

University of Missouri, St. Louis
IRL @ UMSL

Theses

Graduate Works

4-21-2010

Irrelevance and Evaluation: Two Arguments Against Gauthier's Contractarianism

Micah William Svejda
University of Missouri-St. Louis

Follow this and additional works at: <http://irl.umsl.edu/thesis>

Recommended Citation

Svejda, Micah William, "Irrelevance and Evaluation: Two Arguments Against Gauthier's Contractarianism" (2010). *Theses*. 226.
<http://irl.umsl.edu/thesis/226>

This Thesis is brought to you for free and open access by the Graduate Works at IRL @ UMSL. It has been accepted for inclusion in Theses by an authorized administrator of IRL @ UMSL. For more information, please contact marvinh@umsl.edu.

Irrelevance and Evaluation: Two Arguments Against Gauthier's
Contractarianism

Micah W. Svejda
B.A. History, Northwestern College, 2002

A Thesis Submitted to The Graduate School at the University of
Missouri-St. Louis in partial fulfillment of the requirements for the
degree of Master of Arts in Philosophy

May 2010

Advisory Committee

John Brunero, Ph.D.
Chairperson

Eric Wiland, Ph.D.

Irem Kurtsal-Steen, Ph.D.

Copyright, Micah W. Svejda, 2010

Introduction

The purpose of this paper is to provide two arguments against David Gauthier's version of contractarianism. The purpose of these two arguments is to weaken the cumulative case for this view. Contractarianism, in brief, attempts to offer us a theory of morals by agreement. This theory provides answers to the following questions. Which moral claims are we justified in making? Why are these claims justified?

Drawing on the Hobbesian tradition, persons are seen as primarily disposed to pursue their own interests. Thus, any moral claims that inhibit these interests must be justified. Gauthier's test for justified moral claims comes in the form of an agreement. Justified moral claims are those that self-interested persons *would* accept. Thus, the agreement is hypothetical and not actual, or historical. But the parties to the agreement are actual persons.

Not all of our existing moral claims pass this test; a good many of our moral claims fail it and are, consequently, unjustified. But despite this fact, we still have a good reason to comply with *justified* moral claims: it makes us better off. By agreeing to constrain some of our interests, rather than not agreeing to constrain any of them, we are able to gain the benefits of cooperation with others. These benefits are available only to those who so cooperate. So, the two questions are answered. Which moral claims are we justified in making? Those claims that self-interested persons would agree upon. Why are these claims justified? Because they make available benefits otherwise unavailable to us. We are, then, provided with a theory of

morality: morals by agreement. All of this will be fleshed out in a fair amount of detail in the first section.

The first argument of the present paper is as follows. According to Gauthier, if, upon making this hypothetical agreement, we come to realize that the terms of the agreement itself rely on a conception of morality that is not the object of agreement, we will come to view the agreement as tainted, or *unstable*.¹ To re-establish the agreement on more stable grounds, each party to it will, he supposes, consider which moral claims each would accept in a pre-moral, pre-social situation. The agreement, since it then proceeds from a situation that escapes prior influence of any moral conception, will be a purer, or more stable, test for morality.

I will argue that were we to engage in this task of imagining ourselves in such a situation, the resulting imagined person would be quite foreign to ourselves. This hypothetical person, then, is the party to the agreement. The constraints that she would choose in such a situation then appear irrelevant to us. Since these constraints form the object of the agreement, then the hypothetical agreement itself is irrelevant to us. Thus, contractarianism lacks the resources to provide the agreement with the stability that Gauthier insists is needed in order to justify morality.

The second argument responds to a thesis of Gauthier's that morality faces a foundational crisis and that contractarianism is the only plausible

¹ This term is Gauthier's, and it will be explained in due course.

resolution of it. I will argue that Gauthier's resolution to it, deliberative justification, implies an implausible view of moral evaluation.

Moral rationalism, or the view that morality provides good reasons for action, is presented and then bolstered as an alternative view to contractarianism. Once the plausibility of this view is in place, it becomes clear that Gauthier's resolution is implausible.

With these two arguments presented, the cumulative case for contractarianism will be weakened. Section 1 will outline Gauthier's view. I will then provide the two arguments explained above, in Sections 2 and 3 respectively. I will conclude by offering some comments about answering the moral skeptic; it may not be the role of the moral philosopher to do this.

Section I: Gauthier's Contractarianism

Social contract theory, in a distinctly modern form, can be traced back to Thomas Hobbes; it is the idea that morality (or political authority) is grounded, or legitimated, or justified, by an agreement of some sort. John Locke is, of course, another member of this tradition, as is Jean-Jacques Rousseau and Immanuel Kant.²

In the twentieth-century, social contract theorists can be viewed as belonging to roughly one of two camps.³ Contractualists, descending from

² For a detailed treatment of this tradition, see J. W. Gough, The Social Contract: A Critical Study of Its Development. Oxford: Clarendon Press, 1936. See also Jean Hampton, Hobbes and the Social Contract Tradition. Cambridge: Cambridge University Press, 1986.

³ For more on this distinction, see Jean Hampton, "Two Faces of Contractarian Thought" in Contractarianism and Rational Choice. Ed Peter Vallentyne. Cambridge: Cambridge University Press, 1991.

Kant's moral philosophy, hold that it follows from the concept of a rational being that there exist certain constraints on the actions of each. Rational persons are those that seek the interests of everyone. It follows then, that some constraints are necessary for the attainment of the interests of all. For, if there were no constraints, then these interests would not be served.

Contractualists explain these constraints by way of an agreement. John Rawls⁴ and Thomas Scanlon⁵ are two more contemporary proponents of this camp.

Contractarians, by contrast, descend from the Hobbesian tradition. One general assumption of this camp is that humans are primarily self-interested and that morality can be shown to be the most rational way to serve our interests. Gauthier and Jan Narveson⁶ are two influential contemporary proponents of this camp. Each camp holds that morality proceeds from rationality by way of a contract between rational parties. The remainder of this section will be devoted to outlining David Gauthier's contractarianism.

Gauthier provides us with a theory of morals by agreement.⁷ It attempts to derive moral principles from rational choice theory. This latter theory seeks to provide the general principles of choice underlying how

⁴Rawls' view is most fully developed in John Rawls. A Theory of Justice. Cambridge, MA: Belknap Press of Harvard University Press, 1971.

⁵ A good exposition of Scanlon's view is Thomas Scanlon. What We Owe To Each Other. Cambridge, MA: Harvard University Press. 1998.

⁶ Jan Narveson. The Libertarian Idea. Philadelphia: Temple University Press, 1988.

