

Our Admiration for Exemplars and the Impartial Spectator Perspective: Moral Exemplarism and Adam Smith's *Theory of Moral Sentiments*

Karsten R. Stueber

College of the Holy Cross

Abstract

This essay will discuss the philosophical viability of Linda Zagzebski's refreshingly radical theory of moral exemplarism that attempts to elucidate the nature of human morality through an analysis of the structure of our admiration for morally exemplary individuals. After raising some systematic worries about exemplarism, I will turn to Adam Smith and his *Theory of Moral Sentiments*. There are indeed strands in Smith's thoughts that contain an exemplarist flavor. Nevertheless, from the Smithian perspective that I favor, our moral concepts emerge from the everyday practice of holding each other morally accountable through empathic perspective-taking. Such a practice is prior to our admiration for the exemplary person. It takes place in the domain of the "ordinary and vulgar", that is, in the domain of the butcher, the brewer, and the baker. Moreover, our normative commitment to the impartial spectator perspective can be revealed as a regulative ideal only in light of an analysis of such practices. Ultimately, what is truly admirable is tied to our commitment to the impartial spectator perspective, whose normative authority should be established independently of our urge to admire, or at least so I am inclined to argue.

Keywords: Exemplarism, Empathy, Admiration, Adam Smith, Impartial spectator.

Moreover, worse service cannot be rendered morality than an attempt be made to derive it from examples. For every example present to me must itself first be judged according to principles of morality in order to see whether it is fit to serve as an original example, i.e., as a model. But in no way can it authoritatively furnish the concept of morality. Even the Holy One of the gospel must first be compared with our ideal of moral perfection before he is recognized as such.

(Kant 1981: 4, 408)

1. Introduction

Throughout history we have admired the works of exceptionally talented people in the arts, the sciences, the humanities, and even in sports. We are also in awe of the extraordinary deeds by ordinary people—such as a policeman sacrificing his life in trying to save a drowning child—the lifestyle of the rich and powerful, or the perceived accomplishments of our political leaders. We admire these individuals because aspects of their lives exemplify features that we hold dear, that we value, and that are part of our ideals in light of which we orient and regulate our own lives. To recognize the socially and morally beneficial nature of such admiration we only need to think about our reverence for inspirational figures such as Mahatma Gandhi, Nelson Mandela, Martin Luther King, and Abraham Lincoln, whose morally exemplary behavior in the fight for justice led to a variety of mass movements and bent the arch of the moral universe towards justice.

Linda Zagzebski has used these intuitions to develop a refreshingly radical theory of moral exemplarism. She claims that it is best to elucidate the nature of human morality by focusing our philosophical attention on the structure of our admiration for morally exemplary persons. Zagzebski also poses a direct challenge to the above Kantian epigraph by turning it on its head. It is not through prior familiarity with moral concepts that we recognize the moral worth of exemplars. Rather it is by being admiringly attuned to them that our moral concepts get content and gain a motivational and normative hold on our agency. As it is well known, Kant's account of morality is often regarded to fall short of answering the question of why it is that moral commands possess a special normative authority and why our recognition of such authority motivates us to act. Kant himself seems to answer these questions by appealing to a mysterious noumenal realm. He thereby violates the widely accepted framework of naturalism according to which the philosophical explication of basic metaphysical, epistemic, and moral features of our lives and the world must be compatible with what the sciences tell us about human nature and the natural world. In emphasizing the emotion of admiration, Zagzebski is more aligned with the ethical and meta-ethical framework proposed by moral sentimentalists who emphasize that moral concepts are in some sense anchored in our emotional reactivity to each other and to the world rather than being grounded in pure reason (Debes and Stueber 2017: Introduction). Like the moral sentimentalists, Zagzebski is open to insights from the empirical sciences and welcomes an empirical investigation of moral agency.

Yet, regardless of how one thinks about the plausibility of Kant's moral philosophy, the epigraph raises a serious question any exemplarist position must answer, that is, how can our admiration for exemplars ground our moral practices if we can identify exemplars only because of a prior understanding of moral concepts. Moreover, the emotion of admiration is a rather double-edged sword since it also has its dark sides morally speaking. We admire persons for all kinds of reasons ranging from rather mundane traits, such as physical prowess, fame, money to intellectually inspiring and morally elevating features such as amazing historical knowledge, oratory skills, or unexpected generosity, integrity, or courage. As Adam Smith already pointed out, admiration is certainly an emotion necessary for the cohesion of society helping us to “maintain the distinction of ranks and the order of society”. Yet he also was wary of admiration for the “rich and powerful” since it constitutes “the great and most universal cause of the corruption of our moral sentiments” (Smith T.M.S. 1976: 61; Irwin 2015). Here one need only think of the

contribution that the admiration for Hitler made in bringing about the catastrophe of World War II and the Holocaust. Currently, the admiration for people like Putin and Erdogan prop up autocracies all over the world. Closer to home, one could argue that the admiration for somebody like Trump constitutes a serious threat endangering the very foundation of American democracy.

