

Georgia State University

ScholarWorks @ Georgia State University

---

Philosophy Theses

Department of Philosophy

---

5-13-2021

## Religion and Two-Level Utilitarianism in Adam Smith's The Theory of Moral Sentiments

Steven O. Falco  
*Georgia State University*

Follow this and additional works at: [https://scholarworks.gsu.edu/philosophy\\_theses](https://scholarworks.gsu.edu/philosophy_theses)

---

### Recommended Citation

Falco, Steven O., "Religion and Two-Level Utilitarianism in Adam Smith's The Theory of Moral Sentiments." Thesis, Georgia State University, 2021.  
[https://scholarworks.gsu.edu/philosophy\\_theses/296](https://scholarworks.gsu.edu/philosophy_theses/296)

This Thesis is brought to you for free and open access by the Department of Philosophy at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Philosophy Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact [scholarworks@gsu.edu](mailto:scholarworks@gsu.edu).

Religion and Two-Level Utilitarianism in Adam Smith's *The Theory of Moral Sentiments*

by

Steven Falco

Under the Direction of Eric Wilson, PhD

A Thesis submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2021

## ABSTRACT

The questions concerning Adam Smith's religious views and his purported support for utilitarianism have each generated a substantial literature. In this thesis, I propose a response to the first of these problems which I believe also helps resolve the second one. First, I argue that Smith's references to God in *The Theory of Moral Sentiments* reflect his sincere theological commitments, not merely his account of the psychology of religious belief or his attempt to avoid charges of atheism. I then show that Smith conceives of God as a utilitarian creator who designed our moral sentiments so that they would most effectively promote the utility of humanity as a whole. I combine these two claims to develop an interpretation of Smith as a two-level utilitarian where the Impartial Spectator determines the morality of the lower level and God does so for the higher one.

INDEX WORDS: Adam Smith, Impartial Spectator, Two-Level Utilitarianism, Moral Psychology, Moral Sentimentalism, Religion

Copyright by  
Steven Falco  
2021

Religion and Two-Level Utilitarianism in Adam Smith's *The Theory of Moral Sentiments*

by

Steven Falco

Committee Chair: Eric Wilson

Committee: Eddy Nahmias

Electronic Version Approved:

Office of Graduate Services  
College of Arts and Sciences  
Georgia State University  
August 2021

## ACKNOWLEDGEMENTS

I would first like to thank my thesis advisor Professor Eric Wilson, who taught the excellent course Adam Smith's moral theory which inspired this paper and whose ongoing feedback on my work has been crucial for the completion of this project. Additionally, I greatly appreciate my committee member Eddy Nahmias, and my friends and fellow graduate students Connor Kianpour and David Simpson for the time they spent reading and commenting on drafts of this thesis. I also benefited from discussions with my friend Brad Marston, who taught me most of what I know about utilitarianism. I also want to express my gratitude to Marianne Janack, my first philosophy Professor at Hamilton College, who encouraged me to major in philosophy and pursue it in graduate school. Above all, the biggest thank you to my Dad for sparking my passion for philosophy and to both of my parents for their love and support throughout this whole process.

**TABLE OF CONTENTS**

<b>ACKNOWLEDGEMENTS .....</b>	<b>IV</b>
<b>1 INTRODUCTION.....</b>	<b>1</b>
<b>2 UTILITY AND THE IMPARTIAL SPECTATOR.....</b>	<b>3</b>
<b>3 GOD, HUMAN NATURE, AND INDIRECT UTILITARIANISM.....</b>	<b>10</b>
<b>4 RESPONDING TO OBJECTIONS ABOUT RELIGION.....</b>	<b>15</b>
<b>5 RESPONDING TO OBJECTIONS ABOUT UTILITY .....</b>	<b>19</b>
<b>6 CONCLUSION .....</b>	<b>25</b>

## 1 INTRODUCTION

Adam Smith is often credited with developing a sentimentalist ethical theory in his work *The Theory of Moral Sentiments* (1790), in which an action's moral worth depends on whether it would garner the sympathy of an informed but unbiased individual called the "impartial spectator". The impartial spectator here is not meant to be a real person but rather is an imaginative construct we can call upon to gain insight about what we and others ought to do. Smith believes considering the reaction of an impartial spectator allows us to abstract from our particular perspective and properly judge actions. Smith's system departs from those of many other philosophers because it takes our actual moral emotions, and not theoretical, rationally-derived principles, as its starting point. Still, as we shall see, Smith at certain points strongly indicates that he believes that an almighty Deity designed our moral faculties, and that it did so primarily to maximize human happiness. In my paper, I argue that these tenets do not stand in tension with his sentimentalism, but that they serve different roles in his moral system.

Specifically, I interpret Smith as embracing two-level, also known as indirect, utilitarianism, a theory which holds that the best world is the one which contains the greatest total amount of happiness, but that aiming directly at maximizing aggregate utility is often an ineffective way of achieving that end. Richard Hare's original version of this view argued that the utilitarian level of decision making was appropriate for some circumstances but that we should rely on various intuitions and moral rules instead at other times<sup>1</sup>. Smith largely reserves the utilitarian level of decision making for God, who acts only in order to promote the general welfare, and recommends the impartial spectator as our best means of approximating God's

---

<sup>1</sup> Price, Anthony, "Richard Mervyn Hare", *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition), Edward N. Zalta (ed.),



judgment. Because Smith defines morality in terms of what is the proper target of our positive moral sentiments, the content of morality is equivalent to the sum total of the impartial spectator's decisions. However, we only ought to rely on this spectator because we cannot perceive the full consequences of our actions, and following the impartial spectator allows us typically to produce beneficial results despite our imperfect knowledge. One way to understand the thesis I am arguing for here is to say that propriety establishes our morality but utility justifies it. Admittedly, this view doesn't make Smith utilitarian in a traditional sense, but it means that he leans much further in that direction than he appears to at first.