⁷ The summary here provided is drawn primarily from David Gauthier, Morals By Agreement. Oxford: Oxford University Press, 1986.

humans, if they are to be rational, should act. Part of rational choice theory is the claim that persons are disposed to seek their own interest. Since cooperation with others is more likely to bring greater advantage to an individual, it is recognized as a rational choice. However, cooperation with others calls for constraints on each individual's choices, in order to avoid exploitation. For, if each were to act only to serve her self-interest and there were no constraints, then it is likely that there would be many cases of exploitation, since exploitation would be seen as an efficient means to gain greater advantage. Exploitation would be avoided if each were to comply with some constraints. Gauthier identifies these abovementioned constraints as moral principles. Moral principles, then, enable each individual to seek her own interest through interaction with others in a better way than she could without others. I will now flesh out this sparse summary.

Gauthier begins developing his theory by explaining the conception of rationality that underlies rational choice theory. This conception of rationality is instrumental; it is a kind of means-ends rationality. Following Hume, a person's desires are asserted to be neither rational nor irrational. When a person takes means to procure the ends she desires, she is said to be instrumentally rational.⁸

This conception of rationality he calls the maximizing conception of rationality. Since rational actions are those that procure for us the objects of our interest, we are rational if we pursue those objects. Gauthier identifies

⁸ Ibid. p. 21.

interest with preference, so that a person acts rationally to the extent that she seeks to fulfill her preferences. It is important to note here that one can have a preference involving the well-being of others. I may prefer to see that you are happy; if something I do makes you happy, and I am aware of this, then my preference is fulfilled. Lacking this preference, however, I have no reason to do anything to see that you are happy. On the maximizing conception, then, there is no reason for us to be moral if we do not prefer it.

To illustrate these features of the maximizing conception, he draws a contrast with the universalistic conception of rationality, which derives from Kant. On this conception, a person is acting rationally only if she seeks the interests of everyone. Therefore, says Gauthier, the universalistic conception already has the moral dimension of impartial constraint built into it.⁹ This conception of rationality contains an explanation of why a person should be moral: to constrain your actions is a part of being rational. A person is acting rationally if she seeks to satisfy the interests of all rational beings. Hence, constraints on an individual's actions are necessary in order to ensure that everyone's interests are satisfied. The relation between reason and morality is clear: one must be moral in order to be rational.

On the maximizing conception, however, the relation between reason and morals is unclear. For, if a rational choice is one that maximizes our preference-fulfillment, then what need have we for any notion of constraint? There is no explanation for its necessity. The main task of Gauthier's theory

⁹ Gauthier describes the traditional understanding of morality as an "impartial constraint on the pursuit of individual interest." Ibid. p. 7.

is, thus, to explain why morality is required for a rational person. We thus are provided with a theory that clarifies the nature of the relation between reason and morals. It explains how, even though we are primarily disposed to seek our self-interested preferences, morality is needed in order to better promote the fulfillment of those preferences. Morality is a means to their fulfillment; it is (instrumentally) rational to be moral.

For Gauthier, a person is conceived as “an independent centre of activity, endeavouring to direct his capacities and resources to the fulfillment of his interests.”¹⁰ There is no room for constraint; each person initially approaches a choice situation with no limits on what she may or may not do. How, then, does a person come to recognize any constraints on her actions? Where does morality come from if it is not already present to us? Understanding morality as arising out of an agreement offers an answer to this question.

This agreement is not an actual, historical event. Rather, it is an event in a thought-experiment intended to establish why it is better that each of us should choose to live with some constraints, rather than with none. The agreement is hypothetical, but the parties to it are actual people. If actual, rational persons were to engage in a bargain, wherein they came to an agreement about which, if any, constraints would be acceptable to each, then a distinction between what may and may not be done is provided. The agreement is *hypothetical* in that it picks out what each rational person *would*

¹⁰ Ibid. p. 9.

agree to; but the bargainers are *actual*, rational people. If each person were to consider which constraints she would prefer, then an agreement is reached; it is that set of constraints that everyone would choose.

But if everyone were free from constraint, why would anyone engage in such a bargain? Each person will come to see the benefits available to her through cooperation with others. If each person comes to realize that, by pursuing these goods, she will maximize her own preference-fulfillment, then each will voluntarily engage in the bargain. As this is recognized, the place for constraint becomes quite clear to those parties to the agreement. It is to increase the future likelihood of the cooperation of others.

It becomes more complicated, says Gauthier, when we consider that although the agreement is *hypothetical*, it demonstrates the rationality of *actual* constraint. What has the content of some thought-experiment to do with constraints on *my* actions? This challenge must be answered by any theory of morals, says Gauthier.

The theory he offers answers this challenge, and in a more rigorous (and thus superior) way than many past contractarian theories. Gauthier traces the history of moral contractarianism back to the Greek Sophists. Thomas Hobbes is identified as the true progenitor of Gauthier's form of the theory. More recently, he counts G. R. Grice¹¹ and Kurt Baier¹² as co-workers in this enterprise. Gauthier, however, bolsters this conceptual history by the

¹¹ G. R. Grice. The Grounds of Moral Judgement. Cambridge: Cambridge University Press, 1967.

¹² Kurt Baier. The Moral Point of View: A Rational Basis of Ethics. Ithaca, NY: Cornell University Press, 1958.

addition of rational choice theory. Doing so is able to answer a number of questions: “why rational persons would agree *ex ante* to constraining principles [i.e. why we would be motivated to begin bargaining at all], what general characteristics these principles must have as objects of rational agreement [explained below], and why rational persons would comply *ex post* [after the agreement] with the agreed constraints.”¹³ By invoking rational choice theory, then, we are able to see why *we* should agree to comply with constraints: simply because it is rational. This explanation, however, is more rigorous than past attempts, since it corroborates with a separate theory of how *all* rational persons should make choices.

Why is the hypothetical agreement itself rational? The rationality of the agreement becomes clear when the role of society in an individual’s life is considered. Society is stipulated as a cooperative venture, existing for the benefit of each member of it. Contractarians assume that some degree of sociability is a characteristic feature of humanity. Hence, we tend to live in societies. However, only when social interaction is beneficial to each individual is it considered rational. Since we are social creatures, tending to live in societies, it appears that an agreement is rational from the perspective of each; if I am already disposed to live among others, I might as well agree to cooperate with them.

So, Gauthier has shown us why we would be motivated to make an agreement, but why should we comply with the constraints agreed upon

¹³ Gauthier, p. 10.

once the agreement is finished? Contractarianism begins from this presumption against morality.¹⁴ Why agree to comply with the constraints of a hypothetical agreement at all? Gauthier takes pains to make us feel the force of this challenge, for he says, if his theory is successful, it defeats this initial presumption against morality.