In the following, I will critically discuss Zagzebski's exemplarism and investigate whether she can meet the above challenges. In the first section, I will briefly outline the structure of her exemplarist position. In the second section, I will raise three systematic worries about exemplarism before turning my attention to Adam Smith and his *Theory of Moral Sentiments* in the final section. There are indeed strands in Smith's thoughts that contain an exemplarist flavor and raise the same systematic worries as Zagzebski's position. Nevertheless, from the Smithian perspective that I favor, our moral concepts emerge from the everyday practice of holding each other morally accountable through empathic perspective-taking. Such a practice is prior to our admiration for the exemplary person. It takes place in the domain of the "ordinary and vulgar", that is, in the domain of the butcher, the brewer, and the baker. Moreover, it is within the context of an analysis of such practices that our normative commitment to the impartial spectator perspective can be revealed as a regulative ideal. All of this is not to deny that thinking about moral saints is important for our moral life since it reveals that moral action is humanly possible even in extraordinarily challenging circumstances. Yet what is truly admirable is conceptually tied to our commitment to the impartial spectator perspective, whose normative authority should be established independently of our urge to admire, or at least so I am inclined to argue.

2. Zagzebski's Exemplarism: Admiration, the Admirable, and Moral Concepts

Zagzebski weaves an intricate philosophical web consisting of three elementary threads: The notion of exemplars, the analysis of the emotion of admiration, and an externalist and direct theory of reference à la Hilary Putnam and Saul Kripke. In this manner, Zagzebski intends to delineate the complex conceptual landscape of our moral perspective on the world without presupposing a prior conceptual grasp or a normative acknowledgment of moral terminology (117).¹ Most importantly, she wants to ground moral concepts based on "something non-moral" (169).

Exemplars are understood as persons that are at least in some respect "supremely excellent" and therefore "supremely admirable". Additionally, Zagzebski concentrates only on individuals that are exemplary because of acquired excellences rather than natural talents since within the moral realm we have to do with things that are under our control or that could have been otherwise, as Aristotle might express it. While we certainly admire extraordinary natural talents and properties such as perfect teeth, good hair, and a certain height such admiration seems to be of a different type than the admiration for talents that involve some effort in attempting to acquire them. Zagzebski talks specifically about the categories of the hero, who like the Holocaust rescuer is exemplary in showing courage in achieving a moral end; the saint, who shows extraordinary amounts of charity and benevolence; and the sage who, like Confucius, exemplifies the virtue of wisdom.

¹ All page numbers, unless otherwise indicated refer to Zagzebski 2017.

Most significantly, Zagzebski claims that our admiration for individuals tracks their exemplarity without us being able to conceptually articulate why it is that they are exemplary. We are so to speak more certain of their exemplarity than that they are excellent in regard to courage, wisdom, prudence, benevolence, or kindness. From a semantic perspective our access to kinds of moral exemplarity is on par with our access to other natural kinds as suggested by theories of direct reference. We refer to water not because our descriptions of water are necessarily true. Rather our access to water proceeds indexically. It is the stuff to which we are causally exposed in our environment and to which we can demonstratively point as that type of liquid around here. Similarly, the emotion of admiration points us to instances of moral exemplarity and it is through further empirical exploration that we can find out more about its exact nature. Zagzebski mentions specifically narratives, personal experience, but also controlled empirical research as the relevant modes of examination (65ff). Accordingly, when Zagzebski proposes to define value terms such as the notions of virtue, good motive, good end, or good life, and deontic concepts of right, wrong, or duty by referring to exemplars she does not mean to provide us with necessary and sufficient criteria for applying these concepts. In defining virtue as a “trait that makes an exemplar admirable in a certain respect” (113) or a right act as the act that a “person with phronesis [...] would characteristically take to be most favored by the balance of reasons for A in circumstances C” (201), she is quite adamant that such definitions contain an irreducible indexical element (“a person like that”). These definitions presuppose further knowledge gained through the empirical investigation of the lives of exemplars. For this very reason, the moral domain could turn out to be broader than traditionally conceived of since our investigation might make us recognize that intellectual virtues such as epistemic humility or open-mindedness are also traits essential for realizing human exemplarity.

Accordingly, Zagzebski circumvents the Kantian challenge against exemplarism in claiming that we have prior non-conceptual access to exemplars through the emotion of admiration. Emotions for Zagzebski are constituted by an irreducible amalgam of affective, cognitive, motivational, and normative components (Zagzebski 2003, 2015, and 2017: Chpt. 2). Admiring is an appreciative emotion in which we are affectively attuned to somebody, whom we sense to be superior to ourselves, whom we are motivated to be close to, and whose activities we are motivated to imitate. Emotions also have their own unique standards of fittingness and our feeling an emotion makes its object appear to satisfy those standards. In admiring a specific person, we see him or her as being admirable, whereby such seeing cannot be understood as a separate cognitive state that is independent of our admiration. We do not feel admiration because we first judge or perceive another person as admirable in contrast to our seeing ice cream causing a desire to eat it. Rather we only see somebody as being admirable in feeling admiration. A fortiori, our admiration can misfire or can be criticized as being inappropriate because of its inherent appeal to a normative fittingness standard of admirability. It also can be regulated by our reflective capacities. In becoming doubtful about the admirability of the persons whom we admire, our admiration for them diminishes in the same manner that our compassion for the distress of another person might diminish when finding out that the person himself was very much responsible for causing his distress by driving under the influence.