While other scholars have mentioned this interpretation as a possible resolution to the disagreements surrounding Smith's stance on these questions, I am not aware of any previous effort to fully develop this approach and defend both its religious and ethical sides<sup>2</sup>. To do this, I begin by presenting some of the passages on utility that some people have taken as evidence that his views on the topic are inconsistent. Then I show why we should take Smith's religious belief seriously and explain how it helps form the foundation for my understanding of his ethical theory. Finally, I respond to objections which present Smith as non-theological and those which conclude he is an anti-utilitarian.

First, it is worth prefacing my interpretation of Smith by comparing it to a reading of Joseph Butler, Smith's predecessor and influence. According to Robert Louden's article "Butler's Divine Utilitarianism" (1995), Butler believes that the standard that God uses to

---

<sup>2</sup> R.F. Stalley, for example, writes in "Adam Smith and the Theory of Punishment", that Smith's position on utility "is in keeping with his general view that nature is providentially ordered. Its 'Author' intends the happiness of mankind but wisely has not left us with the task of working out how to achieve this. He has, instead, endowed us with natural instincts which lead us to act in ways likely to produce our happiness." (p. 7)

determine the moral truth sometimes diverges from the process by which humans ought to make moral judgments. Specifically, Louden contends that Butler's apparent deontological commitments only describe our moral psychology and actually cloak his deeper philosophical support for utilitarianism. As Louden writes, "Butler's ethics is best viewed as a theological version of what has been called 'oblique' or 'indirect' utilitarianism, according to which human beings were created by God in such a manner (and for the reason) that they will produce the greatest amount of happiness by not trying to produce the greatest amount of happiness." (p. 267) Because we are fallible creatures, we often fail to bring about our intended ends, whereas God's perfection ensures that every goal he has necessarily becomes reality. God can therefore foresee how our attempts to maximize utility may backfire and how non-utilitarian tendencies can counter-intuitively lead to the best possible consequences. I see Smith as advancing a similar, albeit somewhat more moderate, line of argument, by describing our moral psychology as largely but not entirely utilitarian and accounting for the deviations in the way that Butler does. If Louden is right about Butler, the type of reasoning I ascribe to Smith would at least have some historical precedent and therefore would gain additional plausibility by helping fit him into the philosophical thinking of his time.

## **2 UTILITY AND THE IMPARTIAL SPECTATOR**

One potential issue with relying on the impartial spectator as a moral standard is that, as Smith points out, it can contradict itself. This raises the question of which of its judgments we ought to favor when it does so. I will argue that, in at least some important cases, Smith asks us to decide by figuring out which will maximize utility. Something's utility, for Smith, refers to the

totality of positive consequences it will produce. An object or course of action has great utility if it tends to improve people's lives in the long run, and lacks utility if it has no effect or a negative one. Now, let us examine the way utility typically influences the impartial spectator's pattern of judgment. Smith addresses this issue most directly in response to his friend and contemporary David Hume, who thinks that we determine an action's moral worth based on whether it was beneficial or detrimental to society. Here, Smith demurs, asserting that "it is not the view of this utility or hurtfulness which is either the first or principal source of our approbation and disapprobation. These sentiments are no doubt enhanced and enlivened by the perception of the beauty or deformity which results from this utility or hurtfulness. But still, I say, they are originally and essentially different from this perception" (IV.2.3)<sup>3</sup>. Smith gives two arguments for this belief. First, he claims that we evaluate inanimate objects based on how useful they are to us, and it would be inappropriate to judge people's conduct by the same metric we use for our tools.

Additionally, Smith maintains that we often praise people when they display their genius or exercise impressive self-control even when we are unable to ascertain how they are helping society by doing so. For example, we might be impressed with the strong will it took for someone to break a cigarette addiction, an act which mainly helps themselves and has negligible or nonexistent effects on others. Unlike people today, Smith would count our approbation of the person who quit smoking as an example of our noticing their moral virtue. Such a judgment makes little sense on utilitarian grounds, but can be explained by Smith's belief that the impartial

---

<sup>3</sup> While Smith doesn't mention the impartial spectator by name in that statement, when Smith refers to "our" moral sentiments without specifying that they are the sentiments of a particular individual, we can take him to be describing aspects of an impartial spectator.

spectator deems someone's response to their situation proper when he sympathizes with it, or when he feels he would have responded the same way if he were in their situation.

Smith also bases his theory of justice on the reaction of an impartial spectator rather than on a utility calculus. Utilitarian theorists both in his time and today hold that we should only punish someone if doing so will promote the well-being of society, either by preventing the subject of the punishment from committing crimes in the future or by deterring others from engaging in similar acts. Smith rejects this idea, insisting that "it is not a regard to the preservation of society which originally interests us in the punishment of crimes committed against individuals" (II.ii.3.10). Instead, he grounds punishment on our sentiment of resentment. According to Smith, "proper resentment for injustice attempted, or actually committed, is the only motive which, in the eyes of the impartial spectator, can justify hurting or disturbing in any respect the happiness of our neighbour. To do so from any other motive is itself a violation of the laws of justice" (VI.ii.intro.1). Smith defines resentment as a hostile passion directed at those who have intentionally and without good cause harmed us or those we sympathize with. Resentment drives us to harm the target of this emotion specifically for how they wronged their victim, and in this way is more active than hatred, which merely prompts us to wish ill on its target.

Of course, Smith does not believe that we are justified in punishing anyone we happen to resent. Due to our self-centered nature, we often resent people far more strongly than we ought to and are thus tempted to engage in unduly extreme acts of retribution. For instance, we might resent someone who bullied us as a child, and want to ruin their life as an adult. This might feel right at the time given the strength of your memories about how they treated you, but would be criticized by an impartial spectator witnessing your reaction, who recognizes that they probably

changed since you last knew them and no longer deserve your retaliation. Just punishment flows specifically from proper resentment, which is defined as resentment that an impartial spectator could go along with. Nonetheless, this defense of punishment looks backwards from a utilitarian perspective. The fact that someone merits resentment, even among impartial spectators, does not necessarily mean that punishing them would be useful, and a penal policy might in some situations improve public safety by sanctioning acts which fail to provoke our resentment. Insofar as the impartial spectator focuses on retribution for past misdeeds rather than mitigating future harm, it is not aiming at utilitarian ends.