Gauthier identifies four features central to his theory. The first is that of the “morally free zone” of the idealized market.¹⁵ A market, Gauthier argues, is free from constraint. In a perfectly competitive market, mutual advantage is assured if each individual pursues her own interests without constraint. Hence, there is no *rational* need for constraint. Moreover, since each person enjoys the same advantage in market interaction as she would non-socially, there is no partiality within market interaction. Hence, there is no *moral* need for constraint.¹⁶ Thus, the market is morally free. It serves as a foil against which morality becomes clearer; if the world were such a market, morality would be unnecessary. Of course, the world is not such a market. Unethical business practices are engaged in all the time. Hence, some constraint is needed. The idealized market exhibits a *natural* harmony among participating parties; morality makes possible an *artificial* harmony. “Market and morals share the non-coercive reconciliation of individual interest [or, preference-fulfillment] with mutual benefit.”¹⁷

¹⁴ Ibid. p. 8.

¹⁵ Ibid. p. 13.

¹⁶ The assumption here is that morality is a set of *impartial constraints* on individual preference-fulfillment.

¹⁷ Ibid. p. 14.

The second feature central to his theory is as follows. When individual constraint becomes necessary, this is achieved through rational agreement. A necessary condition on this agreement is that it be mutually advantageous, but what is sufficient for such an agreement? This gives rise to the bargaining problem for the bargainers; it consists of deciding which, among a number of possible agreements, is the most advantageous to each bargainer.

Solving this problem leads to two principles that guide both the process of and the content of rational agreement. Beginning from an assumption of the equal rationality of all parties involved, the first principle is arrived at: each party will concede as little as possible relative to what he contributes. Likewise, a second principle is identified: each party will benefit as much as possible relative to what he contributes. Since these principles capture both fairness and impartiality, they provide a basis for justice.

The third feature central to his theory is that of *constrained maximization*. Gauthier considers it a weakness of much past contractarian theory that it has not been able to deal with the rationality of compliance. Granted, it may be rational to make agreements based on the two principles I outlined in the previous paragraph, but is it not equally rational to ignore these if it serves an individual's interests? To answer this question, Gauthier distinguishes between *straightforward maximizers* and *constrained maximizers*. The former is disposed to maximize her own interest in the

particular choices she makes. However, a *constrained maximizer* will comply with mutually advantageous constraints, provided she expects others to do so as well. Of course, this leaves the *constrained maximizer* open to being exploited. However, Gauthier endeavors to show that it is rational to be so disposed, since under most conditions the net benefits of cooperation exceed the risk of being exploited. Thus, it is rational to *internalize*, or to intentionally cultivate as a disposition, moral principles used in decision-making; it is likewise irrational to *straightforwardly* appeal to preference-fulfillment, since fewer net benefits are made available by straightforward maximizing.

The fourth feature of his theory is the proviso on the initial bargaining position (or the list of things that an individual brings to the bargaining table). This was adapted from the proviso of Locke in his version of social contract theory.¹⁸ This proviso prohibits bettering one's own position by worsening that of others during the process of bargaining. At the bargaining table, each person's prospects post-agreement cannot be worse than they would have been in the absence of an agreement. Otherwise, no one has a reason to interact at all. For example, if my options on the bargaining table were either slavery or abject poverty, I would have no reason to make such a choice at all. Clearly, I would be better off not cooperating at all with people who gave me only those two choices! So, since society is viewed as a cooperative venture, with each willingly complying with the agreed-upon

¹⁸ From John Locke, Two Treatises of Government. Second treatise, ch. v. paragraph 27. London, 1690.

constraints, then they must find acceptable the initial bargaining position. The proviso is thus posited as necessary in order for rational persons to bargain at all—it ensures that agreement is made willingly and not coercively.

Gauthier holds that a contractarian theory of morals developed from rational choice theory has certain strengths and certain weaknesses. By demonstrating the rationality of compliance with agreed constraints, morality is “given a sure grounding in a weak and widely accepted conception of practical rationality.”¹⁹ His is the only account that demonstrates that morality can be shown to be rational from non-moral suppositions—without reverting to a *universalistic* conception of rationality.

However, there are weaknesses to it, he admits. The agreement excludes certain people. Only persons who are roughly equal can be parties to an agreement. Not all people are equal. Exploitation of those not party to the agreement, by the parties to the agreement, is therefore possible. And Gauthier admits that his theory has no place to condemn this. The Lockean proviso against bettering one’s own position by worsening that of others only governs the *process* of the hypothetical agreement, which assumes that the parties to it are equally rational. Hence, morality only arises from the agreement of *rational* parties. Since not all people are rational, not all are parties to the agreement. Hence, Gauthier admits, there is no place for condemning all cases of exploitation on his theory.

¹⁹ Gauthier, p. 17.

Western society can be represented well as a cooperative venture. It has discovered, he says, how individuals (each in the pursuit of her own interests) work for mutual benefit. If we consider the vast quantity of material goods, the ever-increasing lifespan of individuals, and the overall level of well-being of individuals (to name just some of the benefits), we can see what has resulted from this discovery.²⁰ However, perhaps too much credit has been given to the idealization of the market in the attainment of these goods, and not enough credit given to the role of cooperation. If this is correct, Gauthier says, then contractarianism is able to express the concern that we do in fact all have in maintaining the conditions necessary to make possible these and further benefits.

Changing technologies have enhanced the well-being of certain persons who contribute little (or not at all) to the mutual benefit of society. Only those who contribute can be parties to the agreement. Thus, not all persons (for example, the handicapped) are parties to the agreement. Intuitively, all persons should be participants in an agreement that justifies moral claims that apply to all of us. On Gauthier's theory, however, not all people are such participants. This is a weakness, he admits.

Despite these weaknesses however, Gauthier holds that morals by agreement is sufficient to explain why we ought to be moral: it serves our self-interest to comply with constraints that have been arrived at through a hypothetical agreement. It provides a reason for all of us to be moral.

²⁰ Ibid. p. 18.

This summary of Gauthier's view serves as an apt introduction to the two arguments that I develop in the rest of the paper. I will now turn to my first argument.

Section 2

As we have seen, Gauthier claims that the hypothetical agreement justifies morality; it provides a reason to be moral. The hypothetical agreement is a thought experiment: if we consider our existing practices, each of us would be led, ultimately, to consider ourselves in a situation devoid of any social practices or moral conceptions. From this imagined situation, we each ask ourselves which constraints we would prefer, since it is clear that some set of constraints are better than no constraints at all. The set we choose together, in light of the self-interest of each of us, forms the content of a hypothetical agreement. This explains the basis of morality; it is a test for which of our moral claims are legitimate, or justified, moral claims. The hypothetical agreement thus gives each of us a reason to be moral.

