

THE SELF IN THE ORIGINAL POSITION*

DANIEL P. THERO

Since the publication of John Rawls's *A Theory of Justice* in 1971, there have been numerous criticisms directed against justice as fairness and liberalism generally. Such criticisms have come from diverse philosophical and ideological orientations. One school of thought which has attacked Rawls is that known broadly as "communitarianism." In this essay I will examine arguments and claims constituting one particular aspect of the communitarian assault: namely, that aspect which focusses on implications of Rawlsian theory for the nature of the human self. I conclude that these particular communitarian criticisms ultimately fail in establishing a decisive critique of Rawls. While a full exposition of the architecture of Rawls's theory would be beyond the scope of this essay, I will begin with a brief treatment of certain features of Rawls's thought that are especially salient for our purposes here.

A central element of Rawls's theory is a construct of hypothetical thought called *the original position*. I would characterize the original position as being defined by two features: (1) The persons ("parties") in the original position have the task of deciding upon the most abstract principles of justice, that will serve to regulate the basic structure of society, and (2) they must do this from behind what Rawls calls the veil of ignorance. The *veil of ignorance* functions as an epistemic screen or filter, denying to the parties behind the veil access to information which might serve to bias the decision procedure in such a way as to further special interests of the party members. Thus, being in the original position, behind the veil of ignorance, rules out knowledge of one's own unique status and circumstances. The only sort of information which "passes"

* I would like to thank professor Jon Mandle of the Department of Philosophy, University at Albany, for his helpful comments on an earlier version of this paper.

through the veil is that which consists of certain sorts of general facts about the world—for example, theories and principles of the social sciences. The motivation behind the introduction of the veil of ignorance is the elimination of any influence exerted by personal contingencies on the reasoning process leading up to the delineation of the basic societal structure.

Rawls says that we can, if we wish, view the choices made in the original position from the standpoint of one person selected at random from among those behind the veil: "If anyone after due reflection prefers a conception of justice to another, then they all do, and a unanimous agreement can be reached."¹ Also, original position parties are different from "us." That is, they are different from real persons living in a society whose basic institutions have not been chosen from behind a veil of ignorance. They are also to be distinguished from persons living in what Rawls calls a well-ordered society, which he identifies with the sort of society which would emerge after the establishment of a just basic structure, constitution, and legislation. In *A Theory of Justice*, Rawls articulates the distinction as follows:

The original position is not to be thought of as a general assembly which includes at one moment everyone who will live at some time; or, much less, as an assembly of everyone who could live at some time. It is not a gathering of all actual or possible persons. To conceive of the original position in either of these ways is to stretch fantasy too far; the conception would cease to be a natural guide to intuition.²

As we shall see, communitarians have sometimes been guilty of conflating the original position parties with persons in less restrictive epistemic situations. This is something which one must be careful not to do if one wishes to avoid misrepresenting Rawls's arguments.

Perhaps the most influential communitarian critic of Rawls, and the one upon which my focus shall be centered, is Michael Sandel. In general, communitarians like Sandel stress that the self is defined through discovered communal attachments, rather than through free choices. The fear is that the Rawlsian liberal's emphasis on the conditions for free choice will prevent us from engaging adequately in the discovery of our constitutive ties. Sandel refers to the Rawlsian self as an "unencumbered

¹ John Rawls, *A Theory of Justice* (Cambridge: Harvard U.P., 1971), p. 139.

² *Ibid.*

self," meaning that the self stands fundamentally alone and autonomous, neither fettered nor enabled in any way by prior commitments to family, social groups, religious institutions, or political units. He explicitly contrasts this with the encumbered self of the communitarian vision, a self which fits neatly into a complex niche not of its own creation. Such a self is said to derive meaning and significance solely from its participation in the life of those communal agencies which serve to constitute the parameters of the niche. In an article entitled *Beyond the Procedural Republic: the Communitarian Liberalism of Michael Sandel*, Terry Hall avers that the encumbered communitarian self is truly a different sort of moral character than is the unencumbered liberal self. This is because an encumbered self must include in all of its deliberations not simply what it would like to choose (to do or to be), but also *who it already is*.³ Thus, "communal selves must take into account the attachments that have already established identity."⁴

