case in which an agent is raised with a particular set of problematic norms. We can imagine a child sufficiently isolated from opposing viewpoints with parents who teach her false moral views that she has no reason to refute. Recall that in the previous chapter I argued that such an agent fails to do the right thing despite it being unreasonable to expect her to know that her actions are morally problematic. Susan Wolf argues that in these special cases, the agent lacks normative competence. The agent in this case would not be able to come to the correct moral conclusions and therefore lacks a moral capacity that is necessary for her to be culpable. However, as I stated before, I think that we have reason to reject this claim. The fact that an agent is lead to make false normative judgements does not necessarily

³³ Calhoun 390-392

mean that she lacks normative competency. These are not sociopaths but rather, a person with a perverse background such that there is nothing wrong with her capabilities. She is still able to judge right from wrong, even if she does so incorrectly. She applied normal faculties of moral reasoning to her problematic circumstances and reached conclusions that followed from these faculties. It is noteworthy that this is not unique to these extreme cases. In fact, agents with “normal” upbringings also make justified moral mistakes when the testimony of their peers, the actions of their parents, and the norms of society are misleading. We do not suggest that this is due to a lack of normative competency, and we should not in the case of those with a limited opportunities for moral education either.

In an abnormal moral context, each agent is disadvantaged in regard to their moral knowledge. Agents in the moral community are justified in holding false beliefs about the moral status of a failure to do more to seek out and eradicate injustice. Our ability to understand moral facts means that we are failing obligations from which we are not exempt. However, insofar as our false beliefs are justified, ordinary people seem to have an excuse for failing to act. This excuse resembles the justified excuses in cases of duress. In the law, duress is an exculpatory circumstance because situational factors are determined to have caused an agent’s action separate from her will. We might imagine a case in which a driver exercising appropriate attentiveness drives over a patch of black ice and is unable to avoid hitting a pedestrian in the crosswalk. Having justified false moral beliefs in abnormal moral contexts is similar to these legal cases of duress because an agent’s will is separated from her wrongdoing by factors beyond her control. The difference here is that abnormal moral contexts are the norm rather than an exceptional situation. Situational factors raise significant barriers to agent’s knowledge by seeming to justify

problematic moral views. On these grounds, abnormal moral contexts might exculpate when agents express no ill will in their ignorance.

However, to say that ordinary people have an excuse in abnormal contexts is not to say that we are not responsible for our actions. Responsibility typically implicates an agent in the possibility of censure and sanction. If we return to the conception of legal liability in cases of duress, the agent who caused the harm is responsible for paying damages to the victim in recognition that she was causally connected to the harm despite not being worthy of censure.³⁴ The abnormal moral context cases may differ in degree, but the connection to a harm can be similarly unwillful. I follow Strawson in claiming that we are not worthy of blame when our failure is not a product of ill will or unreasonable negligence. However, contrary to Strawson, we should not think that sanction must be linked to censure in order to be fitting. In cases of widespread oppression, failure to challenge the moral status quo maintains the harms that marginalized moral agents suffer due to diminished moral consideration. The harms in question are perpetuated by privileged agents' failure to question the moral status quo. The right course of action aims at rectifying situations of oppression.

This is a challenge because oppression and social institutions are closely linked; it is difficult to say what exactly must be done in order to repair harms caused by justified false moral beliefs, but there is an expansive literature on the subject. I will not endorse any particular theory of reparations in this thesis, but it will suffice to say that any reparations must involve attempting to fix abnormal moral contexts by recognizing the moral status of marginalized agents who are systematically denied equal moral consideration. This is the case because without properly fixing

³⁴ Nelkin et. al. 15-18

the moral norms of a moral community, injustice will continue to occur because it does not require that agents willfully do wrong. Privileged agents acting in an “ordinary” way would continually undermine equal moral consideration, a long term goal of reparations.

Section III: Appropriate Self-Censure

I have given an account in which reactive attitudes and blame can be separated from accountability in cases of socially normalized false moral beliefs. I noted that it still seems appropriate for the victims of unwillful wrongdoing to express reactive attitudes towards privileged agents. These reactive attitudes are justified and are valuable in that they indicate that unjustified harms are occurring. The conceptual challenge that I identified about appropriate reactive attitudes concerns conditions of blameworthiness from others. To this point I have considered whether, and under what conditions, others may hold agents responsible for harmful actions that are not connected to an ill will. In bringing this thesis to a close I would like to make a few remarks about what attitudes are appropriate for people to take towards themselves in these cases.

At a first pass, the question of what attitude it is appropriate for a privileged agent to take toward herself might seem to be a nonstarter. We might think that if it is not reasonable for her to believe that she is falling short of obligations, on what basis would she censure herself? My answer is that this basis can be provided by observing the actions of moral exemplars.