The exemplars that Zagzebski has in mind are thus not only people whom we admire but people who are objectively admirable. Moreover, exemplars are

objectively admirable only if our admiration for them survives a process of continuous and conscientious reflection considering additional information. For instance, if we find out that our trusted companions do not admire them, we become more skeptical about our own emotion and might infer that the individuals whom we admire are objectively not admirable (64). Unfortunately, Zagzebski does not say much about what exact type of information might lead us to withdraw our admiration. She seems to think that we uncover it by further investigating our admiring attitudes towards the world. We know that our admiration of a person has to do with the deep structure of a person's character since we admire a person more deeply if we determine that his action is due to an underlying character trait. Our admiration, on the other hand, diminishes if we realize that a person has been mainly motivated by selfish interests (63ff and 107) since a mere selfish motivation would not distinguish him or her from us ordinary folks. Admiration surviving conscientious reflection should therefore be seen as a reliable standard for judging other people as being admirable. Those judgments provide us with good reasons for imitating and emulating the actions and judgments of our chosen exemplars; an emulation that involves taking up their perspective. In simulating their perspective, Zagzebski suggests, we also acquire the motives and reasons for acting that characterize the exemplar (139-40). To make a long story short, exemplarism promises an elucidation of the moral realm that is naturalistically based, that seems to be able to account for the motivational aspects of our moral judgments, and that, in addition, could provide us with means for improving moral education.

3. Systematic Worries about Exemplarism and its Naturalist Credentials

Zagzebski's exemplarism raises, however, a variety of systematic worries that I fear undermine the very foundation of her position. I will focus here on three of them, which are particularly concerning. First, Zagzebski is rather optimistic that different cultures can find common ground by focusing their attention on exemplars (4). After all, human nature is sufficiently similar so that our emotional capacities are very unlikely to track very different kinds of moral exemplars across cultures (17). At the same time, Zagzebski is suggesting that her proposal is a revisionary and a countercultural one since within modernity we not only admire but also vehemently resent extraordinary accomplishments. Not a day seems to go by in recent years without the saintly status of traditional exemplars being challenged, including the "founding fathers" and even Mother Teresa (see, for instance, Michelle Goldberg, *New York Times*, May 21, 2021).

Zagzebski might respond by arguing that this is just part of our ordinary practice of reassessing the admirability of people in order to determine whether the people who we admire are also genuinely admirable. To be honest I tend to be more skeptical than Zagzebski about the power of reflection to separate the truly admirable from the merely admired. Zagzebski points to how most people view Hitler as a moral monster to suggest that we can distinguish the admirable from the admired in light of the emotional reactions of trusted others. Yet given our evolutionary history, as social creatures we are psychologically predisposed to trust our ingroup more than members who we perceive to belong to the outgroup. Accordingly, we are not naturally committed to what I refer to as the moral stance from within which we treat each other as having equal worth and dignity and as

being morally equidistant of each other. Rather we are profoundly moralizing creatures who endow certain of our norms (including norms of loyalty and purity, see Haidt 2012) with an exalted moral status and as such creatures we favor members of our own group. We also tend to conform in our judgments and our emotional attunement to the social world with members of the ingroup. From this perspective, that most people in the world find Hitler to be a monster might be completely irrelevant for Nazis who do not regard most people as members of their trusted ingroup. It is for this very reason that followers of Trump still admire him and find him admirable, despite acknowledging his many moral failings. What they admire about him is that he projects the resentment of their group and that he wants to “stick” it to the liberal elite.

Even if one is less skeptical about the power of reflection to regulate our admiration in light of a conception of the admirable, the constant reevaluations of our former heroes in contemporary times points in my opinion to a central feature of our practices of admiration and of assessing admirability, that is, its essential cultural and historical relativity. While we certainly should admire Mahatma Gandhi, Nelson Mandela, the Dalai Lama and so on as exemplary human beings in their times and within their cultural traditions, it is not so clear that such admiration carries over easily to contemporary times. Yet, why should they then be regarded as the standards for judging what is morally right and wrong? Moreover, Zagzebski distinguishes among different kinds of exemplars each exemplifying a very specific virtue, that is, courage, charity, and wisdom. If this is so, why should our moral judgments be guided in all domains by how exemplars think about these issues? Does exemplarism really commit us into thinking that Mother Teresa would necessarily have any specific authority to make moral judgments about abortion, the death penalty, or our moral obligations to animals? Alternatively, Zagzebski might appeal to the idea of a perfectly wise and virtuous person or an “idea of exact propriety and perfection” that Adam Smith at times talks about (Smith 1982: 248). Nevertheless, it is ultimately doubtful whether such perfectly wise and virtuous person is an embodied one, a person in flesh and blood whom we would be capable of meeting and admiring or whether such person exist merely in our imagination dependent on a prior grasp of the ideal of moral perfection. Relatedly, we do not merely admire moral exemplars but extraordinary achievements in a wide range of domains of human activity. Admiration then does not naturally limit its scope to the morally admirable. To distinguish the scope of the merely admirable from the morally admirable we could, of course, appeal to our idea of moral perfection. It would, however, imply that Kant’s dictum against exemplarism still stands.

Second, Zagzebski claims to ground morality naturalistically since the emotion of admiration is a natural rather than a supernatural phenomenon. The framework of naturalism certainly discourages a philosophical theory to appeal to supernatural properties. Equally important, however, it encourages philosophers to consider what the sciences tell us about human nature (see also De Caro and Macarthur 2010). A fortiori, a naturalist account of human morality would need to look more closely at whether admiration is a phenomenon that from the perspective of evolutionary and ontogenetic accounts can be seen as the basis of human morality. I am more than skeptical in this respect. Scientifically, admiration is regarded to be a “uniquely human emotion”. It is particularly an emotion of societies where rank differences are based on so-called prestige hierarchies rather than dominance hierarchies, which one finds among chimpanzees and which

are imposed by threat, aggression, and mediated by fear (Onu et. al. 2016: 217 and 223; Seetman et. al. 2013). Admiration thus presupposes the foundation of a special form of human social cooperation dependent on the enforcement of social and moral norms (Tomasello 2019, Boehm 2012, Wrangham 2019) as it is only within this somewhat more egalitarian spirit that social differentiations among humans are formed. Ontogenetically, children from the age of three years old are already able to distinguish between conventional and moral norms (Smetana et. al. 2014). They know that not hitting another person is a norm that does not depend on social agreement or social authority and recognize that it would not be ok to hit another person, even if their teacher tells them otherwise. Yet only later do children develop an understanding of the category of the supererogatory, which is the basis for the emotion of moral admiration. In the beginning of their moral awakening, they are fully focused on what is obligatory rather than what is admirable or supererogatory (Dahl et. al. 2020).²