I will claim that social practices and our own moral conception are central to our self-conception. Thus, placing ourselves in a situation devoid of these features results in our imagining a person who is quite foreign to us. If this imagined person is sufficiently foreign to us, then we have reason to doubt the relevance of the agreement to us. Such is the case. Since it is no longer each of us who are parties to the agreement, two conclusions are reached. One: the question of which moral claims are actually justified

remains an open question. Two: the agreement no longer provides a reason for *us* to be moral.

2.1

On Gauthier's account, moral practices are rational to the extent that they would be agreed to by a unanimous hypothetical agreement of rational persons, even though they constrain an individual's preference fulfillment. But why, he asks in a 1991 paper, is it rational to accept only those constraints that would be agreed upon, and not rational to simply accept our existing moral practices?²¹ After all, it does *seem* rational to comply with our existing practices, some might say, since it is compliance with them that gives us a reputation whereby it is likely that others will cooperate with us on future occasions. Thus, it seems that our existing moral practices already provide us with a reason to be moral.

In answering these questions, Gauthier asks us to engage in a thought experiment by imagining a certain society. In this society, there is a set of practices that constrains the choices of the individuals of which the society is comprised. If each individual had taken into account only her desires, aims and interests, she would have chosen to always serve such, and she would have not complied. Also in this society it is *prima facie* advantageous to comply with this set of practices, since those who are not disposed to comply with them are excluded from the voluntary cooperation of the others. Such

²¹ David Gauthier. "Why Contractarianism?", in Contractarianism and Rational Choice. Ed. Peter Vallentyne. Cambridge: Cambridge University Press, 1991. p. 26.

practices are, in this society, constitutive of the morality of that society. But, he continues, we need not consider these practices *morally* justified; they are not the product of agreement. These practices simply constrain the choices of each individual in a way that each finds *rationally* acceptable. It is prudential to comply; compliance leads to mutual benefits.

Next he guides us to imagine that the persons in this society each, individually, examine the practices that constitute their morality. They will test these practices against their own desires, aims and interests (with the knowledge that the others are doing the same). At this point they may begin to question why the existing practices are the ones with which this society complies. Why exactly *these* practices and not others? For, of course, they will recognize that their set of practices is not the only set that will yield mutual benefits; there are many possible moral orders.

Each person in this imagined society can expect to benefit from a disposition to comply with the existing practices of his society. For one who ignores such practices would surely be worse off than one who did not. Since each may have a reason to reject the status quo, two options become clear for each individual: will I benefit more from abandoning morality entirely (and not recognizing any constraint), or will I benefit more from rejecting some set of constraints in favor of an alternative set? Since cooperation will be considered necessary for mutual benefit, the latter option will be deemed superior. At this point, then, the question becomes not whether it is rational to comply at all, but which set of practices is the rational set to choose.

To answer this recent question, each person will examine his own prospects among those sets that he expects are likely to be arrived at from agreement with the others in his society. If his prospect is improved by a set other than the existing one, he will have a real reason to demand a change in the existing moral order. To say this in a more general way: if there are any persons whose prospects could be improved by an agreement, then the existing set of practices is recognized by such a person to be *unstable*.²² Of course those whose prospects will worsen upon an agreement will resist change; they will insist on the status quo. But if each is fixed on maximizing her own preference-fulfillment, this insistence is likely to be ignored by the others.

This thought experiment is relevant to us in the following way. In the actual world, we begin such deliberations from within an existing set of moral practices which constrains our actions. From reflection on such fact, we are led to the notion of an amended set of practices that would obtain the agreement of everyone; this set will be recognized by each to be more *stable* than the existing set. Thus, by reflection, we, as rational deliberators, are led

²² The task of clarifying Gauthier's technical notion of *stability* is in order. *Stability* is an ideal of rationality against which both practices and agreements are adjudged. A practice or agreement is judged to be stable to the extent that it expresses the way in which we would reason about our actions in a natural (or pre-social) environment. Thus, a pre-moral, pre-social consideration of our preferences is considered *stable*. In Gauthier's imagined society, the existing practices were judged to be *unstable* because they didn't adequately reflect how we would reason, were we free from the influence of others. So, a purely stable practice or agreement proceeds from ourselves.

from existing practices to that set of practices that would receive everyone's assent.

By our capacity for reflection, we will be led from here to a consideration of which practices would be agreed to in a pre-moral, pre-social context. For just as *stability* was gained by considering which sets of practices would be agreed to in our *actual* context, even more *stability* will be gained when we consider which practices would be agreed to in a context devoid of *any* actual practices. Put in a different way: just as the status quo turned out to be considered *unstable* in comparison with those sets of practices that would receive everyone's agreement, so this latter agreement turns out to be conceived as *unstable* when compared with what would have received agreement in a pre-social context. Just as *existing* practices seemed to be arbitrary to the extent that they did not correspond to what a rational person would agree to, so the supposed agreement seems arbitrary in relation to what would be accepted in a situation with *no* practices. This arbitrariness should be avoided because it blinds us from our motivation to be moral: it ultimately serves our self-interest.

Gauthier continues his argument as follows. What a rational person would agree to in her actual context depends on what her position is in relation to the others making the agreement with her. Just what does she bring to the table? This negotiating position is affected by the practices of her society. In short, while agreement may well yield some set of practices differing from the existing set, the agreement *itself* will be influenced by that

prior set of practices, which themselves are not products of agreement. This fact must call into question the rationality of the agreed-upon practices. For if a practice is rational so long as it would receive agreement, and the terms of the agreement are tainted (or considered to be *unstable*) by existing moral practices, we have reason to say that the agreement is impure. The agreement then seems arbitrary as well. "The arbitrariness of existing practices must infect any agreement whose terms are significantly affected by them."²³

While rational agreement is a source of *stability*, the *stability* of it is undermined by the circumstances in which it takes place, if such circumstances are arbitrary. To get around this arbitrariness, individuals will be led to move from considering the agreement as actual to considering it as a hypothetical agreement. Individuals will consider which practices they would have agreed to from a position not influenced by existing practices. Such a position is purely rational. From this pre-moral, pre-social context, the agreement regains its *stability*.

Stability is what links agreement to compliance. A stable agreement gives us a clearer reason to be moral.

2.2

Central to Gauthier's recently presented argument is the assumption that imagining ourselves in a pre-moral, pre-social context is a possible task.