At the same time, Smith often supports our sentiments by appealing to their utility and even in some places recommends violating his own principles of justice if doing so has positive consequences. For example, Smith explains how a centinel (military guard) who falls asleep on their watch can rightly be executed for the greater good. As R.F. Stalley notes in “Adam Smith and the Theory of Punishment”, Smith’s analysis of military laws seems to stand in tension with his earlier conception of proper penalties. Smith writes:

A centinel, for example, suffers death by the laws of war, because such carelessness might endanger the whole army. This severity may, upon many occasions, appear necessary, and, for that reason just and proper. When the preservation of an individual is inconsistent with the safety of a multitude, nothing can be more just than that the many be preferred to the one. (II.ii.3.11)

Importantly, this action does not match the dictates of an impartial spectator’s moral sentiments. Smith admits that “the thought of this crime does not naturally excite any such resentment, as would prompt us to take such dreadful revenge. A man of humanity must recollect himself, must make an effort, and exert his whole firmness and resolution, before he can bring

himself either to inflict it, or to go along with it when it is inflicted by others” (II.ii.3.11). In other words, when our sense of what a person deserves conflicts with the needs of society, we may be obligated to overcome our internal resistance for the sake of the greater good. One might be tempted to dismiss this case as a single exception with few implications for the rest of Smith’s theory. However, Smith also allows utility to trump propriety in more common situations. In one passage, he discusses the squeamishness some of us display right before carrying out or witnessing a harsh punishment, even one which is justified. Smith explains that at this point, “the generous and humane ... are disposed to pardon and forgive (the criminal), and to save him from that punishment, which in all their cool hours they had considered as the retribution due to such crimes. Here, therefore, they have occasion to call to their assistance the general interest of society” (II.ii.3.7). Unlike the sentinel, those being executed here have merited profound resentment for their conduct. Yet the “generous and humane” people Smith describes have lost that hostility to the criminals they are judging, and would, if they acted on their current sentiments, spare the person on death row. Smith clearly disapproves of that decision, claiming it is motivated by a “weak and partial humanity” as opposed to the “enlarged compassion” which drives justice. Sentiment has force in both directions here, since, despite being impartial spectators, the witness’ sympathy is torn between the lenient and harsh options. Smith grants utilitarian considerations, which call for executing the criminal, the final say.

Utility also resolves the systematic inconsistency in our sentiments Smith describes concerning intentional and accidental harms. In part two, section three of *The Theory of Moral Sentiments*, Smith argues that our theoretical commitment to judge others’ morals based on their motives conflicts with our tendency to admire and condemn people more strongly for their successful endeavors than for their failed ones. He begins the section by declaring that “to the

intention or affection of the heart, therefore, to the propriety or impropriety, to the beneficence or hurtfulness of the design, all praise or blame, all approbation or disapprobation, of any kind, which can justly be bestowed upon any action, must ultimately belong” (II.iii.intro.3). Here, Smith takes himself to be presenting an uncontroversial, defining feature of a “just approbation”, claiming that we all agree that it is always wrong to allow luck to influence our decisions on these matters.

Smith goes on to describe how, in individual circumstances, the unintended consequences of our actions often impact the credit or criticism society gives us. He writes that:

though the intentions of any person should be ever so proper and beneficent, on the one hand, or ever so improper and malevolent, on the other, yet, if they fail in producing their effects, his merit seems imperfect in the one case, and his demerit incomplete in the other. Nor is this irregularity of sentiment felt only by those who are immediately affected by the consequences of any action. It is felt, in some measure, even by the impartial spectator. (II.iii.2.1)

This inconsistency shapes our legal system, causing the attempt to commit a crime to be punished far less severely (if it is punished at all) than actually committing a crime is. Our tendency to react in this fashion also affects popular opinion even in situations where the law does not apply. It explains, for example, why one military leader gets celebrated for victorious missions whereas another one with identical skill but who lost their battles due to misfortune can be hated by the nation.

After noting this irregularity, one might expect Smith to end the section by reproaching our propensity to judge one’s behavior by its consequences and reminding us of the principle stated at the beginning of the chapter. Rather than do so, interestingly, Smith opts to endorse this

aspect of our morality. We can already tell that Smith does not consider it to simply be a human error because he attributes it to the impartial spectator as well. Given that the impartial spectator is purportedly the moral standard throughout *The Theory of Moral Sentiments*, the fact that it acts inconsistently in this manner raises important issues for Smith's philosophy. As Geoffrey Sayre-McCord describes the problem in his article "Sentiments and Spectators" (2010), "the various irregularities all look as if they disqualify those who suffer them as setting the standard for our moral judgments" (p. 15). Sayre-McCord's article explores the extent to which these irregularities threaten the impartial spectator's value as an arbiter of morality<sup>4</sup>. He concludes that the irregularities' usefulness redeems them in the spectator's own eyes, and since the spectator is the only moral standard we have, it passes the highest test we can put it to. My account, by contrast, provides Smith with a source of moral authority which is superior to the impartial spectator and can render judgment upon it. I will argue that Smith considers God to be this type of moral authority. God approves of the impartial spectator, irregularities and all, because it serves his purpose of inducing us to make the most utility-maximizing decisions. This top-down defense of the impartial spectator is stronger than the self-referential one Sayre-McCord suggests.