Zagzebski acknowledges as much when she says that a human society could not exist without a shared sense of what constitutes an intolerable act (192ff). We would also have to assume that the sense for the intolerable is enforced among members of a society and that it would be backed up by humans being emotionally very sensitive to the violation of the norms of the intolerable. Zagzebski refers to such moral sensibility as “morality light”. I am a bit perplexed why one would call the basis for our social existence derogatively “morality light”. Moreover, I assume that American society would be in a much better shape if people would at least abide by the norms of the intolerable (and refrain from constantly shooting each other). Even more puzzling is the fact that Zagzebski insists that the category of the morally intolerable, of the morally wrong and of moral duty, is determined in respect to what “exemplars cannot tolerate”. Given the forgoing considerations, the reference to exemplars seems to be rather superfluous. Even without the existence of any saints we seem to know perfectly well what is intolerable. Moreover, we would be in no position to be sensitive to what is truly extraordinary and admirable without first having acquired knowledge of what is morally intolerable.

Philosophers are however not merely interested in providing a causal explanation of why it is that humans are normative animals and distinguish between moral and conventional norms. Philosophers are ultimately interested in explicating why we ought to be moral. They want to explain why it is that moral commands have a unique normative authority over us even though their validity does not depend on the particular social practices that we are part of. Exhortations and judgments such as don't be cruel, or slavery is wrong are understood as having universal validity. They do not address us in our particularity as Americans, Germans, or Chinese. They speak to us as human beings from the perspective of the moral stance where we possess equal dignity and value and leave behind the framework of mere personal relations. On behalf of Zagzebski one might argue that in focusing on admiration and admirability, she primarily wants to address the above normative question that is central for a philosophical explication of the moral realm. One could then admit that a conception of the intolerable is causally basic for the functioning of a society without admitting that such conception is also normatively foundational. From this perspective, admiration is motivating

² I was made aware of this research through a talk by Christina Starmans at the 2021 conference of the Society for Philosophy and Psychology, where she presented new and yet unpublished results of experiments that supports this developmental picture.

us to imitate the person whom we admire, and our judgment of admirability (as the result of such admiration surviving a process of conscientious reflection) provides us with normative reasons for imitating such persons.

Yet one wonders whether admirability in its most general form can adequately ground the normative authority of morality. When I was growing up in Germany in the 1970ties, every boy admired the soccer star Gerd Müller. He was an amazingly effective striker (a classic number 9), who was a member of the German national team that brought home the soccer world cup for a second time in 1974 when the championship was played in Germany. Trust me, a lot of boys at the time tried to be Gerd Müller and they had good reasons to do so. After all, he was a truly admirable striker. Nevertheless, it also tended to be perfectly clear to us that while we had all the reasons in the world to imaginatively enact being Gerd Müller, these were not sufficient reasons to become Gerd Müller in real life. There are indeed more important things to do than playing soccer. Such merely optional reasons however are not the reasons that we are after in trying to normatively ground our commitment to morality. Morality does not seem something that is merely optional for us, like becoming a soccer player. Yet why should admiration and admirability of moral exemplars be different than my childish admiration for Gerd Müller? Why does it mean that I have to take the judgments of moral exemplars more seriously in real life? Pointing out that in this case we encounter the moral kind of admirability seems to beg the very question that we are asking of how admirability normatively grounds morality.

To some extent, Zagzebski acknowledges the above points (see 169-70) but dismisses them in that she asserts that the emulation of moral exemplars proceeds via taking up their point of view and simulating their reasons for acting (136-39 and 170). It is exactly in this respect, one could argue, that admiration for the character of a person differs from our admiration of her skills or accomplishments. Yet even if we grant that the admirability for whole persons provides us with reasons for taking up a person's point of view (rather than merely trying to imitate their external behavior)—and Zagzebski never fully explains why this is so—it is not clear why such simulation changes the above equation. Properly understood, imaginatively taking up another person's point of view does not mean that I become the other person, that is indeed a conceptual impossibility (see Goldie 2011). Empathically taking up another person's point of view means that I am at the same time aware of the fact that it is not my perspective that I am simulating. This is particularly true in situations in which I and the other person are otherwise quite different such as is the case with every normal person and the exemplar they admire. Why then should the reasons or motivations of the admired person automatically become my reasons or motivations for acting? Certainly, taking up the perspective of Gerd Müller and reenacting his reasons for becoming a soccer player does not automatically imply that those reasons should be my reasons for acting, even if I admire him as a soccer player. The reasons of the exemplars must therefore be of a very different kind. We might be tempted to say that this is so because they are moral reasons. But such an answer is very much question begging. Accordingly, it is high noon to turn, as promised, to a discussion of Adam Smith. As I interpret him, we should not conceive of Smith as an exemplarist, even if some of his arguments for the impartial spectator perspective at times contain an exemplarist flavor. Most importantly, Smith allows us to weave the various elements that Zagzebski so rightly appeals to in her exploration of the moral

realm—that is, simulation, admiration for exemplars, the basic sense of the intolerable—into a more plausible map of our moral life.

