

FREEDOM HAS NO INTRINSIC VALUE:
LIBERALISM AND VOLUNTARISM

ABSTRACT: Deontological (as opposed to consequentialist) liberals treat freedom of action as an end in itself, not a means to other ends. Yet logically, when one makes a deliberate choice, one treats freedom of action as if it were not an end in itself, for one uses this freedom as a means to the ends one hopes to achieve through one's action. The tension between deontology and the logic of choice is reflected in the paradoxical nature of the "right to do wrong"; and in Rawls's unsuccessful attempts to justify to acting agents their interest in such a right, embodied in his conception of agents as "self-validating sources of claims." A self-validating source of claims is akin to God as envisioned by voluntarist theologians such as Ockham. Leibniz's critique of voluntarism, then, is applicable to the Rawlsian subject: like the voluntarist God, she would be unable to act if it were indeed the case that her action validated itself.

Is freedom—I shall be speaking of freedom of action—an end in itself, or a mere means to other ends? It has been treated as intrinsically valuable by a long line of political theorists culminating, in our own day, in Rawls¹—for whom the logic of the original position dictates that everyone be allowed to pursue whatever ends they prefer. The

Jeffrey Friedman, Critical Review Foundation, P. O. Box 869, Helotes, TX 78023, thanks, for their encouragement and comments on earlier drafts, Ronald Beiner, Erik Davis, Paul DeHart, Shterna Friedman, Paul Gunn, Will Kymlicka, James Bernard Murphy, the late Leszek Kolakowski, Stephen P. Macedo, Russell Muirhead, Christopher Oppermann, Jacob Roundtree, Amartya Sen, Kevin Vallier, Jeremy Waldron, and respondents to presentations at the Northeast Political Science Association meetings, Boston, 1996; the Dartmouth Political Theory Workshop, Spring, 1998; the American Political Science Association meetings, Boston, 1998; and the Harvard Social Studies Workshop, Spring, 1999.

implication of what I will call the logic of choice, in contrast, is (at least arguably) that the freedom to act is reducible to the intrinsic value of the ends pursued by acting. Therefore, freedom has no intrinsic value of its own.

I will explicate the logic of choice by arguing against the deontological-liberal “right to do wrong.” A right to do wrong embodies and protects freedom of action even in pursuit of bad ends or wrong ones. In contrast, someone with free will does not knowingly do what is bad or wrong. So there is a tension, similar to a performative contradiction, between an agent’s free actions and her assertion that freedom of action is intrinsically valuable. In acting freely, an agent affirms the perceived value of the conception of the good that ultimately provides her criterion of choice among possible actions. She thereby denies that her freedom of action is anything but an instrument to the achievement of this putatively good end.

In making this argument, I will not necessarily be endorsing the paternalistic enforcement of “the good.” For obvious reasons, one might not want to give people or governments such enforcement power. Even granting that the paternalist knows (and intends) the good, efforts to impose the good might backfire in a complex world, resulting in more harm than good. In an imperfect world, we might want to treat people *as if* they have the right to do wrong.

It is safe to say, however, that most contemporary liberal theorists believe that people really do have a right to do wrong—not merely that we should treat them as if they do. Part I of the paper reviews a debate between Jeremy Waldron and William Galston about whether there is such a right. What is really at issue in contesting this right, I argue, is the logic of choice. I use several examples to show that from the standpoint of a freely choosing agent, the right to do wrong can never be desirable in itself, because acting on such a right is possible only if one has confused what is right, or good, with what is wrong, or bad.

The heart of the matter, then, is that in making a choice about which action to take, one relies, however implicitly, on what one views as knowledge of a criterion of a good or desirable action. Thus, the problem of choice is inherently cognitive, and the logic of choice is inherently epistemic—a matter of looking for (and thinking one has found) a criterion for action that allows the options for action to be sorted according to their relative desirability. In this view, freedom of action is, at best, instrumentally valuable to the pursuit of what one

perceives to be a good end. A right to do wrong, by contrast, treats the freedom to pursue any end—even a bad one—as intrinsically valuable. Such a right cannot be freely acted upon, because it prescind from the cognitive issue at the heart of any free choice.

In Part II, I try to determine whether the logic of an agent's choice is too narrow a lens through which to view the intrinsic value of freedom. To that end, I analyze Rawls's attempt to combine the perspective of the agent with a deontology of free action. The latter is embodied in principles of justice that impartially defend agents' freedom as a matter of principle, not (as in consequentialist liberalism) as a matter of achieving good ends to the maximum feasible extent. Impartiality among ends is guaranteed by screening out of the original position not merely contractors' knowledge of their historically and genetically contingent place in society, but also their putative knowledge of the good. I call the latter function of the original position "veiling the good."

Rawls has an excellent argument for treating contingent historical and genetic endowments as arbitrary, and thus as appropriately screened out by the veil of ignorance. One is not responsible for one's genes and other favorable (or unfavorable) circumstances and abilities, which therefore do not confer desert; thus, one should not consider these endowments in evaluating the basic structure of society. Part II argues that Rawls's initial attempt to make a parallel argument regarding conceptions of the good commits him to viewing these conceptions as arbitrary and groundless, which is incompatible with the logic of choice. In writings that soon followed *A Theory of Justice* (1971), however, Rawls attempted to square the logic of choice with veiling the good. The effect was to transform the inherently epistemic logic of free choice into an inherently anti-epistemic logic of "self-authenticating" choices (Rawls 1993, 32, 72), according to which a free person is, as such, a "self-originating sourc[e] of valid claims" (Rawls 1980, 543). Therefore, the "conception of the person" advanced by Rawls after *A Theory of Justice* can best be understood as depicting an agent who violates the logic of choice by creating rather than perceiving the desirability of her actions.

By the end of Part II, then, we will have moved from a right to do wrong and the intrinsic value of freedom to liberal voluntarism. Only voluntarist agents, whose actions are self-validating, could have a right to do wrong, for self-validating actions render the concept of "wrong" action moot. In Part III I sketch the roots of voluntarism and its

late-medieval application to God's will. Then I apply to liberal voluntarism Leibniz's profoundly original critique of theological voluntarism. This argument shows that if one's actions are both free and self-authenticating, then one would, like Buridan's Ass, be unable to act. Therefore, the intrinsic value of freedom of action, like the right to do wrong, is incompatible with free action itself. Leibniz allows us to see that individual action as conceived by deontological liberalism must be either heteronomously determined (and thus unfree) or inexplicable: the same dilemma that faces us if we attempt to explain the actions taken by the voluntarist theologians' God.

Since it was Michael Sandel (1982) who first pointed to the voluntarist element in Rawls's theory, the paper concludes with a brief comparison of my argument to Sandel's (Part IV). Sandel saw the problem as Rawls's location of moral authority in the individual; in response, he proposed locating it in the community. I argue, however, that liberal voluntarism is no more the result of excessive individualism than theological voluntarism was the result of excessive theism. The problem, instead, is in both cases a misconception of the nature of choice. Thus, Sandel merely substitutes the community for God or the individual as the self-originating source of valid claims. In Sandel's communitarianism, as in deontological liberalism, the good is subordinated to the right—the right of the community to do what might otherwise be considered wrong. The consequences of this view were so appalling that Sandel repudiated communitarianism in the second edition of *Liberalism and the Limits of Justice* (1998). But he did not repudiate the meta-ethics that, in conferring moral sovereignty on one or another type of authority, entails that this authority's evaluations are either inexplicable or heteronomously determined.

I. THE "RIGHT TO DO WRONG"

In the terminology of deontological liberalism, the "priority of the right over the good" means that conceptions of the good must not conflict with the equal freedom of everyone to pursue their chosen ends (Rawls 1971, 31). This freedom of action is enforced by the basic laws and institutions of a liberal society, and in Rawls's version of liberalism, it is embodied in the two principles of justice. Acting on a conception of the good in a manner that violates others' right to freedom of action is therefore unjust and thus (loosely speaking) "wrong." But conceptions

of the good that do *not* violate others' rights are at worst (in the opinions of those who hold different conceptions of the good) "bad."

Waldron's paper, "A Right to Do Wrong" (1981), therefore departs from deontological liberal usage in calling "wrong" those actions that are based on an unsound conception of "the good." But Waldron's terminology is in line with ordinary language; and in this case, ordinary language clarifies more than it obscures. In ordinary terminology, "right" and "wrong" are antonyms. A "right" to do "wrong" is therefore paradoxical on its face, reflecting the contradiction between saying that something *should not* be done, on the one hand, because it is based on a mistaken conception of the good; and saying, on the other hand, that the same thing *should* be done if someone chooses to do it. This *prima facie* contradiction, which gives rise to Waldron's reflections, is obscured if we employ standard deontological nomenclature, which would transform Waldron's *problematique* into a not-so-problematic-sounding "right to do bad."

To problematize the right to do wrong is to problematize deontological liberalism itself. It is therefore important to follow Waldron's departure from deontological usage and see where it leads.

Waldron (1981, 64) points out that in everyday language it is paradoxical to say that "the action in question is morally wrong, but nevertheless it is an action that the agent has a moral right to do." For "if an action appears arbitrary or capricious" to an observer of the action, it "is not made to appear one iota more reasonable or defensible" if, when questioned about the action by the observer, the agent replies, "I had a right to do it; I was exercising my right" (ibid., 73). This is the crux of the problem with the right to do wrong. Why should one be free to do what one should not do?

Nonetheless, Waldron (1981, 81) concludes that there *is* "a right to do wrong." For such a right protects "individual choices [that] are seen as crucial to personal integrity." Unless there is a right to do what is wrong, Waldron continues, then on all the important questions, individuals could be forced to do what is good, leaving only "the banalities and trivia of human life," which are neither good nor bad, as matters of free decision (ibid., 83). Without a right to do wrong, one would be free to decide what flavor of ice cream to eat, or whether to avoid stepping on cracks in the sidewalk, but only because these are questions of indifference. Actions reflecting one or another conception of the good

could, in principle, be made compulsory, so that they conform to a paternalistic authority's conception of the good.

In his reply, however, Galston (1983) points out that Waldron's defense of a right to do wrong is not really an *argument* for this right; it is instead an explication of the (deontological) liberal theory of rights, according to which important personal decisions should not be compulsory. It is true that if one accepts this theory, then a right to do wrong follows as a matter of course. But what is missing from Waldron's paper is any reason to think that the liberal theory of rights "should be taken seriously" to begin with (*ibid.*, 323) in light of the paradox from which Waldron begins: the paradox of asserting that there is a right to do wrong. This paradox, in other words, exposes the core deontological liberal conviction as incoherent.

In turn, Waldron's rejoinder to Galston suggests that

on the liberal view, it makes a difference how the conception of the good has been adopted. If an individual's conduct has been forced or cajoled into conformity with a conception of the good, then coercion or deception has been driven to the heart of that individual's self-constitution as a moral being; his life has no meaning except a coerced or deceived meaning. It makes no difference to this worry whether the conception of the good in question is a worthy or an unworthy one: the point is that the victim of the enforcement of any moral ideal has had coercion driven into the heart of his moral life. (Waldron 1983, 326)

Waldron thus attempts to shore up his explication of the liberal defense of rights with an explication of the liberal defense of freedom of choice (albeit with "meaning" substituted for Kant's "dignity").

Waldron is not arguing (here) that paternalism is undesirable because it might be based on a mistaken conception of the good or because it might produce bad downstream consequences (as a consequentialist liberal might argue).² Instead, he is saying that the ability to make choices free of coercion or cajolery, including the freedom to make bad (i.e., "wrong") choices, is intrinsically valuable, regardless of the consequences of the actions that result. Moreover, he implies, this freedom must be the *highest* intrinsic value, since it outweighs the negative value of any and all wrong choices that it allows, as well as all the possible bad consequences of those choices. If there is indeed a "right" to do wrong, as Waldron (1981, 80) insists, then the intrinsic

value of the freedom to act without being “interfered with” must be the ultimate value.

That, of course, is exactly how deontological liberalism treats freedom of action. Hence the priority of “the right” over “the good.”