Smith gives several reasons why rewarding and punishing people based on the outcome of their actions, rather than by their intentions, leads to salutary results. First, he argues that it is impossible to know anyone's true intentions. Attempting to discern this information would force us to pry illegitimately into people's minds and cause even innocent people to worry that they will be suspected of wrongdoing (II.iii.3.2). Secondly, Smith believes that valuing success over

---

<sup>4</sup> Although I introduce Sayre-McCord's article here to suggest a potential issue with the impartial spectator's irregularity, I cover the article in greater depth later in this thesis.



perceived effort incentivizes everyone to work harder, knowing that they cannot earn praise by convincing society that they tried to do good (II.iii.3.3). Lastly, the shame attached to causing accidental injury encourages care when engaging with others and reduces negligent harm (II.iii.3.4). In providing this account, Smith is not merely admitting that our deviation from ethical principle comes with some positive side effects; he is attempting to show why it is a crucial part of our morality. He concludes that “That necessary rule of justice, therefore, that men in this life are liable to punishment for their actions only, not for their designs and intentions, is founded upon this salutary and useful irregularity in human sentiments concerning merit or demerits, which at first sight appears so absurd and unaccountable” (II.iii.3.2). Combining this statement with Smith’s earlier insistence that people only earn praise or blame for their intentions indicates that he thinks we, at least in some situations, ought to treat others in ways they do not deserve if doing so will produce desirable consequences. Accepting this interpretation would push *The Theory of Moral Sentiments* much closer to utilitarianism.

### 3 GOD, HUMAN NATURE, AND INDIRECT UTILITARIANISM

So on the one hand, we see that the impartial spectator, whose judgments often appear to be Smith’s ultimate criterion of right and wrong, does not determine the moral value of actions mainly by analyzing their utility. On the other hand, Smith frequently allows utility to either justify or override the impartial spectator. How do we reconcile these aspects of his philosophy, and to what extent should we consider him a utilitarian? I propose that we can best understand Smith’s views on this issue by examining his comments on religion. Smith incorporates teleological and theological language throughout *The Theory of Moral Sentiments*, and suggests

that we construct the impartial spectator out of characteristics given to us by God. The Deity Smith refers to is a utilitarian, and designed us to promote the collective happiness of our species. Our concept of the impartial spectator aids us in this effort, causing us to adopt the moral attitudes which would most effectively further our welfare. We, however, do not always consciously pursue the ends God sets out for us, and Smith does not advise us to. Thus, utility acts largely indirectly in Smith's theory, underwriting actions which we perform for a plethora of complex reasons. I hope to demonstrate that this two-layered reading of utility's place in Smith's work both fits in directly with Smith's idea of God and accommodates his apparently incompatible positions with respect to utilitarianism.

Smith draws on God largely to provide an origin story for our moral sentiments. Although most of his book is devoted to a descriptive account of our psychological tendencies to judge each other, Smith does not think that these tendencies just emerged contingently through a process of evolutionary or cultural development. Instead, he maintains that "Upon whatever we suppose that our moral faculties are founded, whether upon a certain modification of reason, upon an original instinct, called a moral sense, or upon some other principle of our nature, it cannot be doubted that they were given us for the direction of our conduct in this life" (III.5.6) While Smith doesn't explicitly refer to God here, he implies that some higher power designed our mind and cares about our fate. Elsewhere, Smith more clearly specifies God as our creator and adds that he has benevolent aims. As Smith writes:

the happiness of mankind, as well as of all other rational creatures, seems to have been the original purpose intended by the Author of nature, when he brought them into existence. No other end seems worthy of that supreme wisdom and divine benignity which we naturally ascribe to him; and this opinion, which we are led to by the abstract consideration of his infinite perfections, is still more confirmed by the examination of the works of nature, which seem all intended to promote happiness, and to guard against misery. But by acting according to the dictates of our moral faculties, we necessarily pursue the most effectual means for promoting the happiness of mankind, and may

therefore be said, in some sense, to co-operate with the Deity, and to advance as far as in our power the plan of providence (III.3.5.7)

This statement about God is still a very generic one, and does not imply that Smith is committed to any particular religious sect. Although the question of what Smith's personal religious beliefs were remains controversial, his claims here and in the rest of *The Theory of Moral Sentiments* are consistent with any monotheistic belief system. Also, we can begin to see here how Smith's belief that the world is divinely ordered connects his moral psychology to an indirectly utilitarian normative theory. Since God intended our moral faculties, which judge based on many other factors besides utility, to be "the most effectual means for promoting the happiness of mankind", our actions' moral worth can derive from their utility even if we judge them using our sentiments. Smith fleshes out this religious strain in his philosophy by detailing how the more sophisticated elements of our morality helps us more closely mirror God's judgment. As explained earlier, Smith believes our moral sentiments are most reliable when they are filtered through the perspective of an impartial spectator. This spectator expresses characteristically human attitudes when confronted with praiseworthy or blameworthy actions, but does so free of the personal biases and unrelated emotional forces which distort our initial reactions. To Smith, these advantages earn the impartial spectator the titles of "vicegerent of God" (III.5.6) and "demigod within the breast" (III.2.32). Although one can interpret these honorifics as nothing more than hyperbolic compliments for a mundane way to refine our moral perspicacity, they might cohere better with Smith's aforementioned views on religion if we take them to indicate that there is a genuine connection between God and the impartial spectator. If Smith sees human morality in general as a gift from above, the notion that the purest expression of our moral sentiments contains a sacred aspect is a natural next step.