4. Adam Smith, Empathy, and the Impartial Spectator: How to Acknowledge Exemplars without Being Committed to Exemplarism

Let me start my brief exploration of Adam Smith's conception of the moral realm by acknowledging that his conception of virtue is an ambivalent one and that the centrality of it for his moral philosophy has also been disputed interpretive territory. As Smith is one of the preeminent philosophers thinking about human morality within the context of modern commercial society, this fact is not that surprising since the ancient notion of virtue was a controversial one in the modern context. As it is well known, some political and moral philosophers, like Machiavelli and Mandeville, took a decidedly negative even if nuanced view in this respect (see for example Messina 2017). And they were at times quite happy to let moral hypocrisy rule and allow the "invisible hand" take care of the rest, supposedly creating a buzzing, creative, and rich society from which all of us could benefit.

Smith clearly does not belong to this category of thinkers, despite some of his interpreters being puzzled by the so-called Adam Smith Problem, that is, of how to reconcile his *Theory of Moral Sentiments* (TMS) with the *Wealth of Nations*. He understood modern commercial society from a moral point of view both as an opportunity to expand our moral horizon and as a challenge for the education of our moral sentiments. Scholars have, however, been divided in their judgment about how central Smith takes the notion of virtue to be for his account of our moral life and the foundation of our moral judgments. Generally, it has been regarded to play a secondary even if important role since part VI of TMS, "Of the Character of Virtue", was only added to the sixth edition. Only recently has it been suggested that we should read Smith as being closely aligned with ancient and Christian virtue theory even if adjusted for the modern commercial society (Hanley 2009). Moreover, while Smith talks about virtues throughout the book, he uses the notion of virtue in TMS in a decidedly ambiguous manner. On the one hand, Smith seems to allow for the fact that virtue is achievable for most human beings in ordinary circumstances, what Smith also calls the "middling and inferior stations of life" (Smith 1982: 263). Accordingly, Charles Griswold (1999: 13) views Smith mainly as a philosopher defending the "middling human virtue". On the other hand, Smith at times favors a notion of virtue understood as extraordinary human excellence—something that we admire in that our sense of approbation is "heightened by wonder and surprise" (Smith 1982: 20). The paragon of such virtue is the "wise and virtuous man" whose conduct and judgment are not only oriented at the "idea of exact propriety and perfection" but who also fully comprehends that human nature allows at most for an approximation to such an ideal (Smith 1982: 247-48). The wise and perfect man (whose virtues include both ethical and intellectual virtues) in Smith is best seen as a person whose perspective embodies the ideal of the impartial spectator and who therefore also possesses sufficient humility. He recognizes that, even if he is superior in virtue to individuals in the middling and ordinary stations of life, he is ultimately "but one of the

multitude in no respect better than any other in it” (Smith 1982: 83 and 137).³ Insofar as we admire such virtuous person we not only agree with his judgments, but those judgments also “lead and direct” our own (Smith 1982: 20).⁴

However, the exploration of the psychological mechanisms with the help of which we hold each other normatively accountable is at the very heart of Smith’s elucidation of the moral realm.⁵ As he expresses it (Smith 1982: 111), “a moral being is an accountable being”, that is, “a being that must give account of its action to some other”. In holding each other accountable we do not judge an action to be right and wrong independent of an agent’s reasons for acting. Smith strongly objects to Hume who regards “utility or hurtfulness” (Smith 1982: 188) as the primary principle of judging the appropriateness of an action. Indeed, we still blame a person if he has done the right thing for the wrong reasons. Think in this context about an agent who pulls the lever in the famous Trolley case (saving 4 people and letting one other person die in the process) but only because he wanted to get rid of a serious competitor for a job or an award. From a utilitarian perspective we still could judge the action to be the right one. Yet it certainly does not possess any moral worth and the agent is morally blameworthy.

Most importantly, Smith is relevant to contemporary metaethics because he views our practice of holding each other morally accountable as being based on psychological capacities necessary for the constitution of the social realm within which humans live and cooperate. Normative distinctions and normative judgments emerge as effects of our ability to mutually empathize with each other’s thoughts and sentiments. Humans as social creatures are constituted so that they cherish being empathized with. Smith understands such empathy—or what one called sympathy in the 18th century—as imaginative perspective-taking, as putting oneself in another person’s point of view and simulating the manner in which that person thinks about the situation that he has to respond to. While Smith certainly differs from Hume in his conception of the concrete mechanisms of empathy (Stueber 2015), he agrees with him that empathy allows the “minds of men” to be “mirrors to one another” (Hume 1978: 365). We mirror the other person’s thoughts and

³ In this respect Smith’s ideal of the “wise and virtuous man”, even if very much inspired by ancient and Christian philosophers is very much a creature of the modern commercial and cosmopolitan world. I am not so sure how I would classify Zagzebski’s notion of exemplars in this respect as she points to Confucius as the paradigmatic sage.

⁴ My remarks in these two paragraphs have greatly benefitted from the insightful interpretations of Fleischacker 2013, Hanley 2013, and Schliesser 2017 (particularly chpt. 9).