Towards a Logic of Choice

While Waldron (1983, 85) does show that “a right to do wrong . . . is actually required by the way rights function in [deontological liberal] moral theory and the basis on which we [deontological liberals] argue for them,” he concedes that “even after these arguments have been given, the sense of paradox is likely to linger.” He then pinpoints the problem by referring to the logic of choice. “In real life,” he points out (ibid., 87), the

assertion of a right to do A is most often uttered by someone who intends to do A and is responding to moves by other people to prevent his carrying out that intention. There is something odd about having the intention to do A and sincerely believing that A is wrong.

Put more strongly, Waldron’s point is that one would not do something if one didn’t think it were the right (or “good”) thing to do—as opposed to the wrong (or “bad”) one.

That is the starting point for “the logic of choice.” The logic of choice restates Socrates’ idea that we cannot freely do what we take to be wrong. However, the logic of choice does not deny that people can knowingly do what is wrong. They may be forced (by other people or by their passions) to do wrong. In that case, however, they are not acting freely. The logic of choice applies only to the extent that one is exercising free will.

In explicating this logic, I will discuss five hypothetical situations that share three premises: that the agent wants to do A rather than the other possible actions, {B, C, D, . . . n}; that none of the possible actions, {A, B, C, . . . n}, infringe on others’ freedom of action; and that A is objectively the best action in the set {A, B, C, . . . n}.

Situation 1. Action A is allowed; all other actions {B, C, D, . . . n} are prohibited by law.

Situation 2. A is prohibited by law; all other actions {B, C, D, . . . n} are allowed.

Situation 3. A is mandatory; all other actions are forbidden by law.

Situation 4. The entire set of actions consistent with others' freedom of action—{A, B, C, . . . n}—is allowed by law.

Situation 5. B is legally mandated; A and all other actions are forbidden. (This situation will be discussed later.)

In Situation 1, the coercion embodied in the law against {B . . . n} does not affect the logic of the agent's choice because the agent chooses to do A anyway. Since she has no intention of doing the things that are illegal, the laws against {B . . . n} do not deprive her of any valuable option, so the coercion being directed against those options is of no importance to her.

In the opposite situation (Situation 2), where A is prohibited, legal coercion *is* important to her, since she would have chosen to do A. However, the problem with prohibiting A stems from the putative goodness of A, not from the badness of coercion *per se*. The evil of the coercion is, from the perspective of the agent, parasitic on the goodness of the action it prevents (A). Conversely, the huge gain in freedom represented by the agent's newfound ability to do {B, . . . n} is utterly without value to her.

In contrast, in Situation 3, the agent's freedom of action is altogether eliminated. Since she perceives the goodness of action A, however, this drastic reduction in her freedom makes no difference at all. From the agent's perspective, only the perception that A is *not* the best choice in comparison to, say, B would make Situation 3 worse than Situation 2, even though Situation 2 allowed her far more freedom of action. Let us call this Situation 3'.

As in Situation 3, A is mandatory in Situation 3', but unlike in Situation 3, the agent thinks that B instead of A is the best choice. This perception would lead her to value the right to do what is, by our assumption, wrong—B—and to rue the contraction of her freedom of action in comparison to Situation 2. But this is only because she incorrectly perceives what is wrong, or bad (B), to be right, or good. Thus, if she were to become aware of her mistake, returning herself to Situation 3, the much greater coercive scope of the law in Situation 3 as compared to Situation 2 would be of no consequence to her; indeed, as in Situation 1, she would not be coerced at all.

In Situation 4 the agent lives in a society in which freedom of action is treated as if it were intrinsically valuable. Here she has the greatest conceivable degree of freedom compatible with the same freedom for everyone else. Yet since she continues to perceive A as the best thing to do, the radical expansion of her freedom over Situations 1 and 3 is superfluous, because she will not exercise this freedom, i.e., her right to do wrong.³

These examples ground Waldron's paradox in the logic of choice. For they show that the presupposition of choice is that normative distinctions can be made among actions, such that one choice emerges as "better" than the others: this is, in everyday language, the "right" (or "good," or "best") choice. (For expository purposes, I will run together the "better" and the "best" choices in a given situation.)

Since actions distinguished as wrong (or "bad," or "worse") *should not* be done, the logic of choice requires that a "right to do wrong" would never be exercised and thus would never be valued by a choosing agent—unless, of course, she had mistakenly confused the wrong with the right. Even then, however, the freedom of action protected by the right to do wrong is, for the agent, purely instrumental to the putative good: It is merely a means to the end of doing something that she (inaccurately) perceives to be good. Since the agent never does what she *thinks* (at the time) is the "wrong" action, she has no use for a right to take such an action. In noticing the paradox, Waldron recognizes the oddness of actually attempting to make use of such a right. Since the logic of choice dictates that all free actions stem from an agent's perception of what (according to one or another principle of action) she *should* do, no agent would knowingly avail herself of her right to do wrong, since this would entail doing what she (thinks that) she should not do. Therefore she would not value the freedom to act upon this right.

Thus, Waldron's paradox echoes the actual situation facing a freely choosing agent. But his attempt to defuse the paradox by defending the intrinsic value of free choice shows only that this alleged value is as paradoxical as is the right to do wrong.

This is true even if one's choice is more important than which flavor of ice cream to eat. Consider something that may or may not be illegal, depending on the locality: cruelty to animals. If I think it is wrong to be cruel to animals, then I will choose not to be cruel to animals regardless of the law. Suppose that cruelty to animals is illegal in my town but I am

unaware of this fact, and that one day I refrain from being cruel to an animal that crosses my path. It is hard to fathom how the law plays any role in the meaningfulness of my action. If my action is adjudged meaningful, but the next day I discover that, unbeknownst to me, it would have been illegal to treat the animal cruelly, is the meaningfulness of my action now diminished? Similarly, suppose that I think it my duty to fight in a war, but that before I can enlist, conscription is imposed. Is fighting in the war any less meaningful?

In reply, one might suggest that the question of meaningfulness more clearly arises if the law *prohibits* what one thinks would be the best action. But it is not evident that a law requiring me to be cruel to animals, thus prohibiting me from refraining from cruelty, would make my judgment that it is best to refrain from cruelty any more meaningful than if the law allowed me to act on this judgment. The judgment is the same.

The Intrinsic Worthlessness of Freedom

In defending the intrinsic value of freedom, it seems to me that deontological liberals have mistaken a fact for a value. Freedom of action is a fact (for an agent with free will) before she has chosen among options. Prior to her choice, the agent is at liberty to do both what is good and what is bad. But this freedom must be suspended when the choice is made, and superseded when acting upon the choice. Each time an agent deliberately acts on the basis of a free choice, she constrains her freedom of action by means of an antecedently perceived criterion of “the good” (however implicit this criterion might be), treating her freedom of both choice and action as instrumentally valuable to this good.

To choose and to act, the agent needs to supplant her freedom with knowledge of a criterion that will allow her to identify the best action among those she is free to choose. In turn, after the choice has been made, only the freedom to take the chosen option is valuable—not because it was chosen, but because it is perceived as the best choice. Freedom presents to the chooser more than one alternative that must somehow be winnowed down to a single one. The chooser uses her (putative) knowledge of which action is best to *eliminate* freedom of choice (for the moment) by casting aside all the other options. Since only

one action can be taken at a time, she also casts aside freedom of action (for the moment), constraining herself so that she may actually act.

That is why, in cases where the putatively desirable action is legally compulsory, the absence of the freedom to take other actions is irrelevant, from the perspective of the logic of choice: One *needs* constraint (by means of knowledge) if one is to choose and then act. The presence of *legal* constraint changes nothing, and in cases where the needed knowledge is missing (e.g., Situation 3'), legal constraint can accomplish the same thing that constraint by knowledge otherwise would accomplish. In contrast, when the putatively best action is forbidden by law, greater freedom of action would indeed be valuable—but only as a means to the end of doing the best thing. Finally, in cases where one is free to take the putatively right action as well as other (“wrong”) options, the freedom to take the latter options is not just worthless; its only potential value is negative, since its only potential use is to do the wrong thing.

The flip side of recognizing the paradox in asserting a right to do wrong, then, is recognizing that freedom of choice and action are, at best, sometimes instrumentally valuable; that they are in other instances useless; and that in all the cases where they are neither instrumentally valuable nor useless, they are worse than useless, since they may permit bad action. In no case, however, does freedom of action possess intrinsic value. An agent who asserted the intrinsic value of freedom would contradict this assertion every time she acted.⁴

Knowledge vs. Power

By stipulating our hypothetical agent’s perception that she should do A, we saw that she would logically find none of the legal constraints important except the ones imposed in Situation 2 and Situation 3'. This would also be the case in Situation 5, which we have not yet considered. In this case, a relatively bad choice (B) is mandatory. In all three of these cases, however, the problem (from the agent’s point of view) is not a lack of freedom or an excess of coercion; the problem is that the coercion is directed against what she perceives to be the best choice, A, which I have also stipulated is actually the best choice. In Situations 2 and 5, then, the remedy would be morally knowledgeable legislators, not necessarily liberal ones. In Situation 3', the agent, rather than the law, is mistaken.

What is lacking is again knowledge, not freedom, as the agent recognizes once she is (somehow) convinced of her mistake.

These points may seem platitudinous. Of course we need moral knowledge—but getting it is easier said than done, and stipulating it does not create it. That is true enough, but the very real cognitive problem of choice is not solved by allocating to individuals, to legislators, or to anyone else the *authority* to choose. All choosing authorities will face the same need for knowledge; all acting agents will need to know the (putative) truth about which ends they should pursue.

It is wrong, therefore, to see deontological liberalism as somehow more congruent with the scarcity of moral knowledge, and the fallibility of moral agents, than consequentialist or even authoritarian alternatives are. Human agents are fallible, but they still need what they consider to be knowledge if they are to act. The logic of choice simply spells out the implications of this need, and this logic is incompatible with deontological liberalism.

At the end of the paper, I shall return to the impression that deontological liberalism deals better with human fallibility than alternative positions do. For now, I hope that the reader will consider the possibility that in defending a right to do wrong, deontological liberals have come to spurn knowledge (as a constraint on action) in favor of power (in the sense of absence of constraint). The explanation probably lies in the polemical demands of the case for religious toleration during and after the Wars of Religion. The horror of these wars produced an understandable desire for freedom of conscience and of religious practice—as a means to the end of civil peace.⁵ Yet modern liberals came to treat freedom of conscience, of observance, and of action as ends in themselves, regardless of its beneficial consequences. It is always easier to make an argument based on moral principle, such as the principle of freedom of conscience, freedom of choice, or freedom of action, than an argument based on what Rawls (1971, 4) calls the “calculus of social interests”—shifting, tenuous, and contingent as it must be.

II. VEILING THE GOOD

One might object that in the foregoing analysis, I conjured into being the indifference of the agent to freedom through my narrow construal of the agent’s concerns. In each situation, I imputed to the agent the

intention only to pick the best available option. Given this instrumentalist intention, it is no surprise that freedom was valuable to this agent only instrumentally, if at all. But what would happen if, instead, we imputed liberal sentiments to the agent, so that she *believed* in the intrinsic value of the right to act upon bad (wrong) choices? Then, *ex hypothesi*, in each situation the agent would have valued the freedom to take actions that were not instrumental to her conception of the good.

Such an objection, however, is too simple. In essence, Waldron's paradox stems precisely from combining, in the perspective of one agent, the putative desirability of taking a particular action with the putative value of the freedom to take other actions.

This can be seen even though Waldron (1981, 73) states the paradox from the perspective of an observer: If one finds someone else's action to be "arbitrary or capricious," he writes, it "is not made to appear one iota more reasonable or defensible" if the agent justifies the action on the grounds that "I had a right to do it." By way of example, Waldron adduces the example of an agent who marries someone she loathes. Let me expand on this example, since there may be many reasons for such a marriage that make sense from a logic-of-choice perspective. From the agent's perspective, such reasons would be instrumentalist (although not necessarily consequentialist): "I married him for his money," or "Life is a drama and this union will produce a great tragedy," or "My parents promised me to his parents, so by marrying him I fulfilled an obligation." Waldron's point is that in contrast to such reasons, which would make sense from the agent's perspective, the *freedom* to marry him, as expressed in the *right* to marry him, is no reason at all. If, instead of a reason like those just listed, the agent were to say, "I married him because I had the right to," the appropriate response would be, "Perhaps so, but you also had the right to marry someone else or nobody." The paradox is *created*, not dissolved, by inserting into the agent's instrumentalist, action-oriented perspective the deontological-liberal, freedom-oriented perspective. It is from this mixed perspective that the contradiction between deontological liberal sentiments and the logic of choice emerges.