General rules are the other tool Smith suggests for correcting our moral instincts, and he relates these to God as well. According to Smith, “those important rules of morality are the laws and commands of the Deity, who will finally reward the obedient, and punish the transgressors of their duty” (III.5.2). Like the impartial spectator, these rules acquire their holy imprimatur from our sentiments. They are formed inductively, “from the experience we have had of the effects which actions of all different kinds naturally produce upon us” (III.4.9), not proven through reason or derived from first principles. Also like the impartial spectator, their main purpose is to enable us to act on our consciences without interference from temperament or self-interest. When the Deity created us, he instilled in us the capacity to sympathize with others, which we use to measure the propriety of their feelings and actions. The pressure society imposes on us to behave properly typically conduces to our aggregate well-being, as God intended. Yet sometimes, due to a lack of knowledge or unwarranted preconceptions, an individual acquires a reputation they do not deserve. In such cases, “an appeal lies from his sentence to a much higher tribunal, to the tribunal of their own consciences, to that of the supposed impartial and well-informed spectator, to that of the man within the breast, the great judge and arbiter of their conduct” (III.2.32). By imagining ourselves from this point of view and determining the extent to which we sympathize with the action we performed, we render as divine a verdict as is possible for us.

The impartial spectator is imperfect despite being formed from God-given sentiments. Smith is careful to clarify that the impartial spectator is, “like the demigods of the poets, though partly of immortal, yet partly too of mortal extraction” (II.2.33). Even this spectator can be swayed by the opinion of the masses, and, as we have seen, contains irregularities which threaten to make it an inconsistent judge. Our general rules, too, cannot always tell us what we ought to

do. They excel at keeping us from falling below basic standards of decency or committing genuine injustices, but are insufficiently precise to decide the harder ethical dilemmas. In fact, Smith admits that “the general rules of almost all the virtues, the general rules which determine what are the offices of prudence, of charity, of generosity, or gratitude, of friendship, are in many respects loose and inaccurate, admit of many exceptions, and require so many modifications, that it is scarce possible to regulate our conduct entirely by a regard to them” (III.6.8). After exhausting our resources for moral judgment, we may still be uncertain about what we ought to do or whether one of our past actions was right. In these times, our only option is to “appeal to a still higher tribunal, to that of the all-seeing Judge of the world, whose eye can never be deceived, and whose judgments can never be perverted” (III.2.33). Possibly more than anywhere else, Smith here rejects the idea that the impartial spectator is the highest moral authority and reserves total moral knowledge exclusively for the Deity.

According to the interpretation I have been defending, the impartial spectator should be understood as our best proxy for the divine will. The divine will prioritizes utility over other goals, and constructed human nature in order to maximize this end. As limited creatures, we cannot access God’s complete plan for the world and would be better off focusing on more immediate concerns. In our attempt to do so well, we employ sentiments, rules, and imaginative exercises which do not always respond directly to perceived utility but gain their moral legitimacy due to their tendency to increase our well-being.

#### 4 RESPONDING TO OBJECTIONS ABOUT RELIGION

So on the one hand, we see that the impartial spectator, whose judgments often appear to be Smith's ultimate criterion of right and wrong, does not determine the moral value of actions mainly by analyzing their utility. On the other hand, Smith frequently allows utility to either justify or override the impartial spectator. How do we reconcile these aspects of his philosophy, and to what extent should we consider him a utilitarian? I propose that we can best understand Smith's views on this issue by examining his comments on religion. Smith incorporates teleological and theological language throughout *The Theory of Moral Sentiments*, and suggests that we construct the impartial spectator out of characteristics given to us by God. The Deity Smith refers to is a utilitarian, and designed us to promote the collective happiness of our species. Our concept of the impartial spectator aids us in this effort, causing us to adopt the moral attitudes which would most effectively further our welfare. We, however, do not always consciously pursue the ends God sets out for us, and Smith does not advise us to. Thus, utility acts largely indirectly in Smith's theory, underwriting actions which we perform for a plethora of complex reasons. I hope to demonstrate that this two-layered reading of utility's place in Smith's work both fits in directly with Smith's idea of God and accommodates his apparently incompatible positions with respect to utilitarianism.

Smith draws on God largely to provide an origin story for our moral sentiments. Although most of his book is devoted to a descriptive account of our psychological tendencies to judge each other, Smith does not think that these tendencies just emerged contingently through a process of evolutionary or cultural development. Instead, he maintains that "Upon whatever we

suppose that our moral faculties are founded, whether upon a certain modification of reason, upon an original instinct, called a moral sense, or upon some other principle of our nature, it cannot be doubted that they were given us for the direction of our conduct in this life” (III.5.6)

While Smith doesn't explicitly refer to God here, he implies that some higher power designed our mind and cares about our fate. Elsewhere, Smith more clearly specifies God as our creator and adds that he has benevolent aims. As Smith writes:

the happiness of mankind, as well as of all other rational creatures, seems to have been the original purpose intended by the Author of nature, when he brought them into existence. No other end seems worthy of that supreme wisdom and divine benignity which we naturally ascribe to him; and this opinion, which we are led to by the abstract consideration of his infinite perfections, is still more confirmed by the examination of the works of nature, which seem all intended to promote happiness, and to guard against misery. But by acting according to the dictates of our moral faculties, we necessarily pursue the most effectual means for promoting the happiness of mankind, and may therefore be said, in some sense, to co-operate with the Deity, and to advance as far as in our power the plan of providence (III.3.5.7)

This statement about God is still a very generic one, and does not imply that Smith is committed to any particular religious sect. Although the question of what Smith's personal religious beliefs were remains controversial, his claims here and in the rest of *The Theory of Moral Sentiments* are consistent with any monotheistic belief system. Also, we can begin to see here how Smith's belief that the world is divinely ordered connects his moral psychology to an indirectly utilitarian normative theory. Since God intended our moral faculties, which judge based on many other factors besides utility, to be “the most effectual means for promoting the happiness of mankind”, our actions' moral worth can derive from their utility even if we judge them using our sentiments. Smith fleshes out this religious strain in his philosophy by detailing how the more sophisticated elements of our morality helps us more closely mirror God's judgment. As explained earlier, Smith believes our moral sentiments are most reliable when they are filtered through the perspective of an impartial spectator. This spectator expresses characteristically

human attitudes when confronted with praiseworthy or blameworthy actions, but does so free of the personal biases and unrelated emotional forces which distort our initial reactions. To Smith, these advantages earn the impartial spectator the titles of “vicegerent of God” (III.5.6) and “demigod within the breast” (III.2.32). Although one can interpret these honorifics as nothing more than hyperbolic compliments for a mundane way to refine our moral perspicacity, they might cohere better with Smith’s aforementioned views on religion if we take them to indicate that there is a genuine connection between God and the impartial spectator. If Smith sees human morality in general as a gift from above, the notion that the purest expression of our moral sentiments contains a sacred aspect is a natural next step.