With the Rawlsian self, according to Sandel, there is "always a distinction between the values I *have* and the person I *am*. To identify any characteristics as *my* aims, ambitions, desires, and so on, is always to imply some subject 'me' standing behind them, at a certain distance, and the shape of this 'me' must be given prior to any of the aims or attributes I bear."⁵ The supposition is that this fixes the identity of the self permanently by ruling out what Sandel refers to as "constitutive ends."⁶ In other words, this self cannot be conceived to hold membership in any community which is constituted "by moral ties antecedent to choice; he cannot belong to any community where the self *itself* could be at stake."⁷ The self is prior to its ends in the sense of satisfying an epistemological requirement of being independently identifiable.⁸

Another aspect of Rawls's theory which is frequently subjected to communitarian assault is the priority of the right over the good. By "the

³ Terry Hall, "Beyond the Procedural Republic: The Communitarian Liberalism of Michael Sandel." Christopher Wolfe and John Hittinger (editors) *Liberalism at the Crossroads* (Lanham, MD: Rowman & Littlefield, 1994), p. 88.

⁴ *Ibid.*

⁵ Michael Sandel, "The Procedural Republic and the Unencumbered Self," *Political Theory*, Vol. 12, No. 1 (1984), p. 86.

⁶ *Ibid.*

⁷ *Ibid.*, p. 87.

⁸ Michael Sandel, *Liberalism and the Limits of Justice* (NY: Cambridge U.P., 1982), p. 20.

right" Rawls means the *concept* of justice, while "the good" refers to particular *conceptions* of a good life. These conceptions of what constitutes the good, although hopefully reasonable, are intentionally excluded from consideration at the level of the basic structure of society. Only by excluding such competing conceptions of the good from the original position, Rawls thinks, can we hope to achieve the sort of overlapping consensus which he declares to be the desired outcome of political liberalism.

Sandel locates the etiology of Rawls's prioritization of the right over the good in Kant's deontological liberalism. He stresses that for Kant, this prioritization is "grounded in the concept of a subject given prior to its ends, a concept held indispensable to our understanding ourselves as freely choosing, autonomous beings."⁹ Sandel characterizes the Kantian reasoning as follows: "Society is best arranged when it is governed by principles that do not presuppose any particular conception of the good, for any other arrangement would fail to respect persons as being capable of choice; it would treat them as objects rather than subjects, as means rather than ends in themselves."¹⁰ However, especially in *Political Liberalism*, Rawls distinguishes his view from that of Kant on the point that Kant makes metaphysical claims about the self, while Rawls insists that he does not make any such claims. Amy Gutmann, in an article entitled *Communitarian Critics of Liberalism*, characterizes the Rawlsian position as follows: "The major aim of liberal justice is to find principles appropriate for a society in which people disagree fundamentally over many questions, including such metaphysical questions as the nature of personal identity."¹¹ Sandel and at least some of the other communitarians appear either to miss this distinction between Kantian and Rawlsian claims about the self or, more likely, see the distinction but deny that Rawls has legitimate grounds to insist that we take it seriously.

Rawls thinks that parties in the original position will decide upon two principles of justice as the foundation upon which to build the basic structure for society. I am not going to examine the first of these principles in the context of this essay. The most noteworthy element of the *second* principle is referred to as the *difference principle*. Briefly, the

⁹ *Ibid.*, p. 9.

¹⁰ *Ibid.*

¹¹ Amy Gutmann, "Communitarian Critics of Liberalism," *Philosophy and Public Affairs*, Vol. 14, No. 3 (1985), p. 313.

difference principle states that inequality in the distribution of social primary goods is justified only to the extent that such unequal distribution will serve to raise the estate of the worst off members of society.¹²

I will now begin an examination of seven specific arguments targeted at the self as it appears in Rawls's theory. I am by no means claiming that these represent an exhaustive exposition of the ways in which Sandel and other communitarians can and have attacked the Rawlsian self. I am reasonably confident, however, that they cover sufficient conceptual terrain to afford adequate grounds for an evaluation of the overall strength of this one particular branch of the communitarian assault. Some of the arguments and claims attributed to Sandel in the sequel can be found explicitly in his work, while others are in part the product of distillation and interpretation by other commentators.