It would seem, then, that any defense of the right to do wrong must somehow reconcile the agent's orientation toward action with the intrinsic value of freedom. Thus, Rawls devoted increasing attention to such a reconciliation in the years immediately following *A Theory of Justice*, beginning with "Justice as Fairness" (1975) and continuing in "Kantian Constructivism in Moral Theory" (1980). Later, Rawls

dropped this approach in favor of the “political, not metaphysical” turn he took in 1985, culminating in *Political Liberalism* (1993).

In *A Theory of Justice* (1971), Rawls structured the basic institutions of society around protecting, equalizing, and maximizing the “fundamental interest in liberty and in the means to make fair use of it” (ibid., 563). Rawls presented his defense of liberty as a direct competitor with utilitarian theories of justice, which might sacrifice people’s freedom in the interest of beneficial consequences. Instead, Rawls treated freedom of thought and action as the highest intrinsic values. “Systems of ends,” he maintained, cannot be “ranked in value” by contractors in the original position (ibid., 19); if they were, then violations of equal freedom might be enacted by contractors who wanted the highest-ranked system of ends to be implemented in the basic structure of society. Rawls (ibid., 254) adduced the example of hedonism: If “the parties were conceived . . . as pursuing certain kinds of pleasures, then the principles chosen would apply . . . only to persons whose freedom was restricted to choices compatible with . . . hedonism.” Instead, the principles of justice ensure that outside the original position, one’s “liberty in adopting a conception of the good is limited only by principles that are deduced from a doctrine which imposes no prior constraints on these conceptions” (ibid.). For if one’s “ends were restricted in some specific way, this would appear at the outset as an arbitrary restriction on freedom” (ibid.).

Yet when Rawls first tries to reconcile this freedom with the logic of choice, what is striking is the arbitrariness of the actions chosen, and the pointlessness of protecting the freedom to take them.

Veiling Heteronomous Conceptions of the Good

Ignorant of the ends they pursue outside the original position, the contractors know only “that in general they must try to protect their liberties, widen their opportunities, and enlarge their means for promoting their aims whatever these are” (Rawls 1971, 143). Hence the importance of securing a fair distribution of “primary goods,” the “all-purpose means” to any ends that do not conflict with the principles of justice (Rawls 1980, 547). “Whatever one’s system of ends, primary goods are necessary means” (Rawls 1971, 93). Not just the difference principle, however, but also the principles of greatest equal liberty and of

fair equality of opportunity assure everyone “an equal liberty to pursue whatever plan of life he pleases” (ibid., 94). All of the principles of justice, then, are designed to preserve or expand people’s freedom, which Rawls treats as an end in itself: the right that subordinates the good.

Only in a highly abstract sense does Rawls allow the contractors even to think about the good. That they do think about it in the abstract is necessary, however. Otherwise, they would have no reason to care about protecting their liberties, widening their opportunities, and enlarging their means for promoting their aims (e.g., Rawls 1971, 95, 144, 147, 396). That is, they would not care about the freedom that is protected by the principles of justice. For this reason, Rawls needed to introduce the agent’s perspective into the original position, and a constrained reconciliation of the logic of choice and the deontology of liberalism became a subtext of Rawls’s project from the beginning.

Indeed, this reconciliation was so important that Rawls misconstrued his project as part of the theory of rational choice—a construal he withdrew only in *Political Liberalism* (Rawls 1993, 53n7). By that point, he seems to have concluded that his attempts at reconciliation had failed, for *Political Liberalism* replaces such attempts with the idea that “the values of the political are very great values and are not easily overridden” (ibid., 139). “These values express to [sic] the liberal political ideal that since political power is the coercive power of free and equal citizens as a corporate body, this power should be exercised . . . only in ways that all citizens can reasonably be expected to endorse in the light of their common human reason” (ibid., 139–40). Leaving aside the circularity of defending liberalism by invoking the allegedly very great value of liberalism, Rawls has shifted the burden of proof onto the shoulders of those who would “expect others to accept . . . a conception of justice that favors those” who “affirm a particular religious, philosophical, or moral comprehensive doctrine with its associated conception of the good” (ibid., 24). Before, Rawls had tried to give such people reasons to subordinate their conceptions of the good to the right. Now such people must prove that their conceptions of the good *outweigh* the allegedly great values of liberalism. The “burdens of judgment,” which make it difficult to determine what is really good, ensure that this proof cannot be offered. (Therefore a reasonable pluralism of ends is a “fact.”)

In the earlier writings we shall consider, however, Rawls accepted that the burden was on him to show contractors why they should

subtract from the logic of choice its most essential precondition, the cognitive element, thereby veiling the (putative) good. This was a question Rawls addressed more than once.

In “Fairness to Goodness,” he asked “whether by excluding information about people’s conceptions of the good[,] the original position rules out morally relevant information” (Rawls 1975, 268). Rawls’s initial answer to this question, however, not only justifies veiling the good within the original position, but destroys the coherence of the good outside it. Conceptions of the good should be veiled, Rawls argues, because they are as heteronomously determined, and thus are as morally arbitrary, as is one’s luck in the genetic lottery.

“Our final ends,” Rawls (1975, 269) writes,

depend on our abilities and opportunities, on the numerous contingencies that have shaped our attachments and affections. That we have one conception of the good rather than another is not relevant from a moral standpoint. In acquiring it we are influenced by the same sort of contingencies that lead us to rule out a knowledge of our sex and class.

Similarly, in *A Theory of Justice*, Rawls (1971, 252) had asserted that to act in view of “the specific things that” one “happens to want” is “to act heteronomously.” The veil of ignorance, then, merely “deprives persons in the original position of the knowledge that would enable them to choose heteronomous principles.” By cordoning ourselves off from knowledge of these principles in the original position, we achieve autonomy. Autonomy of this sort, however, is not the *aim* of the principles of justice; rather, it is their rationale. The aim is to protect the freedom of action of real people in the real world. This freedom is justified by the fact that autonomous people in the original position—unlike real people—would veil the good, and thus would allow their real-world counterparts the freedom to pursue whatever heteronomously determined conceptions of the good they happened to prefer.

Offhand, it may appear that in veiling conceptions of the good because of their heteronomous determination, Rawls was conflating what philosophers of science call the “context of discovery” with the “context of justification.” A perfectly sound scientific claim may originate in alchemy or a dream or an inexplicable inspiration; the origin of the claim, however, has no bearing on whether it is justified or not. By the same token, in evaluating the accuracy of one’s conception

of the good, its *source* is irrelevant, regardless of whether the source is heteronomous or not; it would be a gross error in reasoning to think otherwise.

Rawls, however, is unlikely to have been making this mistake, for in *A Theory of Justice* and “Fairness to Goodness,” he does not seem to have recognized the possibility that a conception of the good *can* be a truth claim. If he had, he would have noticed that conceptions of the good are qualitatively different from the other heteronomous factors excluded from the original position by the veil of ignorance. When the contractors ignore their knowledge of their gender, class position, and so on, they are excluding knowledge of what is (empirically), not (putative) knowledge of what ought to be. To veil their conceptions of the good, however, the contractors must treat these oughts, too, as if they were simply empirical facts akin to the historical and genetic forces that are irrelevant to justice. Indeed, in the final analysis, Rawls (1971, 552) maintains, “choices” about which ends to pursue are made simply by letting oneself be determined by one’s inclinations, pure and simple. At that crucial point, the logic of choice becomes useless: “Using the principles of rational choice as guidelines, and formulating our desires in the most lucid form we can, we may narrow the scope of purely preferential choice, but we cannot eliminate it altogether.”

Purely preferential “choices” are not properly regarded as truth claims about anything except the way we are—truth claims, at best, about our empirical desires, but not about what we *should* desire. Such truth claims, however, do not serve the purpose of making a choice. Unless one uses the tacit principle that one should do whatever one most strongly desires, one cannot motivate a choice among options merely by tallying what one’s inclinations *are*, since this empirical fact begs the question of what one *should* be inclined to choose. Moreover, the tacit principle that one should act on the basis of the strength of one’s desires cannot itself be justified except as a truth claim that appeals beyond one’s inclinations.⁶

Sandel’s Challenge

Freedom of choice is both cognitive and inherently normative: If it has any relationship with the facts of inclination or extant desire, the relationship is to prompt the question of their normative relevance to the choice at hand. Thus, Rawls’s initial argument for veiling the good

goes too far. If our conceptions of the good are heteronomous, and if this justifies veiling them, then it must be due to the fact that these conceptions are not normative truth claims, since such claims would be evaluated on the basis of their truth or justifiability, not on the basis of the heteronomous context of their discovery. But if they are not normative truth claims, then they cannot serve to motivate choices, and it is unclear what purpose they *could* serve.

Perhaps more pertinently, it is unclear why we should structure society around the unfettered pursuit of heteronomously determined goods. As Sandel (1982, 168) put it, “Once it is conceded that our conceptions of the good are morally arbitrary, it becomes difficult to see why the highest of all (social) virtues should be the one”—justice as fairness—“that enables us to pursue these arbitrary conceptions ‘as fully as circumstances permit.’” Rawls needed more than a reason to veil people’s conceptions of the good; he also needed a reason for people to treasure them. (If this seems paradoxical, it is the same paradox Waldron and Galston analyzed.)

Rawls addressed this challenge more squarely in “Kantian Constructivism” than in any other work. This happened because Rawls (1980, 541) came to recognize the cognitive dimension of conceptions of the good, pointing out that “some religious and philosophical doctrines must be true, even if they merely deny other false or incoherent doctrines.” How, then (he asked), can veiling our conceptions of the good be justified to us “as individuals” who “are convinced” that our moral beliefs “are true” (ibid., 542)? Unless this justification is provided, then we have no reason to subordinate these putative truths to the right by veiling the good.

Rawls’s question nicely poses the problem with veiling the good, especially inasmuch as he situates the problem in our reasoning as agents acting from antecedent conceptions of *true* goods. In answer to his question, however, Rawls turns from conceptions of the good to conceptions of the person. He begins with the fact that people may think of themselves as “inevitably tied to the pursuit of the particular conception of the good and its final ends which they espouse at any given time” (Rawls 1980, 544). Thus, they may “have attachments and loves that they believe they would not, or could not, stand apart from; and they might regard it as unthinkable for them to view themselves without certain religious and philosophical convictions and commitments” (ibid., 545). In short, they may view themselves as if they were

heteronomously determined. However, Rawls now maintains that “none of this need affect the conception of the person connected with society’s public conception of justice and its ideal of social cooperation.” From a political perspective, people are to regard themselves not as heteronomous, but as being able to “stand apart from conceptions of the good and survey and assess their final ends” (ibid., 544). In the revised version of *A Theory of Justice*, Rawls puts it plainly in a passage that did not appear in the first edition: The parties in the original position “do not think of themselves as inevitably bound to, or as identical with, the pursuit of any particular complex of fundamental interests that they may have at any given time. . . . Rather, free persons conceive of themselves as beings who can revise and alter their final ends and who give first priority to preserving their liberty in these matters” (Rawls 1999, 131–32).

It would seem, then, that Rawls is saying that our freedom to revise and alter our ends trumps whatever heteronomy characterizes them originally, at the point when they are first delivered to us by family or culture. We can decline to follow these heteronomously given ends; that is, we can exercise free will in regard to them. This is in sharp contrast to the notion that our ends are inherent to us and inseparable from us, such that in place of free choice we must exercise purely preferential choice.

Self-Originating Sources of Valid Claims

Thus far, “Kantian Constructivism” aligns with the logic of choice—but this logic would justify paternalist and consequentialist violations of “the right” to do wrong if the revision of our ends entails claims about their falseness prior to being revised.