General rules are the other tool Smith suggests for correcting our moral instincts, and he relates these to God as well. According to Smith, “those important rules of morality are the laws and commands of the Deity, who will finally reward the obedient, and punish the transgressors of their duty” (III.5.2). Like the impartial spectator, these rules acquire their holy imprimatur from our sentiments. They are formed inductively, “from the experience we have had of the effects which actions of all different kinds naturally produce upon us” (III.4.9), not proven through reason or derived from first principles. Also like the impartial spectator, their main purpose is to enable us to act on our consciences without interference from temperament or self-interest. When the Deity created us, he instilled in us the capacity to sympathize with others, which we use to measure the propriety of their feelings and actions. The pressure society imposes on us to behave properly typically conduces to our aggregate well-being, as God intended. Yet sometimes, due to a lack of knowledge or unwarranted preconceptions, an individual acquires a reputation they do not deserve. In such cases, “an appeal lies from his sentence to a much higher tribunal, to the tribunal of their own consciences, to that of the



supposed impartial and well-informed spectator, to that of the man within the breast, the great judge and arbiter of their conduct” (III.2.32). By imagining ourselves from this point of view and determining the extent to which we sympathize with the action we performed, we render as divine a verdict as is possible for us.

The impartial spectator is imperfect despite being formed from God-given sentiments. Smith is careful to clarify that the impartial spectator is, “like the demigods of the poets, though partly of immortal, yet partly too of mortal extraction” (II.2.33). Even this spectator can be swayed by the opinion of the masses, and, as we have seen, contains irregularities which threaten to make it an inconsistent judge. Our general rules, too, cannot always tell us what we ought to do. They excel at keeping us from falling below basic standards of decency or committing genuine injustices, but are insufficiently precise to decide the harder ethical dilemmas. In fact, Smith admits that “the general rules of almost all the virtues, the general rules which determine what are the offices of prudence, of charity, of generosity, or gratitude, of friendship, are in many respects loose and inaccurate, admit of many exceptions, and require so many modifications, that it is scarce possible to regulate our conduct entirely by a regard to them” (III.6.8). After exhausting our resources for moral judgment, we may still be uncertain about what we ought to do or whether one of our past actions was right. In these times, our only option is to “appeal to a still higher tribunal, to that of the all-seeing Judge of the world, whose eye can never be deceived, and whose judgments can never be perverted” (III.2.33). Possibly more than anywhere else, Smith here rejects the idea that the impartial spectator is the highest moral authority and reserves total moral knowledge exclusively for the Deity.

According to the interpretation I have been defending, the impartial spectator should be understood as our best proxy for the divine will. The divine will prioritizes utility over other

goals, and constructed human nature in order to maximize this end. As limited creatures, we cannot access God's complete plan for the world and would be better off focusing on more immediate concerns. In our attempt to do so well, we employ sentiments, rules, and imaginative exercises which do not always respond directly to perceived utility but gain their moral legitimacy due to their tendency to increase our well-being.

## **5 RESPONDING TO OBJECTIONS ABOUT UTILITY**

I have hoped to establish so far that Smith believes God designed human morality so that people would maximize their collective happiness without usually having that explicit goal in mind but instead being motivated by moral emotions. I just addressed those who doubt this thesis because they argue it overstates Smith's religiosity. Now, I need to turn to the scholars who instead reject it because they claim Smith was less of a utilitarian than I make him out to be. I am dealing with the objections in this order because I believe the inclusion of God in Smith's theory is crucial for answering some of the criticisms made against utilitarian readings of his work. Geoffrey Sayre-McCord's article "Sentiments and Spectators: Adam Smith's Moral Philosophy" (2010) is particularly useful to examine here because he explicitly addresses the indirect utilitarian theory I put forth in this thesis. In reference to this interpretation, Sayre-McCord suggests that "Smith may be supposing that people perfectly properly regulate their moral judgments by appeal to the standard set by the impartial spectator, having no further end in view, even though it is the resulting contribution to public welfare and to the happiness of mankind that justifies our judging this way" (p. 17). His main problem with this idea is that the moral worth of utility-promoting behavior is still influenced by the context in which it occurs and not merely by

the amount of happiness it produces or by its long-term effects. Sayre-McCord makes this point by writing that “Smith rejects the idea that happiness, no matter whose, no matter how secured, is valuable. That someone might take pleasure in some activity is no defense of that activity, if that pleasure is improper, nor does the fact that some activity might promote happiness work to justify that action, if the happiness depended on doing something wrong” (p. 17). If non-consequentialist considerations can override aggregate happiness in Smith’s determination of the ultimate value of someone’s conduct, it becomes hard to define him as a utilitarian of any sort.