ARGUMENTS CRITICAL OF THE RAWLSIAN SELF

1. The first argument I will discuss is what I shall call the two-way lens argument. Found in Sandel's *Liberalism and the Limits of Justice*, this is a metaphor-driven intuition pump, rather than a rigorous argument. The original position is the device through which all justification leading to the two principles of justice must pass¹³. Sandel asserts that "what issues at one end in a theory of justice must issue at the other end in a theory of the person ..."¹⁴ It is in this context that Sandel introduces the metaphor of a lens:

Looking from one direction through the lens of the original position we see the two principles of justice; looking from the other direction we see a reflection of ourselves. If the method of reflective equilibrium operates with the symmetry Rawls ascribes to it, then the original position must produce not only a moral theory but also a philosophical anthropology.¹⁵

Sandel accuses Rawls of attributing philosophical implications to the product of the lens looking in one direction, while purposely neglecting the product that is obtained by looking through the lens in the reverse direction.

¹² Rawls, *A Theory of Justice*, p. 60.

¹³ Sandel, *Liberalism and the Limits of Justice*, p. 47.

¹⁴ *Ibid.*

¹⁵ *Ibid.*, pp. 47–48.

Before attempting to decide whether Sandel is correct in his assertion that the original position produces a philosophical anthropology, we need to ask what is undesirable with the model of the self which thus supposedly emerges. Sandel insists that the self which issues forth is overly individualistic and simplistic. As he says of this self, "no commitment could grip me so deeply that I could not understand myself without it. No transformation of life purposes and plans could be so unsettling as to disrupt the contours of my identity."¹⁶ Thus, what emerges is, once again, the unencumbered self, with Sandel's emphasis being placed on the unrealistic simplicity of this depiction of the person.

This argument seems initially impressive, especially for its employment of the imagery of the lens. However, in *Political Liberalism*, Rawls insists that the original position does not in fact carry any metaphysical implications, at least not of the strong sort envisioned by Sandel. Rawls expositis as follows:

The description of the parties may seem to presuppose ... that the essential nature of persons is independent of and prior to their contingent attributes, including their final ends and attachments, and indeed their conception of the good and character as a whole.

... I believe this to be an illusion caused by not seeing the original position as a device of representation. The veil of ignorance, to mention one prominent feature of that position, has no specific metaphysical implications concerning the nature of the self; it does not imply that the self is ontologically prior to the facts about persons that the parties are excluded from knowing.¹⁷

It must be admitted that this quote from Rawls, not offered specifically as a refutation for the two-way lens argument, provides only a general sort of answer to Sandel's objection. In the final analysis, perhaps the best answer that Rawls can give to the two-way lens argument is to confront Sandel with the questions: "Why should one think that the imagery of a lens is an appropriate metaphor for the original position?" and, "Even assuming the appropriateness of the metaphor, why think that one must pay equal attention to looking through the lens in both directions?" After all, taking the lens metaphor seriously, nobody criticizes biologists and astronomers for peering through the lenses of microscopes and telescopes in one direction only. The fact of the matter is

¹⁶ *Ibid.*, p. 62.

¹⁷ John Rawls, *Political Liberalism* (NY: Columbia U.P., 1993), p. 27.

that the vast majority of real lenses are intended to be used in one direction only, and this is not normally thought to constitute any sort of flaw in them.

2. A second criticism of Rawls is found in Sandel's article *The Procedural Republic and the Unencumbered Self*. The argument can be outlined as follows: The difference principle is a principle of sharing. This implies that one of the following two possibilities must be the case:

(1) The difference principle presupposes communal ties. If this is true, then it entails the incompatibility of the difference principle with the unencumbered self, since a lack of communal ties is supposed to be a necessary component of the concept of an unencumbered self.

(2) The difference principle does *not* presuppose communal ties, yet forces people to share anyway, which amounts to treating those people as means rather than as ends. Since treating people as ends is requisite to respecting their equality and autonomy, the difference principle fails to respect the equality and autonomy of persons.