Consider that the agent considered in Part I is just as capable of “revising and changing” her “particular conception[s] of the good” (Rawls 1980, 544) as is the agent Rawls envisions in “Kantian Constructivism.” However, the agent in Part I *exercises* her free will by scrutinizing the *accuracy* of her perceptions of a good criterion of choice. We saw this happen when the agent was persuaded that her criterion of choice had been mistaken, such that she then changed her mind, moving from Situation 3’ to Situation 3. If Rawls’s agents were engaged in this type of scrutiny of ends when they revised them, the outcome of the process would be a truth claim about their revised ends and a falseness

claim about their previous, unrevised ends. The agents who revised their ends would be saying, in effect, that previously they had erred. In now taking actions that contradicted their previous errors, they would be saying that it is appropriate to constrain their freedom to take erroneous actions, for that is exactly what they are doing by taking different actions that reflect their revised conception of the good. In principle, then, they should have no compunctions against imposing on others (or trying to impose on others) the ends they have now concluded are truly good, thereby constraining others' freedom of action as they now constrain their own. Once we are making a truth claim about the good, nothing but prudence stands between us and Bishop Bossuet, who famously said, "I have the right to persecute you because I am right and you are wrong."⁷

Without explanation, however, Rawls (1980, 543) avoids this problem—the problem with which he began when he recognized that people think that their conceptions of the good are true, such that they require reasons not to use political power to implement these conceptions—by concluding that people who have (or think of themselves as having) free will are, *ipso facto*, people who are (or think of themselves as being) "self-originating sources of valid claims." Notably, this is not Rawls's jaundiced description of what is in fact going on when we "perceive" what we take to be an antecedent good. That is, he is not saying that such "perceptions" are our own creations, which we unwittingly take to antecede our wills. Rather, Rawls is offering agents a reason to veil their putatively *true* conceptions of the good: Agents are to recognize that they (and thus their fellow citizens), as agents with free will, are *not* perceivers of antecedently valid claims but are, instead, creators of the claims' validity. Presumably Rawls introduces this inconsistency because he understands free will to be, by definition, undetermined—even by antecedent conceptions of the good.

A complication arises at this crucial point in Rawls's argument because he vacillates, often from one sentence to the next, between suggesting, on the one hand, that people really are self-originating sources of valid claims because they actually have free will; and suggesting, on the other, that they are merely to think of themselves this way. The latter, weaker reading seems to be consistent with the "constructivist" thrust of "Kantian Constructivism." However, my goal is less to explicate Rawls than to treat him as representative of deontological liberalism as a whole, so we need not reach a conclusion

about this peculiarity of his theory as of 1980. What is important is that whether free will is supposed to be real or not, Rawls treats it (or the construction of it) as a reason for considering people's conceptions of the good as self-validating, and thus as a reason for protecting their freedom to pursue these conceptions by veiling them (in the original position). Our free will, then, serves to justify the political protection of freedom of action. This is a widely shared deontological view that is (arguably) true to the egalitarian and humanistic intention of liberalism: the intention to treat all human beings as equally dignified and as deserving of equal respect.⁸ Nevertheless, it remains inconsistent with the logic of choice, and therefore is not a legitimate deduction from either the existence or the construction of our free will.

As we have seen, an agent who is thought to possess free will cannot motivate a free choice without an implicit truth claim about the desirability of the choice. If the agent considered in Part I had viewed herself as a self-validating source of valid claims, then *any* end she perceived as good—indeed, any action she chose to take—would have become, ipso facto, good. But truth is necessarily perceived as independent from and prior to choice. Indeed, the lesson of “the logic of choice” could be summarized in this way: Free will, if it is to be exercised as free action, entails an antecedent normative truth claim to eliminate all but one of the alternative actions one might take. The conception of the person that Rawls articulates in “Kantian Constructivism,” then—that of a person with free will—is incompatible with the person as a self-validating source of claims. One whose free will is a self-validating source of claims cannot possibly perceive such claims as valid antecedent to her act of will.

Thus, Rawls (1980, 548) emphasizes that “the freedom of persons as self-originating sources of claims is represented by *not* requiring the parties to justify the claims they wish to make.” That is, these claims are to be considered valid because of their source in free agents' choices, not because of the accuracy of these agents' perceptions of the good: the refusal to demand reasons for claims “is how freedom as originating valid claims is represented” (ibid.). Rawls wants us to think of ourselves as having free will, and to recognize that our conceptions of the good are truth claims, but without recognizing the function of truth claims in constraining free will such that action is possible. For if the “moral power” to stand back from, “critically assess,” and revise our conceptions of the good were exercised on the basis of a truth claim representing

“knowledge of an independent moral order,” as Rawls (1993, 97) later put it, one would be implying that one’s previous conception of the good had been *wrong* in violating that order, and as we have seen, this could justify paternalism, in principle.

It is telling that in “Kantian Constructivism,” the only example that Rawls (1980, 544) can adduce of standing back from our conceptions of the good is when they “conflict with the principles of justice, for in that case, they must be revised.” Here the criterion of revision is not, for example, the discovery a new conception of the good or a new argument for a different conception. The criterion of revision is, instead, the right as a trump against any conception of the good that violates it. It is as if the right has not only taken priority over the good, but has become the supreme good. However, when it comes to making free choices that *are* congruent with the principles of justice, such that the right can no longer play a role in deciding the choice, we are back to Sandel’s challenge. For if truth claims are not the basis of these choices, only whim (“pure preference”) can be. Thus, even in “Kantian Constructivism” Rawls did not banish the shadow of heteronomous conceptions of the good.

In his later writings, too, Rawls mixes acknowledgements of the cognitivist status of conceptions of the good with noncognitivist accounts of what it means to hold or revise such conceptions. Thus, in *Political Liberalism*, Rawls (1993, 313) allows that freedom of action is “the liberty to fall into error and make mistakes.” Yet he also writes of citizens “revising and changing” their conceptions of the good “if they so desire” (*ibid.*, 30). The process sounds so noncognitivist as to be mysterious:

Our conceptions of the good may and often do change over time, usually slowly but sometimes rather suddenly. When these changes are sudden, we are likely to say that we are no longer the same person. . . . We refer to a profound and pervasive shift, or reversal, in our . . . identity. (*Ibid.*, 31; cf. Rawls 2001, 22–23)

This passage warrants scrutiny because it embraces the very position that, in “Kantian Constructivism,” Rawls had used as a contrast against his own. The contrasting position was that taken by people who think of themselves as “inevitably tied to the pursuit of the particular conception of the good and its final ends which they espouse at any given time” (Rawls 1980, 544; cf. Rawls 2001, 22). The person who was, by contrast, able to stand back from and revise her ends becomes, in *Political Liberalism*

(and *Justice as Fairness*), a different person—one with a different identity—by virtue of whatever mysterious event led to the revision. Yet if a sudden change in ends thereby changes the person's identity, there is no longer anyone left who stands back from, assesses, and revises ends. Not only does free will disappear, but so does any coherent conception of the person who is supposed to have made a “mistake” at Time 1 before revising her conception of the good at Time 2. Similarly, in “Fairness to Goodness,” Rawls (1975, 283, *emph. added*) had described the principles of justice as appropriate for agents who “are not identified with their actual or possible plans but are viewed rather as beings that have a capacity for forming, adopting, and changing these plans, *should they be so moved*; and who give priority to preserving their liberty in these matters.” In “Kantian Constructivism,” having added a cognitivist account of conceptions of the good (as truth claims), Rawls effectively retracts it by depicting the process of “being moved” to revise and alter one's conception of the good—which had previously been used to describe heteronomous changes of mind—as if it somehow manifests freedom of the will.

III. VOLUNTARISM, LIBERAL AND THEOLOGICAL

A cognitive understanding of choice, which Rawls arrived at in “Kantian Constructivism,” goes with a certain conception of the person. This person is a human being trying to figure out what to do. She may not be aware of it, but in the process she is searching for reasons that will constrain her freedom of choice and action. However, if she is aware of this, then she cannot make claims on the basis of having chosen them. She makes claims based on what she perceives as good reasons for her choices, however cognizant she may be of past mistakes and however aware of the possibility that, in the future, she will look back on her present reasons for action as mistaken.

In contrast, the Rawlsian person claims equal liberties and primary goods not on the basis of “antecedent principles . . . to which” she is “bound” (Rawls 1980, 548), but on the basis of her “moral personality,” i.e., her ability to stand clear of antecedent principles by pursuing, revising, and altering a conception of the good. Recognizing the same moral personality in her fellow citizens, she grants them the ability to pursue, revise, and alter conceptions that they freely will.⁹ This free

willing, as Rawls understands freedom, cannot be determined either by heteronomous genetic or cultural forces *or* by reasons purportedly drawn from an independent moral order. The Rawlsian person's "choices," then, are as inexplicable as they are free. So, too, are her actions—uncaused causes, freely willed and thus liberated from antecedent determination by a perception of the good.

We might say that with the spontaneity of an uncaused cause, the Rawlsian person creates the good *de novo*. But such a person *cannot* "fall into error and make mistakes," since that would presuppose an antecedent principle of truth against which her actions—and those of her fellow citizens—might, in principle, be judged.¹⁰ Yet an uncaused source of good whose actions can be neither explained nor criticized is not a conception of a person at all. It is a conception of God, as God was understood by voluntarist theologians.

Theological Voluntarism

Theological voluntarism was a product of the crisis in thirteenth-century theology precipitated by Latin Averroism.

The Latin Averroists were philosophers at the University of Paris who, in the 1260s and 1270s, seemed, at least to their critics, to abandon the Scholastic attempt to reconcile the newly translated texts of Aristotle with the dictates of Christian orthodoxy. Aristotle's understanding of man and nature had confronted Christian theology with a series of potentially fatal problems. Among these was a threat to the freedom of God's actions. Francis Oakley (1984, 80) notes that while Aquinas was able "to regard the whole of being . . . as in its own fashion subject to the dictates of the same [eternal] law," one consequence of this view was "that subjection to law could well be seen to extend to God himself, thus threatening his freedom and omnipotence." Aristotelianized Christianity, in short, had the potential of binding God to antecedent criteria of the good, threatening His "right" to do as He pleased. The Latin Averroists pushed this threat openly, and the pope reacted through the agency of the Bishop of Paris, who in the year 1277 issued a proclamation condemning 219 "obvious and loathsome errors" derived from "the accursed pagans" (quoted in Hyman and Walsh 1973, 584).

The lasting effect of the condemnations of 1277 was to turn Christian theology against the notion that reason could elucidate matters of faith.

Within a few decades of 1277, the greatest late-medieval theologian, William of Ockham, had canonized this semi-irrationalist *via moderna* by sharply distinguishing between theology and philosophy. Advocates of the *via moderna* held “that to describe God in terms of a first cause was to regulate His actions; that to attribute to Him a constant mode of operation impaired His absolute sovereignty. Accordingly, emphasis was coming to be placed upon the infinite freedom of His will which defied analysis or explanation” (Leff 1958, 258). This tendency was expressed in theological voluntarism.

Against the Aristotelian view of the universe as a stable order with dependable, epistemically accessible natural laws—laws that bind even God—theological voluntarists endorsed the ongoing possibility that God, should He “be so moved,” might at any moment recreate the world in such a way that its laws, including its moral laws, would be entirely different from those now in place. To restore the autonomy of God, the voluntarists thought that nothing inherent in particular divine actions or their consequences could be considered independently and antecedently good; rather, they are good solely by virtue of their inscrutable divine creation. Voluntarism was thus a crucial part of the late-medieval effort to build an unbreachable wall between what reason can know and what must be taken on faith, an effort that culminated in the Reformation.

Voluntarism expresses one pole of a fundamental contradiction in Christianity: the conflict between God’s omnipotence and His absolute goodness. Long before Christ, of course, Plato (1941, 41) had articulated the basic tension in the *Euthyphro*: “Is what is holy holy because the gods approve it,” Socrates asks, “or do they approve it because it is holy?” If, as Socrates argued, the quality of holiness, or goodness, is independent of divine approval, then the gods would seem to have no choice but to will the good, which suggests a constraint on their power. Transposed to the God of Christianity, this poses a doctrinal problem, for if God is omnipotent, no limits on His power would seem to be acceptable.