Although moral exemplars cannot, and generally do not, blame ordinary people for falling short of obligation, observing them makes us more blameworthy for continuing to do less than what is required of us. Moral exemplars can give us reasons to think that our justified moral beliefs are false or, at least, not so justified. The praiseworthy actions of exemplars who are like

us suggest that we should not be so complacent; we recognize the harms that the exemplar is addressing with her actions, and should recognize that what she is doing would be possible for us as well. In some instances, people feel uneasy when observing the impressive actions of moral exemplars and feel the need to make excuses for their own lack of action. This uneasiness seems to indicate the presence of an internal reactive attitude—shame.

Carbonell argues that shaming might be a productive means of motivating agents to recognize more demanding obligations.³⁵ However, her view only considers shame as a sanction imposed by others. Perhaps the reaction that observers sometimes feel in response to observing moral exemplars is a kind of “self-shaming.” However, I am not convinced that feeling shame is the best, nor the only, way to motivate positive moral development from the observation of moral exemplars. I concede that shame can be a motivating factor to improve one’s actions, but it does not seem like shame leads to the development of appropriate moral dispositions. We can imagine a case where this is apparent. Sam observes a moral exemplar and recognizes that he could do more to alleviate the suffering of others who he had previously failed to appropriately consider as equal moral agents. While he is not shamed by anyone else, he might feel a sense of shame for this wrongdoing. If Sam proceeds to donate a portion of his income, or spend time volunteering, but is motivated by feelings of shame, it does not seem as though Sam is acting in a morally praiseworthy way. We might say that what he did was good, but it does not seem as though we could say that he necessarily developed morally.

We might think that feelings of shame after observing the morally praiseworthy actions of moral exemplars do indicate that an agent developed morally from the observation. On this

³⁵ Carbonell 247

interpretation, the feelings of shame are a symptom of a desire to do the right thing. However, what it takes for moral development is the feeling of shame to become a moral disposition. A feeling has the potential to be temporary, it might be alleviated when an agent thinks that she has done enough to “make up for” the harms caused by her complacency. In order for a feeling to contribute to moral development, an agent must understand the source behind feelings of shame, and resolve to address this out of a sense of duty. In the example with Sam, his self directed shame may well be significant enough to trigger moral development, but in the first instance we should not think that the shame necessarily updates his moral beliefs to the point of consistent action.

In fact, internalized attitudes of shame sometimes fail to motivate action at all. We might worry that agents would manifest what Richard Moran calls the “paradox of self-censure.”³⁶ This paradox arises when an agent comes to recognize that she acted in a morally impermissible fashion. She adopts an attitude of self-censure and responds to herself with negative reactive attitudes. However, the agent believes that the self-censure is itself morally praiseworthy and so she takes a positive attitude toward herself. If we apply this to privileged agents learning that their failure to act constitutes a violation of moral obligations to marginalized people, it can lead to this problematic response to reparations. Imagine the following specific case. Sean is an ordinary person who observes the actions of a moral exemplar. By doing so, Sean receives sufficient evidence to realize that his own actions fall short of obligations he has to consider the moral status of undocumented immigrants. Part of recognizing that this is a failing is the recognition that his former actions were morally impermissible. He feels shame despite the fact

³⁶ Moran, R. (n.d.). 4 Paradoxes of Self-Censure. In *Authority and Estrangement* (pp. 170–183). Princeton University Press.

that his former beliefs were justified by problematic moral norms. Upon reflection, Sean finds his own self-censure morally praiseworthy despite having done nothing to compensate those harmed by his moral failure. This paradox might be further reinforced if Sean recognizes that many of his peers are also failing the obligation he did and that they do not seem to know or feel any shame. This would likely reinforce Sean's praise of himself and make it even less likely that he would act to compensate those that he had harmed.

Ultimately, it seems that although shame might motivate immediate action in particular agents, it is not always warranted or useful. Instead, the appropriate attitude to take towards ourselves after observing moral exemplars is more nuanced. We ought to recognize that we should not blame ourselves for failing to act in accordance with our obligations prior to observing an exemplar, but also recognize that continued complacency afterwards is increasingly blameworthy. Taking this attitude is an appropriate response with practical importance. It is appropriate because it gives fair consideration to the absence of ill will, while recognizing that internally grappling with the problem is not enough to contribute to any substantial reparations. This attitude should encourage privileged agents who are concerned with good moral conduct to lower their credence in the justification for their moral beliefs, be more open to the testimony of marginalized people, and seek out additional examples of moral exemplars.

Chapter 5. Conclusions

It is worth noting that the dialectic of this thesis follows a general conceptual pathway that I advocate each of us take. First, encountering exemplars who are praiseworthy and admirable, and initially seem exceptionally brave or whose actions seem particularly demanding. Second, through continued exposure to these amazing people, we might recognize that they are not that different from us. Third, armed with this realization, we should begin to question our perception of the envisioned sacrifice that holds us back from recognizing our obligations. Fourth, armed with this knowledge, we can identify and challenge moral norms that justify false moral beliefs. Finally, we must find the resolve to reject excuses and act to undermine normalized wrongdoing.