The difference principle begins with the thought that assets I have are only accidentally mine. But according to Sandel, "it ends by assuming that these assets are therefore *common* assets and that society has a prior claim on the fruits of their exercise."¹⁸ Sandel thinks this assumption to be without warrant, arguing that even if one was to grant that *I* do not have a privileged claim on the assets, it does not therefore follow that everyone *collectively* does. He reasons that it is not at all clear why we should suppose it to be more arbitrary to assign to the individual a privileged claim over assets than to assign such a claim to the collective of the would-be liberal society.¹⁹ Sandel continues on to add that, as a principle of sharing, the difference principle "must presuppose some prior moral tie among those whose assets it would deploy and whose efforts it would enlist in a common endeavor. Otherwise, it is simply a formula for using some as a means to others ends, a formula this liberalism is committed to reject ..."²⁰

If we accept Rawls's claim that the difference principle *does not* use persons as means merely, then according to Sandel and his proponents, we need to begin speaking of a group or community subject. As C. Edwin Baker says in his article *Sandel on Rawls*, "Only a group or commu-

¹⁸ Sandel, "The Procedural Republic and the Unencumbered Self," p. 89.

¹⁹ *Ibid.*

²⁰ *Ibid.*

nity subject could both choose the difference principle and, since each person's talents would belong to this larger subject, avoid treating the moral subject as a means."²¹ The upshot of all this is that the idea of an unencumbered subject conflicts with the principle that Rawls wishes to endorse. In other words, he cannot justify the difference principle within the framework of his complete theory.²²

In response to this strain of criticism, a Rawlsian could respond by asking whether "a moral bond" necessarily implies "a constitutive community," or whether "communal ties" must necessarily be constitutive in a strong sense. Could not a *chosen* commitment be a strong moral bond, without being constitutive in the radical sense intended by Sandel? Rawls simply denies that his conception of the subject is one that is antithetical to attachments which in part constitute one's over-all identity as a self.

A second possible Rawlsian response would be to insist that to regard talents as common assets is something which is to be done only for the purpose of the most basic institutional design. The intention is to bring about the result that "no one will be arbitrarily favored or subordinated in the inevitably collective and historical process of deciding what claims to recognize and what activities to favor."²³ In other words, once a well-ordered society is established, it would no longer be necessary to insist on regarding talents as if they were common assets.

If *all* aspects of the self were relevant to the process of formulating the basic principles of justice, then Rawls would need to develop and work with a comprehensive philosophical anthropology. But, so Baker states,

This answer is wrong. We often conclude that various aspects of what we commonly consider to be a person do not properly contribute to the justification of social policy. For example, we commonly consider a person's preferences and personality type to be aspects of the person.²⁴

²¹ C. Edwin Baker, "Sandel on Rawls," *University of Pennsylvania Law Review*, Vol. 133 (1985), p. 896.

²² See Baker, p. 908, and also Chantal Mouffe, "American Liberalism and its Critics: Rawls, Taylor, Sandel and Walzer," *Praxis International*, Vol. 8 no. 2 (1988), p. 198.

²³ Baker, p. 910.

²⁴ *Ibid.*, p. 911.

However, aspects such as personal preferences and personality do not present themselves as being appropriate justifications for social policy. Those aspects of the subject which *are* relevant to the basic structure and to the achievement of an overlapping consensus constitute only a subset of the aspects which may in fact be relevant for making personal decisions or, for that matter, for the drafting of specific legislation within a well-ordered society.²⁵

3. A third communitarian critique of the Rawlsian subject centers around the claim that a sparse self cannot self-reflect. Baker summarizes this argument, which he ascribes to Sandel, as follows: "... Rawls is committed to a thin, denuded notion of the person — a person separate from all ends, commitments, and capacities. This self is so sparse that it cannot constitute an object for self-reflection."²⁶