²³ Ibid. p. 28.

It is not. I will seek to make this claim plausible. Then, I will explain its implications for Gauthier.

A central part of our self-conception consists in the roles that we play.²⁴ We are all sons or daughters; many of us are parents, brothers, sisters, friends, and colleagues. Some roles are chosen reluctantly, others willingly; still other roles are not chosen at all. We have a real choice of whether to play the role of friend to a person. We may not like one of our co-workers, yet we willingly choose to play a role in the story of our co-workers, since we value our continued employment, and thus, we must continue to interact with this disliked co-worker.

Each role is a part of the narrative of some person or persons. We all play one, or several, main or supporting roles in a network of interrelated stories. The importance of our roles to the story of which they are a part is of course contingent on the narrator(s) of each story.

Some roles are partially defined by certain actions we perform. I make a purchase at a convenience store; I play the role of a customer of that store—at least for one visit. I have a minor role in the story of that convenience store. I ride the train; I play the role of a commuter. Once again, I play a miniscule role in the story of public transit. I play a slightly larger role in the story of public transit in my region. I play a still larger role in the story of the particular train on which I am riding. Note that in these

²⁴ I am indebted for much of these thoughts to Alisdair MacIntyre, see After Virtue. Notre Dame, IN: Notre Dame University Press, 1981. See especially chapter fifteen. While I credit him for much of the stimulation of these concepts, they have been extended in the present paper.

examples, we need not be aware of the role we are playing at the time we are playing it.

Roles partially define social practices; conversely, practices partially define roles. We cannot conceive of a racquetball game without racquetball players. A man hitting a ball against a wall with a racquet is not playing racquetball. The practice of racquetball needs multiple players. Racquetball is a practice defined by rules; it is only in the context of the practice of racquetball that the role of a racquetball player is conceivable.

Many practices have a conception of constraint that sustains and partially defines them. Dining at a restaurant requires constraints of the staff, management and patrons alike. The staff must perform certain duties; management must ensure that the staff is performing said duties; patrons must behave in at least a minimally courteous manner. Constraint makes the practice of dining possible.

Like practices, roles possess the quality of constraining certain of our choices. The role of an employee requires fulfilling certain job requirements. These requirements may be ill-defined, yet are central to the definition of any job.

These recent comments are intended to point to the relations of roles, practices, and the constraints necessary to sustain both. They form an interlocking network; each relies on the other for their mutual intelligibility.

I said above that roles form a crucial part of our self-conception. Now it is clear that practices too are of some importance to that conception. For

without practices, the importance of roles to our lives might seem unimportant. But they clearly are, as I have shown. In the same manner, constraint is a part of our self-conception. Lacking constraint, roles and practices cannot be sustained. They are not even intelligible.

With these considerations in mind, I am now in a position to ask of myself if I can accomplish what Gauthier asks of me in his thought experiment. Can I imagine *myself* in a situation in which I have *no* association with other persons, and without *any* aspect of the interrelated network of roles, practices and constraints so integral to our self-conception? The situation described by Gauthier must have these features. A pre-social situation is, of course, one in which there is no association between persons. A pre-moral situation is one in which there is no constraint. So, is it possible to conceive of such a situation?

Certainly, we can imagine such a situation. It is easy to imagine a person existing, in the actual world, without social practices or the moral conceptions which underpin them. However, if I am right in my claims above about the interlocking nature of roles, practices and constraints, then it seems that this imagined person must lack these features. And since these concepts are crucial to our self-conception, then abstracting these concepts away from ourselves, in the process of reflection endorsed by Gauthier, results in our imagining a person who is very, very different from us. For the process of imagining a person who lacks some of the defining characteristics of who I am results in a person who is not me!

Now, if I were to ask this imagined person, abstracted of all conception of roles, practices and constraints, to choose which set of constraints would serve his self-interest, any answer provided seems irrelevant to that set of constraints that would serve my self-interest, since this person is sufficiently foreign to me. Moreover, since this chosen set of constraints forms the content of the hypothetical agreement, it too loses its relevance to me.

In short, abstracting away from our self-conception in the way that Gauthier asks of us leads each of us to an imagined person. Since this person's choice of constraints forms the basis of the hypothetical agreement, and his choices are irrelevant to each of us, the agreement itself is irrelevant to each of us. Moreover, since the agreement is irrelevant, and it is supposed to provide a reason for each of us to be moral, it turns out that whatever reason is suggested in this regard is irrelevant to us as well. Gauthier's morals by agreement turns out to be irrelevant to our individual lives.

2.3

A contractarian might object at this point. I am assuming, he might object, that it is impossible to conceive of ourselves in a different situation than the ones we find ourselves in. But clearly, I can imagine myself living in a fourteenth-century medieval society. But if this is possible, then what difficulty is there in imagining ourselves in a pre-moral, pre-social situation?

I must make clear that I am arguing for the impossibility of *fully* conceiving ourselves in a pre-moral, pre-social situation. From this claim,

however, it does not follow that we cannot conceive of our own existence in such a state. It certainly is possible to imagine the essence of myself in a situation in which I have no contact with others, and no conception of constraint. And such a person would indeed be partially recognizable as myself. But this imagined person lacks the features necessary to make the decision that Gauthier asks us to make. Gauthier makes the mistake of assuming that the *essence* of our selves are capable of possessing the desires relevant to the choice of constraints. This is mistaken.

Our desires, however, are pieces of our individual stories. As such, they cannot be considered in isolation from the network of roles, practices and constraints of which they are a part. Let us suppose that Gauthier is correct in asserting that our essential selves are capable of possessing desires. If this were the case, then it appears that at least some of our desires are an essential part of who we are. It would also be true that these desires would be unchanging, since this is part of the meaning of an essential self.

But most of our desires do change. Remember what your desires were like in your childhood. Are any of those desires the same as your current desires? Probably very few of your desires are the same as when you were a child. This changing nature of our desires suggests that almost all of our desires develop, deepen, and alter throughout our lives. They are not essential features of our selves.

It seems, then, that a better account of our desires is that they form a crucial piece of the story of our lives. This claim accounts for the changing

nature of our desires. If we were to abstract away from the story of our lives, conceiving of only our essential selves, then it is difficult to imagine which desires (if any) we would have. Now, suppose that there were some desires that remained as features of our essential selves. How would we know them? How could we distinguish which desires are the essential ones? Since it is entirely unclear, therefore, which desires our essential self would possess, how is it that we can ask our essential self which constraints would be preferred by this essential self? This is impossible. Were this possible, a justification of constraint would be possible. But such justification would come at the cost of our actual desires, which are a crucial part of our lives. This justification, then, would remain irrelevant to our lives.