My thesis is broadly applicable to challenging the underlying moral paradigm that justifies agents' false moral beliefs. First and foremost, my argument suggests that privileged agents must accept some epistemic costs when it comes to interpreting what moral beliefs are justified. This entails being skeptical of justification for beliefs concerning the treatment of marginalized people. It means standing with victims even when ignoring the harms that befall them might seem justified.

My thesis also has some implications for any effort that seeks to change the problematic status quo. A simple example of this applies to philanthropy. We should consider philanthropy as a praiseworthy exercise of obligation. Referring to it in this way works against the belief that donating money or time is a necessarily supererogatory endeavor. It simultaneously serves as a challenge to those who are not doing more to help others to consider the moral status of their complacency.

Efforts that are more significant, such as those efforts which aim at securing the equality of rights, are a more challenging case. These tend to involve legal remedies and protections as well as efforts to punish actions that had formerly been unpunished. While my thesis does not directly engage with these actions, I do analyze the efficacy of what I take to be one of the most widely used excuses for wrong moral action—the excuse that normalized false moral beliefs justify wrongdoing. I argue that accountability is warranted even if we accept this excuse. My thesis also suggests that increasing diversity within the Academy and other Institutions that contribute to the social, and subsequently, the moral standing of agents in the community is imperative. There is much work to be done before all members of the moral community have equal moral consideration. Increasing the representation of marginalized people involved in advancing the frontiers of moral knowledge by increasing access to the institutions that uphold moral norms is a vital means of challenging a problematic moral status quo.

Finally, I would like to end with a poignant quote that calls attention to the critical function of moral exemplars for which I have argued. The famous author and social critic James Baldwin once said, “Not everything that is faced can be changed, but nothing can be changed until it is faced.” It is our responsibility to look for moral exemplars, who teach us to look critically at a moral community in need of change.

Bibliography

- Archer, A. (2016). Supererogation, Sacrifice, and the Limits of Duty: Supererogation, Sacrifice, and the Limits of Duty. *The Southern Journal of Philosophy*, 54(3), 333–354.
<https://doi.org/10.1111/sjp.12176>
- Brink, D. O., Nelkin, D. K., & et. al. (2013). Fairness and the Architecture of Responsibility. *Oxford Studies in Agency and Responsibility*, 1, 24.
- Buss, S. (2006). Needs (Someone Else's), Projects (My Own), and Reasons. *The Journal of Philosophy*, 103(8), 373–402.
- Buss, S. (2017). *Some Musings about the Limits of an Ethics That Can Be Applied*. Article Draft.
- Calhoun, C. (1989). Responsibility and Reproach. *Ethics*, 99(2), 389–406.
- Campbell, J. (1993). The Hero Today. In *The hero with a thousand faces* (pp. 387–391). London: Fontana.
- Carbonell, V. (2012). The Ratcheting-Up Effect. *Pacific Philosophical Quarterly*, 93, 228–254.
- Chrisafis, A. (2015, August 22). France Train Attack: Americans overpower gunman on Paris express. *The Guardian*. Retrieved from
<https://www.theguardian.com/world/2015/aug/21/amsterdam-paris-train-gunman-france>
- Darwall, S. (2010). Review: Precis: The Second-Person Standpoint. *Philosophy and Phenomenological Research*, 81(1), 216–228.
- Darwall, S. L. (2002). Chapter: In *Welfare and rational care* (pp. 12–15). Princeton, N.J: Princeton University Press.

- Eisenstein, C. (2013). *The more beautiful world our hearts know is possible*. Berkeley, California: North Atlantic Books.
- Harman, L. (2011). Does moral ignorance exculpate? *Ratio*, 24, 443–468.
- Kagan, S. (2002). Chapters 7-10. In *The Limits of Morality* (Reprinted, pp. 231–403). Oxford: Clarendon Press.
- Levy, L., & Howard, T. (n.d.). *How to Be a Hero*. Retrieved from <http://www.radiolab.org/story/how-be-hero/>
- Moran, R. (n.d.). 4 Paradoxes of Self-Censure. In *Authority and Estrangement* (pp. 170–183). Princeton University Press.
- Neiman, S. (2009). Chapter 12: Theory and Practice. In *Moral clarity* (pp. 372–414). Princeton: Princeton Univ. Press.
- Ricard, M. (2018). *Altruism: the power of compassion to change yourself and the world*. S.l.: Atlantic Books.
- The Carnegie Heroes Fund Commission. (n.d.). Requirements for a Carnegie Medal [Website].
- Urmson, J. O. (1969). Saints and Heroes. In Joel Feinberg (Ed.), *Moral Concepts* (pp. 60–73). Oxford University Press.
- Watson, G. (n.d.). Skepticism about Weakness of Will. In *Freedom, Will, and Agency* (pp. 33–58).
- Williams, B. (1981). Persons, Character, and Morality. In *Moral Luck* (pp. 1–19). Cambridge University Press.
- Wolf, S. (1982). Moral Saints. *The Journal of Philosophy*, 79(8), 419.
<https://doi.org/10.2307/2026228>