⁵ In Part VII Smith claims that moral philosophy generally addresses two questions: “First wherein does virtue consist in? [...] And secondly, by what power or faculty of the mind is it, that this character, whatever it is, is recommended to us? Or in other words, how and by what means does it come to pass, that the mind prefers one tenour of conduct to another, denominates the one right and the other wrong; considers the one as the object of approbation, honour and reward, and the other of blame censure and punishment” (Smith: 265). As I read Smith, the first question is the one that he addresses in parts VI and VII in situating himself within traditional virtue theory. Within the context of modernity and its skepticism about the normative domain the second question is, nevertheless, the philosophically foundational one. Accordingly, Smith addresses it in the first sections of the *Theory of Moral Sentiments*. Independent of the question of interpretive accuracy, only in this manner can we understand Smith as providing us with a plausible foundation of morality within the contemporary metaethical context committed to the naturalist framework. See also Stueber 2017. I further elaborate on how to use Smith within the contemporary context in my book manuscript *The Moralizing Animal* (under contract with MIT Press).

sentiments in taking another person's perspective by bringing the other's thoughts and sentiments "home to ourselves", as Smith is fond of expressing it. Equally important though, in resonating with the other person we also hold up a mirror that allows that person to become aware of his thoughts and sentiments as something for which he can be held normatively accountable. Only in a social context is a human being provided with a "mirror which he wanted before" and which allows him to think of "his own character, of the propriety or demerit of his own sentiments and conduct, of the beauty or deformity of his own mind" (Smith 1982: 110).

More specifically, Smith suggests that our ability to reenact another person's thoughts and sentiments is directly tied to judging the propriety and the merit of his actions or even the propriety of his sentiments themselves. Simplistically expressed, our ability to reenact a person's sentiments by taking her perspective leads us to approve of them and to judge her actions to be appropriate or to possess merit. Our inability to do so leads us to disapprove of their actions. For our purposes, Smith's description of the exact and intricate part of the mechanisms leading to such approval is of secondary importance. For reasons I have explicated elsewhere (Stueber 2017), it is best to understand reenactment of another person's sentiments as grasping their thoughts as their reasons for acting (see also Stueber 2006). Such reenactment might lead to our approval since in understanding another person's thoughts as reasons for actions I view them as considerations that from her perspective speak for her actions. If I can indeed bring such thoughts "home to myself" in recreating them in my mind, I can then also understand them as considerations that could be my reasons for acting. *A fortiori* it seems that I myself would then approve of such actions or sentiments.

Yet such approval seems to be rather subjective or at most an approval that reflects the social norms of a particular group, culture, or society. We generally tend to listen to these merely subjective judgments because we all like to be liked by the people we live with. Nevertheless, this fact cannot explain why they have the authority of the moral stance from which we make demands that are normatively binding to all human beings regardless of what group or culture they belong to. That is, it is not at all clear why we ought to take a person's approval and disapproval seriously based on his ability or inability to reenact our thoughts. Ultimately such ability and inability might merely reflect certain limits in a spectator's empathic capacity rather than a moral defect in our agency. As it is well known, Smith attempts to address these concerns by referring to the perspective of the impartial spectator. For him, any evaluative judgment based on the ability of an impartial spectator to empathize with an agent's sentiments provides that agent with a normative and moral reason for taking that judgment seriously.

Here I do not want to spend much time discussing how exactly we should characterize the perspective of the impartial spectator. As it is commonly understood, Smith's impartial spectator is not an omniscient one nor is he a person devoid of normal human emotions. Rather he is a spectator who is removed from the immediate heat of the action, knows all the relevant facts of the circumstances (as they might be accessible to an agent who is diligent enough to pay attention), and has no selfish interest in the outcome of the action. If I am allowed one more soccer analogy: The impartial spectator could be compared to a soccer fan who watches a game on TV between teams whom he normally does not cheer for, just for the enjoyment of the game. It is a person who is emotionally attuned to watching soccer, who knows the game and the emotions it can elicit, but who is one

step removed from being really interested in any of the teams winning.⁶ More importantly for my purposes, however, is the question of why it is that the judgments from such a stance have a special normative authority to make demands on us, why it is that we should take them seriously, or why our actions can be justified only if they can gain approval from that perspective.

One can find two strategies within the Theory of the Moral Sentiments to answer this question. The first is the more obvious and official one (see in this respect also Griswold 1999: 129ff). In light of our discussion of Zagzebski, one could also call it the exemplarist strategy. To motivate the need for the impartial spectator perspective Smith appeals to our experience of being judged wrongly by our peers because they do not fully understand all the relevant factors (including my own mental states) of the situation. The experience of such discordance makes us aware of the fact that we do not merely desire to be praised but that we want such praise to be accorded to us because we are praiseworthy. Besides a desire for praise, human beings are also motivated by a desire of praiseworthiness, a desire that Smith regards to be “by no means derived altogether from the love of praise” (114). It is exactly in this context that Smith refers to our admiration of virtuous exemplars since he sees the

love and admiration which we naturally conceive for those whose character and conduct we approve of, necessarily dispose us to desire to become ourselves the objects of like agreeable sentiments, and to be as amiable and as admirable as those whom we love and admire the most. Emulation, the anxious desire that we ourselves should excel, is originally founded in our admiration of the excellence of others. Neither can we be satisfied with being merely admired for what other people are admired. We must at least believe ourselves to be admirable for what they are admirable. But, in order to attain this satisfaction, we must become the impartial spectator of own character and conduct (Smith 1982: 114).