In a two-level moral theory like the one I ascribe to Smith, the question of the worth of any action hinges on whose perspective it is being assessed from. We judge someone’s actions as morally right or wrong based on our sense of their propriety, which is the extent to which an impartial spectator would find it a fitting response to the person’s situation. Since the impartial spectator sometimes considers actions which decrease aggregate happiness appropriate or ones which increase aggregate happiness inappropriate, the standards of propriety and utility can reach different results. My contention is that God always prioritizes maximizing utility over other goals, which means that relying on the moral faculties he gave us will tend to promote utility. Smith states this most clearly when he insists that “All the inhabitants of the universe, the meanest as well as the greatest, are under the immediate care and protection of that great, benevolent, and all-wise Being, who directs all the movements of nature; and who is determined, by his own unalterable perfections, to maintain in it, at all times, the greatest possible quantity of happiness” (VI.ii.3.5). This characterization of the divine will rules out the possibility that the Deity would sacrifice someone’s happiness merely because it was acquired inappropriately if doing so would not increase others’ happiness by a larger margin. Since everyone’s happiness, regardless of how it was obtained, increases the total amount of it in the world, God must value it

all. Although we place our standards of propriety alongside, and sometimes even ahead of, determinations of utility to figure out what we approve of, God would contradict himself if he ever allowed any other factor to override the utility calculus when he was designing the world.

To better understand the relationship between propriety and utility in Adam Smith's theory, it is worth analyzing his account of the virtues. Smith cites four principal virtues - justice, benevolence, prudence, and self-command. These are the main character traits which we approve of in others, try to inculcate in our children, and seek to acquire ourselves. An important test of my reading of Smith is to see how these virtues relate to the pursuit of happiness. Does Smith think that behaving virtuously necessarily redounds to the benefit of society as a whole? If and when it doesn't, does Smith have any advice on how to weigh the consequences of an act against its virtuousness or lack thereof? Smith addresses the relationship between utility and virtue in his comments on the history of philosophy in a way that makes it seem as though his virtue theory is a problem for his utilitarianism. He writes that "the only difference between (a system based on utility) and that which I have been endeavoring to establish, is, that it makes utility, and not sympathy, or the correspondent affection of the spectator, the natural and original measure" or the virtuous degree of a sentiment (VII.ii.4.1). If we read him here as saying that, for us, the impartial spectator's sympathy is the best way to tell what amount of each sentiment we should display, I think this passage becomes consistent with my interpretation. As long as the virtuous degree of each sentiment is the one which most effectively conduces to overall well-being, we can read Smith as an indirect utilitarian. We will see that all the virtues tend to promote happiness in general, but that there are certain times when behaving virtuously can have undesirable consequences. Such cases appear somewhat troubling for the interpretation I have

been promoting, but can be incorporated into my theory with the proper conception of the function of our sentiments in a two-level moral system.

Of the four virtues, self-command can most easily conflict with the principle of utility. Smith closely associates the other three with the happiness they bring about, claiming that “the virtues of prudence, justice, and beneficence, have no tendency to produce any but the most agreeable effects. Regard to those effects, as it originally recommends them to the actor, so does it afterwards to the impartial spectator” (VI.conc.6). Smith admittedly implies elsewhere, for example in his discussion of resentment and punishment, that non-consequentialist sentiments sometimes help explain our judgments regarding these virtues as well, but overall the impartial spectator behaves in line with utilitarian maxims when praising people for being prudent, benevolent, or just. By contrast, Smith writes that “in our approbation of the virtues of self-command, complacency with their effects sometimes constitutes no part, and frequently but a small part, of that approbation. Those effects may sometimes be agreeable, and sometimes disagreeable; and though our approbation is no doubt stronger in the former case, it is by no means altogether destroyed in the latter” (VI.conc.7). Although the first part of this statement, that we approve of self-command for non-utilitarian reasons, fits into an indirect utilitarian picture without much difficulty, his second claim that we actually approve of displays of self-command even when they are detrimental to society’s overall well-being represents a greater departure from utilitarianism.

How exactly does self-command lead to bad consequences? The main cases where this occurs, and the ones Smith brings up, arise when people use their ability to subdue their fear or anger to more effectively carry out evil plans. Smith describes the negative side of self-command by arguing that “the most intrepid valour may be employed in the cause of the greatest injustice.

Amidst great provocations, apparent tranquility and good humour may sometimes conceal the most determined and cruel resolution to revenge. The strength of mind requisite for such dissimulation, though always and necessarily contaminated by the baseness of falsehood, has, however, been often much admired by many people of no contemptible judgment” (VI.iii.12). Even here, utility plays a significant role in our judgment process, as it renders our approval more qualified than it would otherwise be. Still, Smith thinks that we naturally view their courage as a redeeming feature, and that we should not be faulted for doing so. Even when we decide that a brave or calm-tempered person is vicious on the whole, their self-command remains a virtue - just one that has been outweighed by their lack of benevolence or justness. A typical utilitarian, on the other hand, would regard a dangerous person’s self-command as an exacerbating characteristic. They would judge the self-controlled reprobate more harshly than the weak-willed one, since the former can do more harm than the latter. Smith’s contrary tendency indicates that utility need not always be the most important good for him.

We can understand why Smith finds it acceptable for us to approve of harmful behavior by clarifying exactly what aspect of our moral psychology must be sensitive to utility-based criteria. My claim is not that we always rank actions according to the nature of their results, but just that our moral sentiments themselves consistently promote the greater good. The relevant question, then, is whether it is beneficial to praise villainous characters for possessing a high degree of self-command. Most important to recognize on this issue is Smith’s beliefs about how crucial self-command is in general to the functioning of society. He argues that our ability to control our actions based on a sense of propriety is so important that “without the restraint which this principle imposes, every passion would, upon most occasions, rush headlong, if I may say so, to its own gratification. Anger would follow the suggestions of its own fury; fear those of its

own violent agitations” (VI.concl.2). Since self-command is such a valuable trait overall, increasing its prevalence benefits society on net.

This benefit justifies our approbation of wicked but self-controlled individuals. Our admiration for them is founded on an appreciation for their valuable skills, which in turn inspires a desire to improve ourselves in that respect. As an example of this attitude, Smith describes how “when a robber or highwayman is brought to the scaffold, and behaves there with decency and firmness, though we perfectly approve of his punishment, we often cannot help regretting that a man who possessed such great and noble powers should have been capable of such mean enormities” (VI.iii.6). This reaction is precisely the one a utilitarian God would want us to have. On the one hand, our respect for them does not prevent us from taking the necessary measures to deter future crimes and protect the public, but on the other, our hope that better people will share their most useful attributes means that some positive side effects can emerge from witnessing their example.