I will advance what I think are two reasonable responses from the Rawlsian perspective. The first is something we have already encountered: the admonition to keep original position parties, members of a well-ordered society, and us, in separate conceptual compartments.²⁷ Rawls would clearly want to maintain that he does *not* think that we are thin, denuded selves, or that members of a well-ordered society would be emaciated subjects either. It is only in the original position that the moral subject is discussed as if it lacks any unique personality or thick conception of the good. And even here, it is not apparent that we need to think of original position parties *as actually lacking* a thick conception of the good which might well include highly significant communal attachments. We must remember that the veil of ignorance is an epistemic screen rather than a metaphysical condition. Its purpose is to limit access to information in such a way that the choosing subjects are rendered for all practical purposes as having only a thin concept of the good in relation to their role as architects of the most basic principles of distributive justice. In other words, I do not think that it would belie Rawls's project to think of the original position parties as owning thick conceptions of the good, but being rendered helpless by the veil of ignorance to access (and thus to make decisions based upon) those conceptions during the negotiation process.

²⁵ *Ibid.*

²⁶ *Ibid.*, p. 896.

²⁷ Rawls, *Political Liberalism*, p. 28.

A brief machine analogy might prove helpful here. Think of each original position party as being like a computer, possessed of a hard drive filled with data. In each case, the data would include the specifications for the computer system itself. These specifications would serve as indications of the computer's capacities and limitations, the types of applications for which it would be most suited, and its compatibility with other systems and with various possible operating environments. Let us designate this as Data Set 1. Each computer also contains all sorts of information pertinent to standard operating requirements for all computers, such as the need for a power source and for protection against electrical surges, as well as general information about the physics of electrical circuits. Label this Data Set 2. Imagine that each computer is also provided with a program that draws upon the available data to design an "ideal world" for that computer and for other computers with specifications identical to its own. Finally, to interject an analogy with the veil of ignorance, imagine that another piece of software is introduced which blocks access to Data Set 1, while allowing unlimited access to Data Set 2. Thus, as each machine carries out its mandate, it can draw only upon the data set relevant to all computers, just as original position parties can draw upon only that information relevant to all human subjects. The other data set is still present, however; it is just that access to it is blocked in all and only those instances when the computer is running the "ideal world" program. The rest of the time, the computer can access all of its stored data. In fact, let us assume that the information in Data Set 1 plays a vital role in the operation of many of the other programs that our imaginary computers happen to run most of the time. Analogously, original position parties are not metaphysically deprived entities suffering from an acute dearth of those features which allow one to occupy a meaningful place within a fully articulated human community. Rather, the Rawlsian is simply proposing that the set of data which could serve to individuate the choosing subject be rendered inaccessible for the duration of the negotiation procedures to occur in the original position.

Baker asserts that "Rawls only needs a theory of those aspects of the person or of human interaction that are relevant to his enterprise. Sandel's error ... lies in assuming that those few universal qualities that Rawls emphasizes reflect a complete Rawlsian theory of the person."²⁸ In fact,

²⁸ Baker, p. 896.

regarding those aspects of the self highlighted in Rawlsian discussions of the original position, Baker says:

these general aspects ... need not be the ones that are most important to what a fully constituted person takes herself to be. The claim that these aspects lead to basic principles of social organization only means 'basic' in a trumping or priority sense, not that these principles are necessarily experientially most important or most burdensome.²⁹

Furthermore, despite his emphasis on persons as rational, autonomous agents concerned to advance their own interests, "Rawls does not imply or assume that this conception of the person is empirically or historically accurate, or even that it is a relevant conception of the person for other purposes."³⁰

4. A fourth communitarian critique of the Rawlsian subject, which Amy Gutmann distills from Sandel's work, is what could be called the "encumbered selves won't want justice because they have community" argument. Gutmann expounds that "Sandel seems to mean that communally given ends can so totally constitute people's identities that they cannot appreciate the value of justice."³¹ Perhaps the most direct response to this criticism would be to ask the empirical question of whether or not this is really the case. After all, even people in the ancient world wrote about justice—long before the advent of liberalism as a political philosophy. Admittedly, in some traditional societies I expect that there is scant concern for justice as such, or at least not for distributive justice. If this is correct, perhaps it is because people do not feel oppressed or taken advantage of in these particular societies. Also, there may be fewer primary goods to be distributed, so that Rawls's difference principle would not find a significant venue for application. However, in traditional societies in which grossly unfair practices are suddenly instituted without precedent (as for example when a particularly bad leader comes to power), I suspect that a concern for justice (or for something quite like justice) would arise rather quickly. But this really amounts to an anthropological hypothesis, and thus calls for empirical investigation. In order for communitarians to give teeth to the criticism that properly encumbered selves will find a concern for justice to be superfluous or even detrimental to community, they need to produce

²⁹ *Ibid.*, p. 899.