Without doubt Smith’s account of our desire for praiseworthiness (that is our desire to be praised from the perspective of the impartial spectator) has a very exemplarist flavor. For that very reason, it also encounters all of the philosophical worries that we talked about in the last section. Ultimately, Smith regards persons to be virtuous to the highest degree because they are embodying the impartial spectator perspective. They will thus also be praised from that perspective. Accordingly, our admiration for the “wise and virtuous” enables us to causally explain the desire for praiseworthiness. The central philosophical question that we try to answer is, however, not a causal one but a normative one. We want to know why we should have the desire for praiseworthiness and why we should accept the perspective of the impartial spectator perspective as having normative authority for the evaluation of our character and actions. Or to ask the question differently, if truly excellent people embody the impartial spectator perspective, why does that fact make them admirable? Pointing to our admiration to these exemplars as an answer appears to be begging the question.

Smith’s text, however, allows us to reconstruct a philosophically more promising strategy for answering the normative question. For that purpose, I take my departure from Smith’s conception of the impartial spectator perspective as reason or principle (Smith 1982: 137). It is the highest tribunal to which we implicitly have

⁶ I think this is a better analogy than a comparison to the referee of the game as that person is still too close to the “action” on the field.

to appeal in order to negotiate the comparative strength of our reasons for actions (Smith 1982: 128ff) within our practice of mutual empathic perspective-taking.⁷ Central to my argument is the fact that for us to properly simulate another person's perspective we have to take into account differences between us and the target of our empathy. We have to imaginatively adopt the attitudes that we do not share with the other person and quarantine our own attitudes that the other person does not share with us for our reenactment to provide us with reliable insights into the other person's mind. Yet, and here we have to be a bit more careful than Smith (and also Zagzebski), bringing another person's case home to myself in this manner does not automatically constitute approval of his actions since simulating his reasons does not automatically mean that they would be reasons I would act on in his situation. In recreating his perspective, I am at the same time aware of the fact that our perspectives on the world differ in relevant respects. I recognize his thoughts as potential reasons that I would act on only if my own perspective would also undergo relevant changes. It is exactly in this situation, however, that our reenactment of another person's reasons addresses us as a critical, reflective, and therefore self-critical reasoner. Reenacting another person's perspective and his reasons makes a demand on us that requires a rational response. It demands an answer to the question of why it is that we do not make his perspective our perspective, given the fact that his reasons are perfectly intelligible to us. And it is exactly in this context that we implicitly appeal to the normative authority of the perspective of the impartial spectator within which we conceive of ourselves as equal reasoners, or so I would like to argue. The impartial spectator perspective as the "highest tribunal" within which we adjudicate between our reasons for acting, is therefore neither "our own place nor yet his". It is a stance where we look at our reasons "neither with our own eyes, nor yet with his, but from the place and with the eyes of a third person, who has no particular connection with either and who judges with impartiality between us" (Smith 1982: 135).

To fully understand the demand that the reenactment of another person's reasons makes on us, it is important to grasp that in reenacting another person's perspective we reenact a holistic web of attitudes within which a person's thoughts constitutes a reason. Moreover, as already Aristotle understood, our reasons tend to be hierarchically organized. Not only do we have first-order reasons we also have reasons for having those reasons. We not only recognize that somebody likes a neat office. Such recognition would indeed not put much pressure on us to change our messy ways of "taking care" of our office. Additionally, we can recognize that the other person has a reason for keeping his office neat such as that cleanliness is next to godliness. In imaginatively taking up another person's point of view, we ultimately reenact a differently structured framework of reasons. I would suggest that in this manner we enlarge our own possibilities of conceiving of rational agency and of considerations that could count as reasons for acting. In reenacting them in our own mind, we imagine them as reasons we can "live" by, that we might feel at home with. Such reenactment ultimately sharpens our sensitivity to our common humanity as rational agents in our local distinctiveness. It is a sensitivity that does not yet constitute full approval. It constitutes a somewhat appreciative engagement with the "vitality" and "life potentiality" (Lipps 1903) that lies in the reenacted perspective.

⁷ More specifically, the passages that I find in this context most interesting were taken out by Smith for the sixth edition or were parts of a draft. They were added by the editor for the Glasgow editions of his works. See Smith 1982: 128-30.

Appreciating another person's perspective in this manner has a very positive valence when we try to reenact a Buddhist perspective with its emphasis on sympathy. It might however also resonate with us in a negative and almost scary manner such as when we try to reenact the perspective of a Holocaust perpetrator.

To conclude my discussion of contemporary exemplarism within the context of Smith's Theory of Moral Sentiments, I admit that our recognition and normative acknowledgment of the impartial spectator perspective causally involves a quasi-aesthetic dimension, an appreciative sensitivity, and an appreciative grasp of the intricacies of another person's point of view. Such appreciative component allows us to grasp the strength of another person's reasons requiring us to call for a normative judgment from the perspective of the impartial spectator. I would also acknowledge that appreciative emotions like admiration, but also awe and reverence, for persons with extraordinary achievements at times facilitate our understanding of another person's reasons. Admiration can prime us to think highly of another person's point of view even before we fully engage in simulating that person's perspectives. That probably is a good thing if we admire moral exemplar. Yet as already Smith pointed out, it can also contribute to moral corruption if the wrong person is admired. Admiration for moral exemplars thus should be thought of as being able to play a role in moral education if properly constrained by the impartial spectator perspective. But Kant still seems to have a point in claiming that reference, even an admiring one, to exemplars cannot ground the conceptual framework for our moral life. For that purpose, it is best to follow Smith's analysis—or my favorite reading thereof (Stueber 2017)—of how we hold each other accountable among rather ordinary folks as such practices implicitly commit us to the ideal of the impartial spectator perspective.