That was the conclusion theologians reached after 1277. The Bishop of Paris had effectively foreclosed the Socratic view, which would have subordinated God’s omnipotence to His goodness by ruling out the possibility that He could will something independently defined as wrong. The right of God to do as He pleases must, it was now thought, contradict any constraint on His autonomy, and thus any God-independent moral order.

Thus, the first voluntarist theologian, Duns Scotus, “began the slide to ethical relativism” by arguing that “something is good because God as good wills it rather than it is willed by him because it is good, and since he can will whatever is not self-contradictory, the only thing he cannot will is hate of himself as the highest good” (Leff 1976, 54). This constraint on God’s will is parallel to revising one’s conception of the good when it conflicts with a principle of justice. It sets a very broad limit on the agent’s freedom, but cannot acknowledge, let alone explain, constraints (and thus actions) within that limit. Ockham, however, Scotus’s most brilliant follower and critic, while continuing to accept that God could not will what is self-contradictory, went farther, denying that the constraint of logic rules out even God’s hatred of Himself:

The hatred of God, theft, adultery, and actions similar to these according to the common law . . . can be performed by God without any evil condition annexed; and they can even be performed meritoriously by an earthly pilgrim if they should come under a divine precept, just as now the opposite of these in fact fall under a divine command. (Ockham 1317–18, 56–57)

This difference between theological and liberal voluntarism can be explained as follows. In liberal voluntarism, there is a large number of “self-validating sources of claims,” i.e., individual sovereigns, each of whose freedom of action is to be protected; in theological voluntarism, there is only one. The constraints set in liberalism by “the right,” or “justice,” or “the political” are attempts to ensure that each sovereign’s authority is respected. Theological voluntarists do not have to deal with this problem, so contrary to Scotus, there need not be a constraint on God that parallels the right—Rawls’s sole constraint on conceptions of the good.

Theological voluntarism was an understanding of normativity premised on the intrinsic desirability of freedom of action for the normative sovereign. If God’s freedom of action is intrinsically desirable, and thus somehow expresses God’s dignity, then it disrespects God and derogates His sovereignty to suppose that He is “bound” by “antecedent principles.” Therefore, such principles—normative criteria—must be superfluous to God’s actions, which become unconstrained by anything (except, perhaps, the law of non-contradiction). God’s right to act as He is moved displaces any conception of the good that might constrain His action.

Leibniz's Critique of Voluntarism

The definitive answer to theological voluntarism did not appear until 1709, with the publication of Leibniz's *Theodicy*. The Leibniz of the *Theodicy* is, of course, the model for Voltaire's Dr. Pangloss, who was so deftly satirized in *Candide*. But for our purposes, what is interesting about Leibniz is *why* he was driven to "gloss over all" the evils of the world.

Leibniz's claim that ours is the best of all possible worlds was a direct response to theological voluntarists' denial that we could judge God's work as good according to an independent criterion of goodness. A consequence was skepticism about the ultimate goodness of Creation—reviving the theodicy problem: How could an omnipotent and good God create a world of manifest imperfection, suffering, and evil? One way to solve the problem is to deny that the evils of the world are real. This was Leibniz's Panglossian alternative. In the course of defending it, however, he produced a logic of choice that explains the incompatibility of freely willed action with the intrinsic value of freedom.

"If there were not the best (optimum) among all possible worlds," Leibniz (1709, 128, sec. 8) wrote, "God would not have produced any." Tied together in this brief statement are both Leibniz's Panglossianism and his unique meta-ethics. These can easily be separated, rendering the latter profoundly relevant to deontological liberalism.

Let us dispose of the Panglossianism immediately. God seeks the good, Leibniz contends; and it is a postulate of faith that, being infallible, He finds what He seeks. "A lesser good is a kind of evil if it stands in the way of a greater good; and there would be something to correct in the actions of God if it were possible to do better" (Leibniz 1709, 128, sec. 8). Since, *ex hypothesi*, God cannot need correcting, the world He has created must be the best world possible. Leibniz is committed by this argument to explain away all the enormities of earthly existence. This is what Voltaire found so odious, but it follows only if one combines Leibniz's meta-ethics with the premise that the hypothesized Christian God, omniscient and omnipotent, exists. Substitute for God any fallible, free-willed agent and Leibniz provides a perspective on the nature of choice that undermines all forms of voluntarism, not just the theological one.

By virtue of His omniscience, Leibniz (1709, 323, sec. 327) points out, God "always wills what is most to be desired." The theological dimension of this assertion is the assumption that a God exists who

infallibly knows what is desirable (e.g., *ibid.*, 428, sec. 21). Bracketing the epistemic issue, however, why might we believe that God, *if* He is omniscient, always *wills* what is desirable? Leibniz's answer applies to any agent with free will. "Taking it in the general sense," he explains, "one may say that will consists in the inclination to do something in proportion to the good it contains" (*ibid.*, 136, sec. 22). Leibniz's reasoning is simple: By the principle of sufficient reason, everything has a cause (*ibid.*, 147, sec. 44; 407, sec. 3). Thus, "every will supposes some reason for willing, and . . . this reason is naturally anterior to the will" (Leibniz 1686, 5). This reason is the desire to do whatever is "best" according to *some* criterion of the good: Such a criterion must be present to motivate any deliberate action. This is the logic of choice.

Leibniz proves that this logic is inescapable by contrasting it against alternative explanations for an action. The first alternative, which Rawls apparently endorses in "Kantian Constructivism," is that one might simply "will" an action. But this is no explanation at all. "We do not will to will; else we could still say that we will to have the will to will, and that would go on to infinity" (Leibniz 1709, 151, sec. 51). When discussing the human "will," one might be inclined to translate this infinite regress into the endless chain of physical causes that embodies the operation of the principle of sufficient reason under the laws of nature. This would mean treating the will as heteronomously determined. In the case of God, however, that possibility is barred by theological doctrine. So there must be another cause that determines God's will—and our own wills, to the extent that they are not determined but free.

Leibniz (1709, 236, sec. 174; 345, sec. 367) calls this other cause "moral determination," as distinct from natural or physical determination. Moral determination is determination by an antecedent conception of the good (*ibid.*, 313, sec. 310):

Modern philosophers have improved upon the opinions of the Schoolmen by showing that, according to the laws of corporeal nature, a body can only be set in motion by the movement of another body propelling it. Even so we must believe that our souls (by virtue of the laws of spiritual nature) can only be moved by some reason of good or evil. (*Ibid.*, 407, sec. 3)

Thus, Leibniz satisfies the principle of sufficient reason by establishing that a free will must be "moved"—that is, determined—by an antecedent criterion of goodness, rather than by the types of forces

that affect, say, the position of a planet (which I will call “mechanical” determination, including all forms of heteronomy).¹¹ Mechanical determination leaves no space for freedom of the will, while “moral” (normative) determination does: When one is morally determined, one is still mechanically capable of taking alternative actions, but one judges them to be inferior to the chosen action. The antecedent criterion of the good that is the cause of one’s choice “determines” one’s action without mechanically necessitating it.

The voluntarists’ God, in contrast, lacking moral determination, must, of necessity, be determined mechanically. Just as Rawls seeks to preserve a right to freedom of action by initially denying freedom of the will, the voluntarist theologians, in attempting to liberate God from antecedent criteria of the good, inadvertently deprived God of a free will. Conversely, then, if we are to preserve God’s free will, i.e., His capacity to stand back from, assess, and choose ends without mechanical determination, we have no option but to deny His right to unconstrained freedom of action. In turn, this entails denying His sovereignty—i.e., His status as a self-originating and self-validating source of claims. For if God is not mechanically determined, then the source of His commands is one and the same as the reason for their validity: their antecedent goodness, the criteria of which winnow down to one the infinite number of actions that he might take. God’s moral determination thus entails that His commands themselves are not the source of their validity, which antecedes them.

In the case of Creation, for example, God could have chosen “another sequence of things” than the present order of our universe (Leibniz 1709, 148, sec. 45; cf. Adams 1994, 33). His decision to create this universe instead of another one, therefore, was not mechanically determined. But precisely because God’s will was free of such strict determination, it had to be (by the principle of sufficient reason) subject to normative determination instead: that is, determination “only by the preponderating goodness of the object” (ibid, 148, sec. 45). In order “to preserve the privilege of free will” both of God and of ourselves, then, Leibniz concludes that it is enough that “we have been so placed at a fork in the road of life, that we can do only what we will, and can will only what we believe to be good” (quoted, Adams 1994, 11). With us, unlike God, the belief may be false. “But free will tends toward the good, and if it meets with evil it is by accident, for the reason that this evil is concealed beneath the good and masked, as it were” (Leibniz 1709, 220, sec. 154).

Leibniz is not suggesting that freely chosen actions are intrinsically valuable, such that the heteronomous determination of these actions would be intrinsically wrong. It was the theological voluntarists who treated God's actions as inviolate—producing the paradox that the actions must in fact either be heteronomous or inexplicable. Instead of valorizing freely made choices, by contrast, Leibniz simply undertakes an epistemological investigation of them, asking what might have “moved” God to do one thing rather than another.

The Incompatibility of Voluntarism with the Logic of Choice

In reply to Leibniz, one might say that determination is determination, and that if he is going to insist on the principle of sufficient reason, he cannot get out of it by conjuring up a new species of “determination,” normative or “moral” determination, that is consistent with free will. Either God's actions, and our own, are uncaused causes, free of all determination, or they are not. If they are not, as Leibniz is saying, then they are not free.

Voluntarists, however, face the same dilemma. Leibniz is, in effect, noting that their horn of the dilemma—the free will as inexplicably determined—is much worse than the horn he is defending. If God's will is neither mechanically nor morally determined, then, by the principle of sufficient reason, it is unintelligible, because it is an uncaused cause. An uncaused cause is not just a mystery; it is a contradiction in terms. So if voluntarists want to insist on treating God's will as sovereign, hence as moved neither by mechanical causes *nor* by antecedent principles, they render logically impossible any actions such a will might take (Leibniz 1709, 149, sec. 48). Better to make a distinction between normative and mechanical determination that allows us to explain God's actions coherently—even at the expense of denying God's sovereignty.

Grasping Leibniz's horn of the dilemma, we can explain (in the abstract) how God chose to do one thing rather than another: to create this universe, to issue the Decalogue rather than the anti-Decalogue, and so on. He did it by discovering what he (infallibly) perceived to be antecedent criteria of the good. In the voluntarist understanding of God's will, by contrast, where His action is self-originating and self-validating, it is itself the criterion of goodness, depriving God, prior to the action, of the grounds of choice among possibilities. *If every divine action is*

self-validating because the divine will is sovereign, God would be unable to choose which action to take.

This is why Leibniz (1709, 128, sec. 8) maintained that “if there were not the best (optimum) among all possible worlds, God would not have produced any.” Before God’s decision to create the universe, if there had not been a reason (its optimality) for choosing one world over another, God would not have been capable of deciding to create any worlds at all. Assuming that God’s will is free, then if not from a desire to match His action to this independent criterion of goodness, how else would God have been “so moved” as to choose one possible world over the others? Lacking an antecedent criterion of the good, He would have been unable to constrain His freedom so He could act.

Inaction, then, is the only other alternative—apart from mechanical determination—to normative determination.¹² Leibniz (1709, 148, sec. 46) calls inaction based on the lack of a perceived criterion of the good “an indifference of equipoise.” An agent stuck in this familiar position would be like the legendary Buridan’s Ass, which starved because it was unable to choose among what seemed to it identical bales of hay. “This equipoise in all directions,” Leibniz declares, is “impossible,” for no two options can be entirely equal (*ibid.*). Substantiating this claim requires Leibniz to engage in implausible speculation about the shape and consistency of the universe, such that no two bales of hay (or fields of hay, as Buridan’s dilemma is often posed) could be equal.

It might have been better for Leibniz to have conceded that equipoise is quite possible. We often experience indecision, whether because the options before us are actually equal or because they appear to be. However, this fact supports rather than undermining Leibniz’s account of free actions, which cannot take place in the absence of a *perceived* advantage for one option over the others. When we lack such a perception, we find it impossible to choose. Whenever there is action, there must first have been either a choice based on the putative goodness of the action or its consequences; or there must have been mechanical determination.