In short, fully imagining ourselves in a pre-moral, pre-social situation is impossible. If it were possible, then our essential selves would possess desires. They do not. But even if they did, which constraints would be chosen by our essential self would be impossible to know, since it is entirely unclear which of our desires would belong to our essential selves. In any case, then, Gauthier's hypothetical agreement remains irrelevant to our lives.

2.4

In the next section, I will provide an argument that will serve as a challenge to Gauthier's conclusion that self-interest grounds our moral claims. Moral rationalism is the view that morality is able to provide reasons for action. If this view is correct, then morality is able, directly, to provide reasons for action. If this is true, then it is difficult to see the need for

Gauthier's supposition that by abstracting from our moral practices, we are able to *fully* account for reasons for action. For, if our moral practices *already* provide normative reasons for action, regardless of whether or not it serves our self-interest, then there is no need to demonstrate, as Gauthier supposes, that morality proceeds from our preferences in order to also demonstrate that it gives us reasons.

On moral rationalism, morality gives us reasons directly. For Gauthier, morality provides reasons only because it satisfies our preferences. It seems important, then, that he is able to demonstrate this. If however, moral rationalism is correct, then such demonstration is unnecessary. The next section is devoted to expanding this argument.

Section 3

Gauthier's thought experiment, discussed in the last section, is presented in a 1991 paper. The thesis of that paper is that contractarianism provides the most plausible resolution to the foundational crisis that morality faces.²⁵ His resolution to this purported crisis is implausible. By putting contractarianism in conversation with moral rationalism, we are able to see that *deliberative justification*, his proposed resolution to the foundational crisis, implies an implausible consequence about moral evaluation.

I will begin by briefly sketching Gauthier's argument for *deliberative justification*. I will then present one argument for moral rationalism. I will

²⁵ Gauthier, p. 15.

attempt to bolster this latter argument, and then to show that, in light of it, *deliberative justification* leads to an implausible consequence about moral evaluation. I will conclude by pointing to the difficulty of answering any questions about what grounds our moral claims, and to one view that might be taken in light of this difficulty.

3.1

“Morality faces a foundational crisis. Contractarianism offers the only plausible resolution of this crisis.”²⁶ Gauthier presents this straightforward thesis. The foundational crisis of morality, for Gauthier, is that there is an inconsistency between what our moral language presupposes—“objective values that help explain our behavior, and the psychological states—desires and beliefs—that, given our present world view, actually provide the best explanation”²⁷ of our behavior. This worldview affirms both that it is our preferences which explain our behavior and that morality is conceived as a set of constraints on such behavior. If we are primarily disposed to do what we prefer, then morality, constraining our preferences, stands in need of justification. Otherwise, it stands in danger of being considered as no “more than an anthropological curiosity,”²⁸ a relic of a bygone age. The project of contractarianism, therefore, is to provide a justification of morality that is consistent with our present worldview.

²⁶ Ibid. p. 15.

²⁷ Ibid. p. 16.

²⁸ Ibid. p. 16.

In short, Gauthier conceives of the crisis of morality as that of *justifying constraint*. “From the standpoint of the agent, moral considerations present themselves as constraining his choices and actions, in ways independent of his desires, aims, and interests . . . [This] reveals clearly what is in question—the ground of constraint.”²⁹

For Gauthier, such a crisis of morality is a part of our present worldview. This worldview provides the best explanation of our behavior. Such explanation takes our actions as a direct result of our preferences. We conceive of ourselves as agents who make choices and who perform those actions that tend to maximize preference fulfillment. This is integral to our “deep sense of self;”³⁰ is partly constitutive of who we take ourselves to be. Since preference-determined action is basic to our self-conception, any constraint on our actions (and consequently, our preferences) must be justified. Hence, morality must be justified. This argument should be familiar by this point to my readers.

Contrary to this conception of morality is that of an older worldview that has been abandoned: there exists objective values and these values help explain our behavior. For example, the claim, “genocide is evil” is seen to be, in some sense, true of the world. Our moral language, by and large, reflects this view of the world as containing true, objective moral claims. However, Gauthier insists that this view of the world is inadequate to explain certain facts about our behavior.

²⁹ Ibid. p. 16.

³⁰ Ibid. p. 18.

This line of thinking is closely aligned with that of Gil Harman.³¹ For Harman, if values were real features of the world, they would possess some explanatory power. However, they lack this power. Therefore, values are subjective.

Gauthier too holds that values are subjective. They are relative to each individual's considered preferences.³² He holds, like Harman, that if objective values were to exist, they would explain certain facts about our behavior. Since the worldview that explains our behavior in terms of objective values is insufficient, we have abandoned it for one that denies objective values and instead seeks to explain our behavior in terms of our preferences.

But why, we might ask at this point, should our preferences be constrained at all? No such justification is available in our present worldview, as it was on the older one, continues Gauthier. What possible reason, then, could we have for accepting any constraints on our preferences? "[W]hat justifies paying attention to morality, rather than dismissing it as an appendage of outworn beliefs? We ask, and seem to find no answer."³³ There simply is, asserts Gauthier, no extra-moral foundation upon which to answer such questions. There may be a moral foundation that provides an answer; we may have a moral reason for being moral. But there

³¹ See Gilbert Harman, Explaining Value: And Other Essays in Moral Philosophy. Oxford: Clarendon Press, 2000.

³² He develops the notion of a *considered preference*; this is the result of sufficient reflection on our preferences. See Gauthier, Morals By Agreement, chapter 2.

³³ Gauthier, p. 16.

is no way of answering such questions that doesn't already presuppose a moral framework. However, since there is no moral framework, these questions are unanswerable. Thus, it is in this sense that the crisis of morality is foundational. The language of morality assumes a conception of morality; our explicit worldview expresses a differing, and superior, conception.

Gauthier proceeds to provide a reply to an objection. Perhaps morality needs no justification, rather it is the job of moral philosophy simply to systematize our moral judgments "and so to give us a deeper understanding of what moral justification is."³⁴

The problem with this, Gauthier says, is that such an objection recognizes only one mode of justifying our actions: that of justifying our actions within a moral framework. But this clearly is an inaccurate picture of the way that we usually justify our actions. We primarily engage in *deliberative justification*. This is understood as follows. In making choices, we often have before us two conflicting desires. In such a case, "[w]e order our preferences, in relation to decision and action, so that we may choose in a way that maximizes our expectation of preference fulfillment. And in so doing, we show ourselves to be rational agents, engaged in deliberation and deliberative justification. There is simply nothing else for practical rationality to be."³⁵

³⁴ Ibid. p. 18.