References

- Boehm, Chr. 2012, *Moral Origins: The Evolution of Virtue, Altruism, and Shame*, New York: Basic Books.
- Chalik, L. and Rhodes, M. 2020, "Groups as Moral Boundaries: A Developmental Perspective", in Benson, J.B. (ed.), *Advances in Child Development and Behavior*, vol. 58, Cambridge, MA: Academic Press, 63-93.
- Dahl, A., Gross R.L. and Siefert, C. 2020, "Young Children's Judgment and Reasoning about Prosocial Acts: Impermissible, Suberogatory, Obligatory, or Supererogatory?", *Cognitive Development*, 55, 3, <https://doi.org/10.1016/j.cogdev.2020.100908>
- Debes, R. and Stueber, K. 2017, *Moral Sentimentalism*, Cambridge: Cambridge University Press.
- De Caro, M. and Macarthur, D. 2010, *Naturalism and Normativity*, New York: Columbia University Press.
- Fleischacker, S. 2013, "Adam Smith and Equality", in Berry, Chr., Paganelli, M.P. and Smith, C. (eds.), *The Oxford Handbook of Adam Smith*, Oxford: Oxford University Press, 484-500.
- Goldie, P. 2011, "Anti-Empathy", in Coplan, A. and Goldie, P. (eds.), *Empathy: Philosophical and Psychological Perspectives*, Oxford: Oxford University Press, 302-17.
- Griswold, C. 1999, *Adam Smith and the Virtues of Enlightenment*, New York: Cambridge University Press.

- Haidt, J. 2012, *The Righteous Mind: Why Good People are Divided by Politics and Religion*, New York: Vintage Books.
- Hanley, R.P. 2009, *Adam Smith and the Character of Virtue*, Cambridge: Cambridge University Press.
- Hanley, R.P. 2013, "Adam Smith and Virtue", in Berry, C., Paganelli, M.P., and Smith, C. (eds.), *The Oxford Handbook of Adam Smith*, Oxford: Oxford University Press, 219-40.
- Hume, D. 1978 [1739], *A Treatise of Human Nature*, Oxford: Clarendon Press.
- Irwin, T.H. 2015, "Nil Admirari: Uses and Abuses of Admiration", *Proceedings of the Aristotelian Society, Supplementary Vol.*, 89, 223-48.
- Kant, I. 1981, *Grounding for the Metaphysics of Morals*, Indianapolis: Hackett.
- Lipps, Th. 1979 [1903], "Empathy, Inner Imitation and Sense-Feelings", in Rader, M. (ed.), *A Modern Book of Esthetics*, New York: Holt, Rinehart and Winston, 374-82.
- Messina, J.P. 2017, "Kant, Smith, and the Place of Virtue in Political and Economic Organization", in Robinson, E. and Surprenant, C.W., *Kant and the Scottish Enlightenment*, London: Routledge, 267-85.
- Onu, D., Kessler, T., and Smith, J.R. 2016, "Admiration: A Conceptual Review", *Emotion Review*, 8, 218-30.
- Schliesser, E. 2017, *Adam Smith: Systematic Philosopher and Public Thinker*, Oxford: Oxford University Press.
- Seetman, J., Spears, R., Livingstone, A.C., and Manstead, A.S.R. 2013, "Admiration Regulates Social Hierarchy, Dispositions, and Effects on Intergroup Behavior", *Journal of Experimental Social Psychology*, 49, 534-42.
- Smetana, J.G., Jambon, M., and Bell, C. 2014, "The Social Domain Approach to Children's Moral and Social Judgments", in Killen, M. and Smetana, J.G. (eds.), *Handbook of Moral Development* (2nd edition), New York: Psychology Press, 23-45.
- Smith, A. 1982 [1759], *The Theory of Moral Sentiments* (Raphael D.D. and Macfie A.L. eds., The Glasgow Edition of the Works, and Correspondence of Adam Smith), Indianapolis: Liberty Fund Press.
- Stueber, K. 2006, *Rediscovering Empathy: Agency, Folk-Psychology and the Human Sciences*, Cambridge, MA: MIT Press.
- Stueber, K. 2015, "Naturalism and the Normative Domain: Accounting for Normativity with the Help of 18th Century Empathy-Sentimentalism", *Rivista Internazionale di Filosofia e Psicologia*, 6, 24-36.
- Stueber, K. 2017, "Smithian Constructivism: Elucidating the Reality of the Normative Domain", in Debes and Stueber 2017, 192-209.
- Tomasello, M. 2019, *Becoming Human: A Theory of Ontogeny*, Cambridge, MA: Harvard University Press.
- Wrangham, R. 2019, *The Goodness Paradox: The Strange Relation between Virtue and Violence in Human Evolution*, New York: Vintage Books.
- Zagzebski, L.T. 2003, "Emotion and Moral Judgment", *Philosophy and Phenomenological Research*, 66, 104-24.
- Zagzebski, L.T. 2010, "Exemplarist Virtue Theory", *Metaphilosophy*, 41, 41-57.
- Zagzebski, L.T. 2015, "Admiration and the Admirable", *Proceedings of the Aristotelian Society, Supplementary Vol.*, 89, 205-21.
- Zagzebski, L.T. 2017, *Exemplarist Moral Theory*, Oxford: Oxford University Press.