This perspective renders value pluralism, or what Rawls (1980, 542) calls “impartiality” among ends—the foundational commitment of deontological liberalism—untenable for free, acting agents. In Rawlsian liberalism, value pluralism takes the form of veiling the good. In the original position, veiling the good is merely meant to protect freedom of action. However, impartiality among all ends would also be the effect if

an agent outside the original position took her freedom of action to be a self-originating source of self-validating claims. This would effectively make the agent like the voluntarist God, unconstrained by any antecedent norms, such that, “until the will has determined itself, there would be no reason for its determination, either in him who chooses or in the object; and one would not choose what pleases, but in choosing without reason one would cause what one chooses to be pleasing” (Leibniz 1709, 406, sec. 1). Even if one believed that by choosing one bale of hay from among the others, it would thereby be transformed into the best bale of hay, one would be unable to motivate this choice if one genuinely believed that the bales were equally valuable before the choice, as a value pluralist believes that several incommensurable values cannot be ranked.

Alternatively, if one *could* choose without ranked antecedent perceptions of the good, one’s “choices” would (by the principle of sufficient reason) have to be determined by mechanical causes, not free will. Yet it was, I hypothesized, to escape the untenable position of having originally grounded the supreme value of freedom of action in its heteronomy that Rawls came to emphasize free will in “Kantian Constructivism,” in the form of our “moral power” to stand back from heteronomous conceptions of the good. Leibniz’s argument allows us to see clearly the lack of any basis for moving forward (toward action) from this autonomous standing-back unless one is normatively determined to do so by an antecedent reason, rather than by some mysterious process of being “moved.” For Rawls (1971, 254), any antecedent reason would constitute an “arbitrary restriction on freedom.” But without such a reason, freely willed action would be impossible. “The mere desire to make use of one’s freedom has no effect of specifying, or determining us to the choice of one course or the other” (Leibniz 1709, 311, sec. 305).

The Asymmetry between Deontology and Consequentialism

One might argue that none of these points are relevant to the *political* sovereignty of the individual. Perhaps the individual’s actions are determined by a perceived antecedent criterion of the good; perhaps not. Perhaps such criteria are claims about moral truth that are, as such, sound or not independent of any authoritative agent. But why should any of this matter to political theorists? Axiology and theories of action

need not be relevant to politics; indeed, the genius of liberalism is to separate politics from such contentious issues.

Thus, one might continue, Waldron erred by inquiring into agents' reasons for their actions. Only this inquiry made his paradox seem to be more than a linguistic oddity, opening a Pandora's box of questions about agency and axiology that have nothing to do with political theory proper. Perhaps, as well, Rawls was wrong to try to merge the logic of choice with deontology. A stricter deontologist might simply have ignored whatever determines private agents' choices, the better to seal off activities in the private realm from inspection, criticism, and coercion. This is, in a sense, what the "political, not metaphysical" Rawls finally did.¹³

In other words, one might respond to my argument by insisting on the traditional liberal dichotomies: political versus metaphysical, public versus private, right versus good. The "conception of the person," including the question of free will, would be consigned to the private realm along with "the nature of the good," and the logic of the choice of action would be hermetically sealed from the logic of the choice of principles of social order; the personally normative would have nothing to do with the politically obligatory. A truly "political, not metaphysical" liberal would not offer the freely choosing individual a rationale for treating freedom as intrinsically valuable; she would just treat it that way, period, by asserting the "very great value" of liberal freedoms. A bloody-minded liberal deontologist might do well, therefore, to follow the example of theological voluntarists, who acknowledged that God's ways are mysterious and left it at that (ultimately producing Protestantism). Not only should we not try to combine the deontological and "logical" perspectives, but we should cordon off the latter in a separate branch of philosophy and insist that it has no implications for justice.

This response to my argument, however, creates a new problem for deontological liberalism: Even given the traditional liberal dichotomies, and therefore the priority of political justice (whatever its content) over private conceptions of the good, why should anyone endorse "justice" *defined as freedom of action*? Why should we endorse the right to do wrong even if we do not consider it paradoxical, or even if we are willing to bracket the paradox? In short, why should we endorse the priority of *the right* even if we accept that its logic is distinct from that of the good?

In principle, the consequentialist liberal does not face a similar problem, as she can point to goods (such as civil peace) that may be

achieved by liberal justice. Still, if the deontologist prioritizes the right over the good, the consequentialist merely reverses these priorities, so it may seem that we have reached a standoff. However, the consequentialist does not in fact prioritize the good over the right. She subsumes the right within the good. She proceeds from an independently perceived (non-voluntarist) good to the collective means (if any) of achieving it, just as she proceeds from this perception of the good to individual action when she makes a “personal” decision. The logic of consequentialist political choice is *the same* as the logic of “private” choice.¹⁴ Therefore, for the consequentialist liberal, there is no conundrum paralleling the question that besets liberalism: Why should we endorse the right over the good?

Thus, the “metaphysical, not political” Rawls acknowledged again and again that the burden was on him to show why we should *deviate* from the logic of choice when we think about politics. In other words, Rawls recognized that the deontological liberal is at a disadvantage when compared to the consequentialist. This is why he repeatedly tried to justify veiling the good to the agent who is (of necessity) motivated by the good. His answers, however, were contradicted by the implications of free will as expressed in the logic of choice. If the liberal then falls back on traditional liberal dichotomies in order to discount the relevance of the logic of choice, we are entitled to ask why we should take these dichotomies seriously—as Galston asked. Until we get an answer, deontological liberals are begging the question.

IV. THE COMMUNITARIAN DEAD END

There are many similarities between Sandel’s critique of deontological liberalism and mine, starting with the fact that Sandel identified deontological liberalism with voluntarism¹⁵ and voluntarism with inexplicable, arbitrary “choices” among ends (Sandel 1982, 180).

Oddly, however, like Alasdair MacIntyre (1981), Charles Taylor (1992), and Michael Walzer (1984), Sandel asserted that communally constituted ends are somehow less arbitrary than those constituted by individuals. Liberal theorists eagerly showed otherwise. For example, in response to MacIntyre’s apotheosis of communally defined virtues, Stephen Holmes (1993, 114; cf. Shapiro 1990, 146) asked whether “we want to call ‘virtues’ those traits that enable a Mafia hit-man to act

as a Mafia hit-man is supposed to act.” Nancy Rosenblum (1989, 54; cf. Gutmann 1985, 319) pointed out that a community can easily impose relationships “that may be experienced as suffocating rather than embracing.” Holmes (1993, 105, 110) drew the logical conclusion from this indeterminacy, equating communitarianism with “sociological relativism” or “cultural relativism”—even while noting that “relativism is not what MacIntyre has in mind” (*ibid.*, 110).

Sandel eventually affirmed that relativism was not what he had had in mind, either. In *Liberalism and the Limits of Justice*, he had equated the good with what is identified as good by one’s community. In reaction to Rawls’s emphasis on the logic of choice in “Kantian Constructivism,” Sandel (1982, 179) had argued that “we cannot regard ourselves as independent” enough of our ends to be able to form and revise them “without great cost to those loyalties and convictions whose moral force consists partly in the fact that living by them is inseparable from understanding ourselves as the particular persons we are—as members of this family or community or nation or people, as bearers of this history, as sons and daughters of that revolution, as citizens of this republic.” In other words, he had argued that communally constituted identities are *self-originating sources of valid claims* (for the members of given communities). But sixteen years later, in the preface to the second edition of his book, Sandel (1998, xi) repudiated communitarianism: “The mere fact that certain practices are sanctioned by the traditions of a particular community is not enough to make them just.” This is exactly what the liberal critics had been saying, but it contradicted Sandel’s criticism of Rawls for subordinating the good to the right—i.e., to justice: Now Sandel was doing the very same thing.¹⁶

While Sandel eventually felt compelled to resubordinate the good to the right, however, he did not withdraw his claim that communally constituted goods were somehow an “indispensable aspect of our moral experience” (Sandel 1982, 179).¹⁷ One might have expected that to prove this claim, he would have produced a logic of choice showing the necessity of communal attachments. Instead, he simply asserted that personal identity—“understanding ourselves as the particular persons we are”—entails “enduring attachments and commitments” to “this family or community or nation or people” (*ibid.*).

Plainly this is untrue. A great many of us feel few or no commitments to such groups, nor do we define ourselves “as bearers of this history, as sons and daughters of that revolution, as citizens of this republic” (*ibid.*).

And even those of us who do define ourselves in this way may feel that other obligations, unconnected to any group or personal history, trump these identities. Moreover, at the normative level, as the liberal critics pointed out and as Sandel eventually conceded, a blanket grant of authority to one's felt communal attachments would license any variety of injustice, depending on which community one felt attached to (a slave-owning community? a Nazi community?). It would also license any and all contradictory ends, as long as each was upheld by at least one community.

Communitarian Epistemology

Instead of providing a communitarian logic of choice, Sandel simply embraced the understanding of the good that Rawls had contrasted against the liberal view of the self in "Kantian Constructivism." That is, Sandel treated as a logical necessity some people's tendency to "have attachments and loves that they believe they would not, or could not, stand apart from," and to "regard it as unthinkable for them to view themselves without certain religious and philosophical convictions and commitments" (Rawls 1980, 544). In addition, he contended that constitutive communal attachments provide deeper, more normatively compelling grounds for action than are available to a self that is distanced from its ends. "To imagine a person incapable of constitutive attachments such as" those to family, community, nation, or people, he wrote, "is not to conceive an ideally free and rational agent, but to imagine a person wholly without character, without moral depth" (Sandel 1982, 179).

Sandel's two points are not, strictly speaking, compatible with each other. If communal attachments are literally indispensable, then the purely preferential commitments that Sandel (1982, 163, 180) criticizes as shallow and arbitrary cannot exist. I take it to be obvious that communal attachments are not literally indispensable, so let us focus on his claim that decisions based on such attachments are less arbitrary than are purely preferential choices. This claim depends on the peculiar sense in which he believes that communitarianism is cognitivist or epistemological, while deontological liberalism is voluntaristic.

"For the subject whose identity is constituted by ends already before it," according to Sandel (1982, 152–53), "agency consists less in

summoning the will than in seeking self-understanding. Unlike the capacity for choice, which enables the self to reach beyond itself, the capacity for reflection enables the self to turn its lights inward upon itself, to inquire into its constituent nature, to survey its various attachments and acknowledge their respective claims.” This “cognitive” possibility, Sandel maintains, is denied to the Rawlsian agent, for the question “‘What ends shall I choose?’” is

addressed to the will. Rawls’ subject would thus appear epistemologically impoverished where the self is concerned, conceptually ill-equipped to engage in the sort of self-reflection capable of going beyond an attention to its preferences and desires to contemplate, and so to re-describe, the subject that contains them. (Ibid., 153)

For Sandel, then, knowledge of the good is equated with knowledge of the self. “Epistemology” investigates the empirical question: What ends already “constitute” us? The model for such investigation is when I, “uncertain which path to take,” consult “a friend who knows me well, and together we deliberate, offering and assessing by turns competing descriptions of the person I am” (ibid., 181).

If I were engaged in the logic of choice, I would ask what I *should* do. In Sandel’s moral epistemology, however, the category of the normative is effaced by the allegedly deeper inquiry into my predetermined (if perhaps ambiguous or conflicting) commitments. An epistemology of heteronomy is thus Sandel’s “cognitive” alternative to Rawls. It seems appropriate, then, to take Sandel’s eloquent criticism of the arbitrariness of purely preferential (heteronomous) choice and apply it to Sandel’s model of choice guided by personal identity.