³⁵ Ibid. p. 20.

The deliberative justification account of practical rationality, continues Gauthier, is part of our present world-view. He takes this account to be rationally more fundamental than the objector's view. It appears then, that morality (as a mode of justification purportedly on par with deliberative justification) is pushed on the defensive, and must explain why it can make any claims on those who refuse to recognize the existence of *any* moral framework. For Gauthier, there simply is no extra-moral foundation that is able to answer them.

So, how is it then that morality can be justified? If morality is to be saved, "[t]he first step is to embrace deliberative justification, and recognize that morality's place must be found *within, and not outside*, its framework."³⁶ How is this to be done? The resolution is his version of contractarianism. On this account, the set of constraints which constitute morality are limited to only those actions which maximize preference fulfillment. In a society, however, our preferences often conflict with those of others. Thus, only those actions that would be agreed by everyone to be mutually advantageous are moral actions. Thus, the constraints on our behavior are voluntary. Morality is then justified.

We have then, for Gauthier, a foundational crisis of morality. That is, morality, if it is to be saved, must be justified in light of a more accurate conception of how we make choices: deliberative justification. It is thus that contractarianism is able to resolve the crisis.

³⁶ Ibid. p. 22. Emphasis mine.

3.2

By placing deliberative justification in conversation with moral rationalism, the former begins to appear less plausible as a resolution to the foundational crisis. I will expand this claim by giving some brief explanation of moral rationalism. Then, I will present one simple argument for its plausibility, as outlined by Russ Shafer-Landau. At the end of this discussion, moral rationalism will be seen to be superior for two reasons. One: it provides a plausible account of moral evaluation (where deliberative justification will be shown to fail). Two: it allows us to be silent about whether or not there exist objective values.

“Moral rationalism is the view that moral obligations entail reason for actions.”³⁷ Suppose my friend Alison performs an act. She explains that she did it because it was the right thing to do. Suppose that she is correct about this; the reason she performed the act was because it was the right thing to do. Now, what justifies her in performing this act? Usually, in such a case, the justification of her act is not called into question. But if an act’s rightness did not provide a reason for performing it, then it seems that we would call it into question. Now, if moral rationalism is false, then the justification that Alison provides for her action would be incoherent. Since it is not incoherent, moral rationalism is not false. These considerations imply that the rightness of an action is a good reason for performing it.

³⁷ Russ Shafer-Landau, “Moral Rationalism” in Ethical Theory. Malden, MA: Blackwell Publishing, 2007. p. 174.

Now, Shafer-Landau says, consider immoral acts. When we assess someone's behavior as morally unjustified, it is implied that she has violated some moral order. Suppose that this, or any other, moral order did not provide a reason to act. Then, we no longer have a reason to avoid certain actions solely on the grounds that they are immoral or improper. But, intuitively, this seems both conceptually confused and very unfair. To say, for example, that Alison ought not to have stolen from her grandmother because it violated some code of conduct, and still affirming that she had no reason to not steal from her grandmother, seems conceptually confused. And it seems unfair to blame Alison for stealing, and yet also to admit that she had no reason not to steal. "The fairness and appropriateness of moral evaluation rest on an agent's attentiveness to reasons."³⁸ If a person claims correctly that she did not ignore any reasons to act, she cannot be blamed for violating any moral code. This is a plausible claim. But this claim is true only if moral rationalism is true.

We are left, concludes Shafer-Landau, with two choices. One: endorse moral rationalism. Two: admit that an agent's attentiveness to reasons has no bearing whatsoever on moral evaluation. Those who deny moral rationalism, and commit themselves to the second option are further committed to two tasks. They must explain what serves as a basis for moral evaluation other than reasons. They also must explain how this basis avoids

³⁸ Ibid. p. 174-175.

the seeming unfairness of blaming persons for actions they had no reason to avoid.

Now, if moral rationalism is correct, then Gauthier's deliberative justification seems implausible. For deliberative justification affirms that morality does not provide reasons for actions. The ultimate reason for performing any action, according to contractarians, is whether or not that action serves our self-interest by maximizing our preference-fulfillment.

So, is moral rationalism correct? I will now seek to bolster Shafer-Landau's argument by considering an example. Suppose that my friend Alison is asked to explain why she donated money to her favorite charity. She explains that she did this because it was the right thing to do. Suppose again, that she is not lying about this; this truly was her reason for acting. If the rightness of donating her money in this way did not provide a reason for her to do it, then she would not be admired for doing it. Suppose that instead of citing the rightness of donating as her reason, she said that she did it simply because she preferred to do it, given the circumstances. It seems implausible to say that she would be admired for simply maximizing her preference-fulfillment in this way.

Now suppose that Gauthier is correct in asserting that deliberative justification is the fundamental mode of how we reason practically. If this were the case, then the rightness of donating would provide no reason for Alison to donate the money that she did. Her reason for donating the money would then be simply because she preferred it. Then, clearly, she would not

be admired for simply doing what she preferred. Likewise, if Alison were to steal money from her grandmother, she did this because she preferred it. And, similarly, it is unlikely that stealing money in this way would be disapproved. Clearly, this is both conceptually confused and unfair.

3.3

Let me consider a few objections. The first objection to the claim that moral rationalism is true denies Shafer-Landau's claim that, were moral rationalism false, we would be left with affirming conceptually confused statements regarding the moral evaluation of an action. Shafer-Landau claims that to say, for example, that someone should be kind to his friends, while at the same time affirming that he has no reason to be kind to his friend is conceptually confused. But we need not be confused, says this objector. Perhaps, instead of affirming moral rationalism, we can evaluate an action from two perspectives: from the point of view of reasons and from the point of view of morality. So, from the point of view of morality, being kind to your friends is the right thing to do. But from the point of view of reasons, if someone has no desire to be kind to her friends, then she has no reason to act in this manner. We can affirm both statements without confusion. So goes the objection.

If we deny moral rationalism in this way, however, the confusion would remain. If we look at a case of extreme wrongdoing, it is clear that Shafer-Landau's argument is unaffected by this objection. When we say that someone should not have sexually molested her own children, we are both

saying that it is wrong for her to have done so, and that she had a real reason to avoid doing this. However, if moral rationalism is false, and she did not have a reason to avoid molesting her children, then we could not disapprove of her action. But, of course, we do disapprove of her action. It is clear that she had a reason to not molest her children: because it is wrong.