³⁰ *Ibid.*, p. 901.

³¹ Gutmann, p. 313.

evidence to this effect. Simply speculating about what is the case in this regard will not be sufficient to refute Rawls or to convince persons who are not already sympathetic to the communitarian stance.

5. A fifth criticism which follows from Sandel's thought starts from the claim that *if* the self is fundamentally a choosing subject, then it must follow that it is "most itself" when unencumbered.³² This is because "encumbrance" restricts its freedom of choice by imposing boundary conditions. Following from this, the self forfeits nothing essential to its human agency when it withdraws from a voluntary attachment.³³ To quote Sandel, "For the unencumbered self, what matters above all, what is most essential to our personhood, are not the ends we choose but our capacity to choose them. The original position sums up this central claim about us."³⁴

In response to this criticism, the Rawlsian liberal could remind Sandel that Rawls does *not* claim that the human agent is fundamentally just a choosing subject. Furthermore, *even if* Rawls did claim this, it is not at all clear how it follows that it is most itself if totally unencumbered by any prior commitments or attachments. Why can we not see the prior commitments as the material from which choices are "fashioned"? Additionally, why could we not have just as much grounds for saying that the fact that I *choose* a commitment means that I forfeit *more* when I abandon that commitment than if I did not choose the commitment for myself? In fact, it may be the case that there is no significant correlation between the degree to which a commitment is chosen, as opposed to "discovered," and the nature and amount of what is forfeited if that commitment is abandoned. That is, some voluntarily chosen commitments may be much deeper than some discovered commitments, and *vice versa*. Why assume that there exists anything like a direct proportionality between the seriousness of a commitment and the degree to which it is discovered rather than adopted?

In *Political Liberalism*, Rawls espouses the view that the stability of one's *public identity* has no direct bearing on the stability or changeableness of the elements which make up one's encumbered self. He is fully aware that persons in our society, and presumably persons living in a well-ordered society as well, will "regard their final ends and attach-

³² Hall, p. 81.

³³ *Ibid.*, p. 82.

³⁴ Sandel, "The Procedural Republic and the Unencumbered Self," p. 86.

ments very differently from the way the political conception supposes."³⁵ He goes on to say that when changes occur in our conception of the good,

... we are likely to say that we are no longer the same person. We know what this means: we refer to a profound and pervasive shift, or reversal, in our final ends and commitments; we refer to our different moral (which includes our religious) identity Yet such a conversion implies no change in our public or institutional identity.³⁶

We are thus advised to maintain separate conceptual spaces for the unencumbered public self and the encumbered complete self, the idea being that if we can accomplish this, we may avoid the error which appears to underlie this particular communitarian criticism.

6. A sixth objection concerning the Rawlsian self is that it precludes there being any warrant for positing a plurality of parties in the original position, given the description of what agents are like behind the veil of ignorance. Sandel asks rhetorically what could individuate original position parties, given that they possess "radical equality as choosers."³⁷ If one cannot effectively individuate the agents, then what warrant is there for speaking of more than one agent being present? If it is necessary to have a plurality of agents in order to carry on negotiations regarding the foundational principles for society, then the possibility of any such negotiations is placed in serious jeopardy. Terry Hall, speaking of the original position parties, voices this objection as follows:

... Rawls wants to say that they all make the same choice; they all choose the same principles of justice. But this is just because they are identical *as agents of choice*. And their capacity to choose exhausts their identity here.