My response to Sayre-McCord reveals how the two-level, God-based utilitarianism I argue Smith endorses describes his theory more effectively than a straightforward utilitarian interpretation would. Sayre-McCord understandably points out that Smith portrays our moral judgment, even at its best, as multifaceted and sensitive to many factors other than utility. This feature of his theory would completely contradict the idea that Smith is a direct utilitarian, since if he were, he would have to regard non-utilitarian moral evaluations as errors. However, my interpretation only requires Smith to think that each of our moral sentiments have an overall positive utilitarian impact, not that each of our actions does. To demonstrate that Smith’s theory meets this criteria, I discussed his four cardinal virtues, focusing on self-command. By explaining how even Smith’s account of self-command, the least utilitarian virtue, is compatible

with our creation by a utilitarian God, I showed why Sayre-McCord fails to rule out all utilitarian interpretations of *The Theory of Moral Sentiments*.

## 6 CONCLUSION

The possibility for particular actions or character traits to count as virtuous despite failing to promote collective happiness raises a broader question about what Smith takes to be the criterion of morality. I've maintained throughout this paper that utility is the sole standard for God's judgment, whereas the impartial spectator, who values utility together with other factors, is the standard for ours. But which of these standards actually determine what qualifies as a moral action? One strategy would be to say that God, as an omniscient being, is always right about morality, and so utility is the correct standard to use when evaluating actions. On that view, the impartial spectator would simply be wrong when it issues non-utilitarian judgments, even if we can recognize on reflection that it is good for it to judge this way. Relying on the impartial spectator could still be our most reliable means for moral improvement, but it would nonetheless be a fallible tool and not the true moral standard. The best textual evidence against this interpretation comes when Smith compares morality to our other senses, and writes that "what is agreeable to our moral faculties, is fit, and right, and proper to be done; the contrary wrong, unfit, and improper. The sentiments which they approve of, are grateful and becoming: the contrary, ungraceful and unbecoming. The very words, right, wrong, fit, improper, grateful, unbecoming, mean only what pleases or displeases those faculties" (TMS III.5.5). Here, he is explicitly defining morality in terms of our sentiments, and presumably primarily in terms of those of the impartial spectator, since it most aptly expresses those sentiments. Under this



definition, it is not possible to call an action we approve of wrong just because it doesn't meet utilitarian requirements for morality, even if God abided by those requirements in his own plan of creation.

However, this approach might generate the concern that by defining morality so that it can conflict with God's own motives, we risk the implication that God is morally wrong when he acts in ways which offend our sentiments. The idea that we could legitimately condemn the Deity, and that we could do so using faculties he gave us, is unappealing for obvious reasons. Additionally, admitting that the moral criterion differs from the principle of utility might undermine my attempts to view Smith as a utilitarian, even an indirect one. I believe the best way to look at this issue is to understand the impartial spectator's reactions as determining which actions are moral, but utility as ultimately explaining why we make the moral judgments we do and why morality truly matters. The impartial spectator only finds some actions proper and others improper because God thought it would benefit us if it did so. So, while generating the sympathy of the impartial spectator is in an immediate sense what it means for an action to be right, utility makes behavior right or wrong in the deeper sense that it is the reason we sympathize with it or fail to do so in the first place.

Lastly, we can see how Smith connects the human standard of propriety to the divine standard of utility through his description of our opinion of God's moral status. While we cannot morally judge God himself, we can decide whether or not we think following him is proper.

Smith argues that God passes our own test, declaring:

That our regard to the will of the Deity ought to be the supreme rule of our conduct, can be doubted of by nobody who believes his existence. The very thought of disobedience appears to involve in it the most shocking impropriety. How vain, how absurd would it be

for man, either to oppose or to neglect the commands that were laid upon him by Infinite Wisdom, and Infinite Power! How unnatural, how impiously ungrateful not to reverence the precepts that were prescribed to him by the infinite goodness of his Creator, even though no punishment was to follow their violation (III.5.12).

So, when we come to learn that God designed our moral faculties to promote utility, our devotion to God leads us to view that as the chief end of our action. We continue to center our moral lives around judgments of propriety, but we begin to do so mainly due to our faith that that is what God would want from us and that he has our best interests in mind.

## REFERENCES

- Smith, Adam, D.D Raphael, and A.L Macfie. *The Theory of Moral Sentiments*. Oxford: Clarendon Press, 1976. Print.
- Louden, Robert B. "Butler's Divine Utilitarianism." *History of Philosophy Quarterly*, vol. 12, no. 3, 1995, pp. 265–280. [www.jstor.org/stable/27744666](http://www.jstor.org/stable/27744666). Accessed 1 Mar. 2021.
- Stalley, Richard F. "Adam Smith and the Theory of Punishment". *Journal of Scottish Philosophy*, vol. 10, issue 1, 2012, p. 69-89.
- Sayre-McCord, Geoffrey. "Sentiments and Spectators: Adam Smith's Theory of Moral Judgment". *Essays on the Philosophy of Adam Smith*, edited by Vivienne Brown and Samuel Fleischacker, Routledge, 2010.
- Kopatijic, Lauren. "The Vicegerent of God? Adam Smith on the Authority of the Impartial Spectator". *Journal of Scottish Philosophy*, vol. 17 issue 1, p. 61-78.
- Heydt, Colin. "The Problem of Natural Religion in Smith's Moral Thought. *Journal of the History of Ideas*, vol 78 issue 1, p. 73-94, 2017
- Price, Anthony, "Richard Mervyn Hare", *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition), Edward N. Zalta (ed.) <https://plato.stanford.edu/entries/hare/>