It is equally clear that an appeal to desire in this case would be inappropriate. Suppose that morality didn't provide reasons for action, and preference-fulfillment provided the basis of moral evaluation instead. If this were the case, then we would have to say that parents who do not molest their children, avoid doing so simply because they prefer not to do so. Most would affirm, it seems, that parents have a reason, regardless of their desire, not to molest their children: simply because it is wrong.

Now, if we were to evaluate the above action in the way suggested by the objector, we would affirm the following two separate propositions about one action. One: the parent was wrong to have sexually molested her children. Two: she had no reason to avoid sexually molesting her children. It seems clear to me that almost no one would affirm the second proposition. While it may appear that in some cases (such as being kind to your friends), that the suggested analysis might work, when we consider an extreme case such as this one, it becomes clear that there is an intimate connection between the claims of morality and the reasons it provides.

A consideration of this extreme case also provides an answer to a second possible objection. One might object that the rightness of an action

only provides a reason for those who already have another reason to be moral. For those who view themselves as outside any moral framework, however, adherence to morality might appear as a kind of fetish. Since such persons have no desire for morality, the rightness of an action fails to provide a reason for them to perform that action.

But if we consider that no one would affirm that someone would ever have a reason to molest her children, then it becomes clear that morality is more than a fetish; there are some claims that apply to all. The questions of which moral claims are true, and which moral claims apply to whom and under which circumstances have not been discussed. It is clear, however, that morality provides reasons for action.

3.4

These considerations about moral evaluation point to the plausibility of moral rationalism and the implausibility of deliberative justification. But where does this leave the foundational crisis, with which I began the section? If moral rationalism is correct, then it appears that morality does not indeed stand in need of justification.

But, we might ask, what about the apparently strong intuition that morality does indeed need justification? This line of thinking is of course not new. Not only can elements of it be found in much of twentieth-century moral philosophy; we can think of Hume, of course, and before him Hobbes. It can also be traced further back to Thrasymachus of Plato's Republic. We

can also identify many of our friends and family as a part of this tradition.

What can the moral rationalist say to these people?

It seems, of course, that there is a competing intuition that might also be strong; many people have no trace of moral skepticism at all. Many of these same people simply comply with the moral practices of their society without question. For these people, morality needs no justification.

These considerations suggest that perhaps we should say nothing to the moral skeptic. Certainly, it is possible to argue that one moral code is superior to another. But perhaps morality, on the whole, is such that no justification of it can be given. Morality might simply be a crucial part of who we are, as humans.

I claimed at the beginning of this section there are two advantages that moral rationalism has over deliberative justification. It avoids the implausible claim about moral evaluation discussed in Section 3.2. But the second advantage is that it allows us to remain silent about the existence of objective values, which was one of Gauthier's critiques of the older worldview. Shafer-Landau reminds us that questions about the ground of normativity are prevalent, and very difficult to answer. He also admits that moral rationalism fares no better than other accounts in attempting to answer these questions.³⁹ Perhaps these questions will always remain a feature of our inquiries into morality. But by affirming moral rationalism, we

³⁹ Ibid. p. 179.

are able to simply assert that morality is able to provide reasons, without committing to whether or not moral claims are “objective,” or “subjective.”

Conclusion

In this paper, I have presented two arguments against David Gauthier’s contractarianism. The first argument was as follows.

Gauthier’s hypothetical agreement is the object of a thought-experiment wherein we are asked to consider which constraints we would prefer in a pre-social, pre-moral situation. A consideration of constraints from this situation proves more stable, since it relates deeply to how we conceive of ourselves. However, I have argued, that since roles, practices, and the constraints that sustain and partially define them are central to our self-conception, the imagined person resulting from an abstraction of our social situations is sufficiently foreign to us so as to render the hypothetical agreement irrelevant to us.

The second argument was that since deliberative justification implies an implausible view about moral evaluation, it fails as a resolution to the foundational crisis. Moral rationalism was presented as an alternative superior resolution.

I have sought to make plausible these claims in the hope of weakening the cumulative case for Gauthier’s contractarianism. If I am correct in my claims, what does this mean? If my claim that roles, practices and constraints are central to our self-conception, and that the hypothetical agreement fails to attain any relevance to us, as actual beings, it seems that I have begun a

sketch for a richer account of us as actual, moral beings. This account begins from the plausible intuition that most of us conceive of ourselves as existing within a moral framework, and that this framework is interwoven into our self-conception.

On this account, then, a justification of morality is not necessary. Nor is this descriptive account intended (primarily) to provide any answer to the moral skeptic. It may, however, perform these functions. If this account is developed, the moral skeptic may perhaps come to see morality as basic to her own self-conception, and may perhaps from this infer that she has a good reason to be moral.

Am I evading the question? Is it part of the role of the moral philosopher to provide a justification of morality? In Section 2, I argued that the interlocking network of roles, practices and constraints are central to our self-conception. If I am correct about this, and my recent comments are on target, then it seems that morality is simply a part of who we are. Maybe this is all we can say on the matter. After all, we must end our inquiries somewhere, just as the present paper has now ended.

Bibliography

Baier, Kurt. The Moral Point of View: A Rational Basis of Ethics. Ithaca, NY: Cornell University Press. 1958.

Gauthier, David. Morals By Agreement. Oxford: Oxford University Press. 1986.

-----, -----, "Why Contractarianism?", in Contractarianism and Rational Choice. Ed. Peter Vallentyne. Cambridge: Cambridge University Press. 1991.

Gough, J. W. The Social Contract: A Critical Study of Its Development. Oxford: Clarendon Press. 1936.

Grice, G. R. The Grounds of Moral Judgement. Cambridge: Cambridge University Press. 1967.

Hampton, Jean. Hobbes and the Social Contract Tradition. Cambridge: Cambridge University Press. 1986.

-----,-----, "Two Faces of Contractarian Thought" in Contractarianism and Rational Choice. Ed Peter Vallentyne. Cambridge: Cambridge University Press. 1991.

Harman, Gilbert. Explaining Value: And Other Essays in Moral Philosophy. Oxford: Clarendon Press. 2000.

Locke, John. Two Treatises of Government. London. 1690.

MacIntyre, Alisdair. After Virtue. Notre Dame, IN: Notre Dame University Press. 1981.

Narveson, Jan. The Libertarian Idea. Philadelphia: Temple University Press. 1988.

Rawls, John. A Theory of Justice. Cambridge, MA: Belknap Press of Harvard University Press. 1971.

Scanlon, Thomas. What We Owe To Each Other. Cambridge, MA: Harvard University Press. 1998.

Shafer-Landau, Russ. "Moral Rationalism" in Ethical Theory. Malden, MA: Blackwell Publishing. 2007.