From Voluntarism to Heteronomy

Sandel (1982, 158, quoting Rawls 1971, 552, 551) points out that in Rawls’s view, “we must simply choose. ‘We may narrow the scope of purely preferential choice,’” according to Rawls, “‘but we cannot eliminate it altogether. . . . We eventually reach a point where we just have to decide which plan we most prefer without further guidance from principle.’” To which Sandel (ibid., 162) trenchantly replies that

a choice that is a choice “in the sense that” it “often rests on” (is determined by?) my existing wants and desires is a choice only in a peculiar sense of the word. For assuming with Rawls that the wants and desires on which my choice “rests” are not themselves chosen but are the products of circumstance (“We do not choose now what to desire now” [Rawls 1971, 415]), such a “choice” would involve less a voluntary act than a factual accounting of what these wants and desires really are. And once I succeed in ascertaining, by “direct self-knowledge,” this piece of psychological information, there would be nothing left for me to *choose*. . . .

Such a “decision” decides nothing except how accurately the agent has perceived something already *there*, in this case the shape and intensity of his pre-existing desires.

However, the same thing could be said of Sandel’s account of agency, as long as we remove the word *choice* and substitute “identity” for “desires.” Sandel (1982, 180) is well aware of this, and closes his book by addressing the matter. In contrast to purely preferential choices, he writes, “When I act out of more or less enduring qualities of character,”

my choice of ends is not arbitrary in the same way. In consulting my preferences, I have not only to weigh their intensity but also to assess their suitability to the person I (already) am. I ask, as I deliberate, not only what I really want but who I really am, and this last question takes me beyond an attention to my desires alone to reflect on my identity itself. . . . Although there may be a certain ultimate contingency in my having wound up the person I am—only theology can say for sure—it makes a moral difference none the less that, being the person I am, I affirm these ends rather than those, turn this way rather than that. While the notion of constitutive attachments may at first seem an obstacle to agency—the self, now encumbered, is no longer strictly prior—some relative fixity of character appears essential to prevent the lapse into arbitrariness which the deontological self is unable to avoid.¹⁸

Thus, the depth and profundity that Sandel attribute to communally constituted agency merely come down to its stability over time. If values are inculcated by nurture or culture, as they may be by one’s family, religious community, or nation-state, then they are likely to last longer—perhaps for a lifetime—than if they are arrived at through solitary contemplation. One *is less likely to change one’s mind* if one is so

steeped in the culture of one's upbringing that one cannot even conceive of "standing back and assessing" the ends taken for granted in that culture. So one will be more predictable.

However, the persistence of an end has nothing to do with whether it is sound, and the predictability of one's behavior has nothing to do with whether one's character is admirable. Sandel rescues only the appearance of non-arbitrariness and confuses depth of character with obliviousness to alternatives—that is, obliviousness to what might be called "the fact of choice." Indeed, Sandel (1982, 164) equates choice itself with an irrational, capricious will; he is therefore in no position to offer a logic of choice.

Sandel sees, in effect, that a criterionless Rawlsian chooser is in the position of Buridan's Ass, but he does not see that this implicates the ideal of freedom by pitting it against the logical prerequisites of freely willed action. Nor does he see that there is any alternative to criterionless choice except mechanical determination. Therefore he is pushed into portraying the ends inculcated heteronomously by our families, communities, and societies as indispensable to personal identity. This closes off the genuine epistemological concerns apotheosized in Rawls's acknowledgement that people see their "religious and philosophical doctrines" as "true" (Rawls 1980, 541). If one sees a doctrine as true, then one's normative "identity" must conform to it—not the other way around.

Granted, a doctrinal truth may *seem* to be so self-evident that one cannot imagine its being false. But one does not believe it to be true (or rather, one does not think that one believes it to be true) simply by virtue of its originating in one's communally constituted identity or even one's individually constituted identity. There is a difference between thinking that A is so obviously good that it is inconceivable that {B, C, D, . . . n} are better, on the one hand; and, on the other, being so comfortable with having always thought that A is good that one literally cannot imagine believing otherwise. In the former case, which is a more recognizably "cognitive account" (Sandel 1982, 153) of agency than is communitarianism, one does not retreat from recognizing one's freedom of choice, and thus from acknowledging that one believes that one's criterion of choice is sound. In the communitarian version of cognitivism, however, the subject displays such a deep failure of imagination that she cannot envision herself believing not-A *even if not-A were demonstrated to her to be true*. That is, her beliefs are somehow

compelling not because they seem to be sound, but because they seem to be (contingently and arbitrarily) “hers.”

Sandel (1998, 187) does recognize the “voluntarist” dimension of liberalism: the fact that it embraces as the ultimate good a criterionlessness form of action that can only be capricious and heteronomous. However, Sandel misconstrues the source of the problem as liberal individualism, and thus offers, as the alternative to the “radical” choice facing a voluntarist individual (Sandel 1982, 164), the acceptance of “ends we have not chosen—ends given by nature or God, for example, or by our identity as a member of a family or people, culture or tradition” (ibid., 150). The “epistemological” dimension of such an acceptance is that to embrace their communal identities, people must first come to know “what they *are*, not a relationship they choose (as in a voluntary association) but an attachment they discover” (ibid.). But there is no substantive difference between this position and Rawls’s description of ends as heteronomous in *A Theory of Justice* and “Justice to Goodness,” excepting the locus (communal rather than individual) of heteronomy. Purely preferential choice, too, has an epistemic dimension: the discovery of what one truly desires. If there is an epistemic deficit in liberalism, then, it does not stem from inattention to the discovery of what *is*. The problem, instead, is inattention to the discovery of what *should* be.

The Narcissism of Small Differences?

That said, one might object that the Leibnizian “normative determination” I have defended is no more different from “purely preferential choice”—labels aside—than is communitarianism. After all, what Rawls calls our preferences and Sandel calls our constitutive attachments serve the same function as what I call our antecedent criteria of the good. Whatever they are called, they are ends that determine one’s choice, and they do so through what might be seen as the same epistemic process: whether the criterion of (say) happiness or beauty is based on a perception of “what should be” or instead on an “existing” preference or an attachment, one has to perceive it, or recognize it (at least tacitly), before one can act upon it. Moreover—despite Sandel’s attempt to

blame arbitrariness in the “choice” of ends on the distancing of self from ends—there is in fact, in all three views, a distinction between knower and known, or observer and observed; even if one is trying to “kno[w] oneself” (Sandel 1982, 181), there has to be an object (in this case, the self) whose constituent parts or implications an observer-subject is trying to perceive accurately. This is also true of the fourth position discussed here, theological voluntarism.

Given these similarities, the distinctions between the alternatives may seem negligible, in the grand scheme of things. For in the logic of choice, as in the two non-theological versions of voluntarism, the causes of action are only putatively “sound”; what one perceives to be one’s communal identity or one’s overriding preference or the antecedent good may be an illusion. Thus, all three non-theological approaches—including my own, which purports to substitute normative for mechanical determination—are potentially heteronomous in the objectionable sense of conflating justified normative standards with the contingent products of faulty perceptions. (This is not true of theological voluntarism only because it is stipulated that God is infallible.)

There are two elements of truth in this objection, I think, but in both cases the effect is ultimately to buttress my position instead of undermining it.

The first element of truth is that *if* an agent is capable of deliberate action, then the agent—whether the agent posited by Ockham, by Leibniz, by Rawls, or by Sandel—must be separate from, and able to perceive, the values that guide its action. However, this similarity shows that the logic of choice (unlike, say, the recognition of communal attachments) truly is “indispensable” in an accurate account of agency. For the voluntarist theorist, the agent is a self-originating source of valid claims. But for the voluntarist agent, self-validation would make action impossible, so in her own mind, she has to be a seeker of truths that are “outside” herself. This breach between the agent’s point of view and that of the theorist mirrors the divergence between agent and theorist that is evident in the deontological-liberal insistence on the agent’s right to do wrong.

The other element of truth in the first objection, however, leads to a second objection.

Fallibilism and Its Simulacra

In all four views except theological voluntarism, the agent's perception of the truth may be faulty. Yet non-theological voluntarisms may seem to be more cognizant of this fact and more accepting of it than is "the logic of choice," since the voluntarists would not impose the mistakes made by one individual or community on another.¹⁹ Perhaps one can say that the voluntarist views are more genuinely "epistemological" in the sense that they take to heart the fundamental problem of epistemology, the possibility that what we "know" may not be true. Toward the end of *Political Liberalism*, as we have seen, Rawls (1993, 313) even invokes the liberty to "fall into error and to make mistakes."

However, it does not follow from human fallibility that individuals should be left free to pursue whatever ends they perceive to be true, since no given individual is (a priori) less fallible than those who might try to impose their perceptions on us all.²⁰ The *logical* source of voluntarist tolerationism is its designation of the individual or the community as authoritative or sovereign. When there are multiple sovereigns, their authority is protected by policing their borders. (For this reason, in theological voluntarism, where there is only one sovereign agent, there is no illusion of toleration.) Where one designates a normative authority, however, one effectively treats it as infallible. Hence, as Waldron demonstrated, the very concept of "wrong," or "should not," begins to lose its meaning in liberal usage. The idea of "error" cannot make sense for a self-validating source of claims.

By contrast, the agent who determines her action by means of a criterion that she takes to be genuinely good need not insist that this perception is infallible. We must act on the basis of implicit truth claims even if we are aware that they *may* be false. On the other hand, if our confidence in the accuracy of our axiological perceptions is great enough to justify constraining our own freedom of choice, then we must also think that it justifies constraining others' freedom of choice (barring other considerations, such as relevant differences between our situation and theirs). Confidence in an axiological truth claim works in the other direction, too. If we did not have enough confidence in such a claim to impose it on others, we could not impose it on ourselves and, like Buridan's Ass, would be unable to act.

Deliberate action, then, presupposes that one thinks oneself more likely than not to have identified an accurate criterion of good action.

But this does not preclude acknowledging that one may be wrong. An agent pursuing the logic of choice is therefore open, in principle, to the possibility that her perception of the good may be mistaken, as may any other perception of part of the “external” world. Indeed, in the logic of choice, it is the distance between the perceiving subject and the putatively perceived object that creates space in which error may occur. Where an agent’s identity and its ends are inseparable, one’s recognition of an error about ends would entail the extinction of one’s self.

Despite the fact that “the” logic of choice opens up *theoretical* space for a recognition of error, however, one may still wonder whether the epistemic element of this logic would not lead to endless and potentially violent political conflicts about the good among those who see their opponents as horribly mistaken. Taken alone, the logic of choice would certainly lead in that direction. However, the conflict between the logic of choice and liberal voluntarism leads naturally to consequentialism as the alternative. If human happiness or the minimization of human suffering is taken to be the good, a strong case can be made for liberal principles of justice that would treat people as if they had a right to do wrong. The rationale for these principles, however, would be that they are instrumental to the good—not that they are oblivious to it.

NOTES

1. This characterization applies to *A Theory of Justice* and Rawls’s later works through 1980. However, in *Political Liberalism*, Rawls (1993, 24) defines as “unreasonable” the desire to expect others to accept principles of justice that affirm one’s “particular religious, philosophical, or moral comprehensive doctrine with its associated conception of the good.” While the explicit rationale for this definition is that “one of our considered convictions” is the unreasonability of such a desire (*ibid.*, 8, 24, 97, 247), the implicit rationale is that unreasonability, as defined, would lead to endless and perhaps violent civil strife (*ibid.*, xxii, xxiv–xxvi, 151, 159, 243, 301, 329). The first, explicit rationale is a liberal version of communitarianism (see Part IV below); the second, implicit rationale is tacitly consequentialist and, if explicitly stated, would open the door—in principle—to restraints on freedom in the name of civil peace or the larger good consequence served by civil peace.
2. Waldron (1981, 74) defends the “right” on consequentialist grounds elsewhere in his 1981 paper.
3. Thomas Hurka (1987, 376) disputes this line of argument by suggesting that we would always prefer to have more rather than fewer options: “To have ten options rather than just the best among them is to be able to say no rather than

yes. It is to be able to say no nine times, and to be responsible for the fact that no was said.” This is to assume, however, that we already accept the intrinsic value of the freedom to say no (which is, of course, also the freedom to say yes), and that is precisely what is at issue. In addition to begging the question, Hurka’s argument also reveals, I think, the juvenile conception of normative psychology that unwittingly underwrites deontological liberalism in its attempt to defend the dignity and respectworthiness of each individual. Only children relish the right to say no for its own sake. Moreover, genuine respect for other agents *per se* would recognize that they are acting from what they take to be an accurate conception of the good; that the premise of their choices is the intention to do good; that, therefore, if they seem to be mistaken about the good, their own premises of action suggest that they should, in principle, be corrected. Children, however, resent being corrected, regardless of whether they are right or wrong; and an adolescent view of “respect” is that one should gain others’ esteem for doing whatever one wants to do, such that there is no right or wrong applicable to “personal morality” (Rawls 1993, 15), and such that political attempts to correct mistaken views of the good are seen as being pushed by people who “simply . . . insist on *their own* beliefs,” rather than on what they consider to be true beliefs (*ibid.*, 61, *emph. added*). The alternative would be to view not one’s dignity or respectworthiness as being challenged by paternalism, but rather the truth of one’s convictions; and to recognize, as Weber did, that since there is no infinite regress of argumentation when it comes to values, which are in this respect like mathematical axioms, political disagreement about values is ultimately a matter of warring faiths—faiths, however, in competing perceptions of the truth (Weber 1922, 153; Weber 1918, 117).

These considerations do not apply to Waldron’s notion of “meaningfulness,” since the distinct idea here seems to be an element of thoughtfulness or deliberation. One must think harder if one must choose rather than being coerced. However, sustaining this view would require an account of the good life as involving thoughtfulness or deliberation, and such an account would render the “right” to do wrong merely instrumental to the requisite conception of the good life. Therefore I bracket it as inconsistent with the intrinsic value of freedom that is suggested by treating the so-called right to do wrong as real, rather than as a consequentialist “as-if” right.

4. One might think at first that some small measure of intrinsic value of freedom might be preserved in the face of a deliberate (freely willed) action, even if one grants that this intrinsic value is outweighed by the greater intrinsic values putatively achieved, or intended, by the action. But since all freely chosen action entails the constraint of one’s freedom of action, any leftover intrinsic value of freedom of action must be smaller than the intrinsic value of *any* freely chosen action, no matter how inconsequential. I cannot say if this preserves some quantum of intrinsic value for freedom of action; but for *practical* purposes, the intrinsic value of freedom of action is nil, at best.
5. Zagorin 2003 provides a survey of the confused and contradictory arguments for toleration that were advanced. Allen 1928 gives more respectful attention than

does Zagorin to the consequentialist concerns of the *politiques*, and the uncomfortable theological questions to which these concerns led.

6. Rawls (1971, 416) writes: "To be sure, there is one formal principle that seems to provide a general answer" to the problem of choice of ends. "This is the principle to adopt that plan which maximizes the expected net balance of satisfaction. . . . But this principle fails to provide us with an explicit procedure for making up our minds. It is clearly left to the agent himself to decide what it is that he most wants and to judge the comparative importance of his several ends." Thus, "we must finally choose for ourselves in the sense that the choice often rests on our direct self-knowledge not only of what things we want but also of how much we want them." One can see that by "choice" Rawls does not mean the voluntarist credo, "Whatever I choose becomes good by virtue of the choice." One's "several ends" are already given to the "agent" heteronomously. The voluntarist element is, as of 1971, incompletely developed as the idea that the agent should have untrammelled freedom of action. But this idea is unintegrated into the theory of choice because heteronomy, hence the *absence* of choice of any kind (voluntarist or not), is, at this point, Rawls's rationale for veiling the good.
7. Mendus 1989, 7, quoted in Rawls 1993, 61.
8. However, see n3 above.
9. Again, it is a non sequitur to conclude from the fact (or construction) that we possess free will that it is good and thus should be protected. But that is a separate point.
10. In *Political Liberalism*, from which the quotation is drawn (Rawls 1993, 313), Rawls had abandoned as "metaphysical" any account of "the person." But that does not solve the problem.
11. Leibniz also distinguishes a third form of determination that might apply to God, logical determination by "metaphysical" necessity. This is not discussed in the text because it is plain that no human agent can do what is logically impossible.
12. However, see n11 above.
13. However, Rawls managed this only by tacitly appealing to the consequentialist fear that without political liberalism, society would degenerate into an analog to the Wars of Religion (see n1 above). Otherwise he would have had to allow that there are many different rational methods of dealing with the fact of reasonable religious and philosophical pluralism—not just toleration and individual freedom but paternalism, inquisition, heresy hunting, and civil warfare. If we reject these methods, it must be because we judge their temporal consequences to outweigh whatever costs to our souls, or others', might be incurred by tolerating their errors. Since this type of weighing is consequentialist, it might lead to paternalism, so Rawls's explicit argument in *Political Liberalism* is instead a communitarian one that appeals to the given sentiments of a liberal community. "Political liberalism does not . . . claim to discover what is true by its own distinctive methods of reason apart from any tradition"; instead it "help[s] to put in order our considered convictions of justice," our "shared political understandings, as Walzer might say" (Rawls 1993, 44–45). The communitarian rationale for liberalism, however, overtly conflates what *is* (what we believe)

- with what ought to be: “standards are so formulated and arranged that they are freely recognized as ones we do, *or should*, accept” (ibid., 44, *emph. added*). On communitarianism as ultimately dependent on this conflation, see Friedman 2002.
14. The consequentialist/deontological dichotomy is incomplete, as one may view certain personal or political actions as inherently good (not good because they produce good consequences) without violating the logic of choice. What the logic of choice rules out are claims for the intrinsic value of the authority (the freedom) to decide whether to take actions.
 15. See Sandel 1998, 22, 58–59, 113, 119, 121, 122, 127, 152, 154, 161, 176–77, 179.
 16. In Friedman 2002, ch. 2, I show that Sandel had in fact proposed a particularist, communitarian defense of Rawlsian tenets of justice on the ground that such a defense was less vulnerable than were Rawls’s universalist, individualist assumptions to being seized upon by libertarians such as Nozick. Thus, Sandel’s communitarianism (like that of Taylor, Walzer, and MacIntyre) was supposed to serve as a stronger foundation for the ideal of equal *individual* freedom. Therefore, when in the preface to the second edition Sandel repudiated blanket grants of normative authority to “the community” because such grants might license *injustice*, his action was not so different in spirit from the arguments of the first edition: in both cases, his aim was to defend justice as defined by liberal theory. However, if the philosopher is able—as Sandel did in the preface to the second edition—to stand above communally constituted identities and sort them according to whether they are just or unjust (and, indeed, whether they are just or unjust according to the liberal criterion of equal freedom), then nothing remains of communitarianism as a critique of liberalism. What does remain is the idea, which spread widely among political theorists in subsequent decades, that communally constituted identity is the source of the good.
 17. Cf. Sandel 1982, 22, 50, 53–59, 132, 150, 153–54, 164, 171, 176–77, 181.
 18. Given the argument Sandel is making, it seems unlikely that by “theology” he means theological voluntarism rather than the doctrine of predestination.
 19. In *Political Liberalism*, Rawls is able to adduce tolerance for mistake as a fundamental reason for affirming the principles of justice. However, this is because while Rawls asks the question of veiling the good in *Political Liberalism*, he answers it from the perspective of the liberal theorist, not from the perspective of the acting agent.
 20. Nor does Rawls claim otherwise: He invokes “the burdens of judgment” only to show that would-be paternalists cannot meet the (liberal) test of “reasonableness” by which Rawls tries to shift the burden of proof onto their shoulders.

REFERENCES

- Adams, Robert Merrihew. 1994. *Leibniz: Determinist, Theist, Idealist*. New York: Oxford University Press.
- Allen, J. W. 1928. *A History of Political Thought in the Sixteenth Century*. Totowa, N.J.: Rowman and Littlefield.

- Freeman, Samuel, ed. 1999. *John Rawls: Collected Papers*. Cambridge, Mass.: Harvard University Press.
- Friedman, Jeffrey. 2002. "The Politics of Communitarianism and the Emptiness of Liberalism." Ph.D. diss., Yale University.
- Galston, William A. 1983. "On the Alleged Right to Do Wrong: A Response to Waldron." *Ethics* 93 (January): 320–24.
- Gerth, H. H., and C. Wright Mills. 1946. *From Max Weber: Essays in Sociology*. New York: Oxford University Press.
- Gutmann, Amy. 1985. "Communitarian Critics of Liberalism." *Philosophy and Public Affairs* 14: 308–22.
- Holmes, Stephen. 1993. *The Anatomy of Antiliberalism*. Chicago: University of Chicago Press.
- Hurka, Thomas. 1987. "Why Value Autonomy?" *Social Theory and Practice* 13: 361–82.
- Hyman, Arthur, and James J. Walsh. 1973. *Philosophy in the Middle Ages: The Christian, Islamic, and Jewish Traditions*, 2nd ed. Indianapolis: Hackett.
- Idziak, Janine Marie. 1979. *Divine Command Morality: Historical and Contemporary Readings*. New York: Edwin Mellen Press.
- Leff, Gordon. 1958. *Medieval Thought: St. Augustine to Ockham*. Chicago: Quadrangle Books.
- Leff, Gordon. 1976. *The Dissolution of the Medieval Outlook: An Essay on Intellectual and Spiritual Change in the Fourteenth Century*. New York: New York University Press.
- Leibniz, Gottfried Wilhelm. 1686 [1953]. *Discourse on Metaphysics*, trans. Peter G. Lucas and Leslie Grint. Manchester: Manchester University Press.
- Leibniz, Gottfried Wilhelm. 1709 [1951]. *Theodicy: Essays on the Goodness of God, the Freedom of Man, and the Origin of Evil*. London: Routledge and Kegan Paul.
- MacIntyre, Alasdair. 1981. *After Virtue*. Notre Dame: University of Notre Dame Press.
- Mendus, Susan. 1989. *Toleration and the Limits of Liberalism*. Atlantic Highlands, N.J.: Humanities Press.
- Oakley, Francis. 1984. *Omnipotence, Covenant, and Order: An Excursion in the History of Ideas from Abelard to Leibniz*. Ithaca: Cornell University Press.
- Ockham, William of. 1317–18 [1979]. *On the Four Books of the Sentences*. Excerpted in Idziak 1979.
- Plato. 1941. "Euthyphro." In *Plato on the Trial and Death of Socrates*, trans. Lane Cooper. Ithaca: Cornell University Press. Pagination from excerpt in Idziak 1979.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.
- Rawls, John. 1975. "Fairness to Goodness." In Freeman 1999.
- Rawls, John. 1980. "Kantian Constructivism in Moral Theory." *Journal of Philosophy* 77(9): 515–72.
- Rawls, John. 1985. "Justice as Fairness: Political Not Metaphysical." In Freeman 1999.
- Rawls, John. 1993. *Political Liberalism*. New York: Columbia University Press.

- Rawls, John. 1999. *A Theory of Justice*, rev. ed. Cambridge, Mass.: Harvard University Press.
- Rawls, John. 2001. *Justice as Fairness: A Restatement*. Cambridge, Mass.: Harvard University Press.
- Rosenblum, Nancy L., ed. 1989. *Liberalism and the Moral Life*. Cambridge, Mass.: Harvard University Press.
- Sandel, Michael. 1982. *Liberalism and the Limits of Justice*. Cambridge: Cambridge University Press.
- Sandel, Michael. 1998. *Liberalism and the Limits of Justice*, 2nd ed. Cambridge: Cambridge University Press.
- Shapiro, Ian. 1990. *Political Criticism*. Berkeley: University of California Press.
- Taylor, Charles. 1992. *Sources of the Self*. Cambridge: Cambridge University Press.
- Waldron, Jeremy. 1981. "A Right to Do Wrong." In idem, *Liberal Rights: Collected Papers 1981–1991*. Cambridge: Cambridge University Press.
- Waldron, Jeremy. 1983. "Galston on Rights." *Ethics* 93 (January): 279–325.
- Walzer, Michael. 1984. *Spheres of Justice: A Defense of Pluralism and Equality*. New York: Basic Books.
- Weber, Max. 1918. "Politics as a Vocation." In Gerth and Mills 1946.
- Weber, Max. 1922. "Science as a Vocation." In Gerth and Mills 1946.
- Zagorin, Perez. 2003. *How the Idea of Toleration Came to the West*. Princeton: Princeton University Press.