... In consequence, not persons but only a single subject is found behind the veil of ignorance.³⁸

In response to this assessment, I think that a Rawlsian could say that the fact that all of the parties agree, and could perhaps be thought of functionally as one subject, actually *supports* the claim that the decision procedure is rational. A high degree of agreement on matters concerning

³⁵ Rawls, *Political Liberalism*, p. 31.

³⁶ *Ibid.*

³⁷ Sandel, *Liberalism and the Limits of Justice*, p. 131; also see Hall, p. 84.

³⁸ Hall, p. 84.

the principles of justice underlying the basic structure would seem to be in line with Rawls's assertion that he is doing "moral geometry." Just as two geometers who start out with the same information and rules should arrive at the same answers, so too will the original position parties, each having the same restraints placed on his or her knowledge, arrive at the same abstract principles. Rawls says that any one of us can at any time imagine himself or herself as being an original position party, with all of the requisite knowledge constraints in place, and think our way through to pretty much the same conclusion. While the two principles of justice do not follow deductively from the setup of the original position, Rawls clearly expects that they *will* be the product of the reflection of anyone who honestly imagines assuming the mask of the veil of ignorance.

7. A final objection based in Sandel's thought is that "self discovery" is more in line with our deepest intuitions than is the Rawlsian view of making one's self through autonomous choices. This might be referred to as the "self-perception argument." This argument is highlighted by Will Kymlicka in his article *Liberalism and Communitarianism*. If the self is really prior to its ends, says the argument, then introspection should allow one to see through or past one's particular ends to an "unencumbered self." But this is not what we find; "our deepest self-perceptions always include some motivations, and this shows that some ends are constitutive of the self."³⁹

There are several possible Rawlsian responses to this argument. First, the issue of "perception" is quite possibly misleading here. As Kymlicka says, "what is central to the liberal view is not that we can *perceive* a self prior to its ends, but that we understand ourselves to be prior to our ends, *in the sense that no end or goal is exempt from possible re-examination.*"⁴⁰ In fact, this does not even require that we can perceive a self totally unencumbered by any ends. Rather, we ought to consider that the process of ethical thinking in general is one of "comparing one 'encumbered' potential self with another 'encumbered' potential self."⁴¹ The self must have *some ends* when it reasons ethically, but it fails to

³⁹ Will Kymlicka, "Liberalism and Communitarianism," *Canadian Journal of Philosophy*, Vol. 18 no. 2 (1988), p. 190.

⁴⁰ *Ibid.*

⁴¹ *Ibid.*

follow from this "that any *particular* ends must always be taken as given with the self."⁴²

A second and somewhat "aggressive" Rawlsian response would be to assert that it is actually Sandel who is violating our deepest self-understandings. According to this line of reasoning, people really do not think that self-discovery "replaces or forecloses judgments about how to lead one's life."⁴³ Even when we feel deeply intertwined and implicated in some tradition, practice, or form of life, we still feel that we can, if we wish, question whether the practice is in fact valuable, either to us personally or in general. Such questioning is not meaningful on Sandel's account.⁴⁴ As Kymlicka states: "The idea that moral reasoning is completed by this process of self-discovery (rather than by judgments of the value of the attachments we discover) seems pretty facile."⁴⁵

This concludes my examination of arguments against the Rawlsian self or subject that have issued forth from Sandel and his expositors. A problem which I mentioned specifically at several places but which I find to be virtually ubiquitous in the literature is the confusion of persons in a non-well-ordered society, original position parties, and persons in a well-ordered society. I am inclined to attribute at least part of this confusion to the fact that all but one of the communitarian sources examined were published prior to *Political Liberalism*, in which Rawls is more overtly concerned than in *A Theory of Justice* to keep separate the three types of agents.

Before attempting to draw any conclusion regarding the over-all merit of Rawlsian liberalism versus communitarianism, we must keep in mind that the arguments examined above deal with only one of several aspects of the Rawlsian theory which could serve as possible loci for communitarian assault. Needless to say, the totality of this debate will not come to a resolution here. Before concluding, however, I wish to make mention of a challenge facing communitarians in general. As we have seen, communitarians such as Sandel speak constantly of commitment and embeddedness, of encumbered versus unencumbered selves. While they appeal to often powerful human intuitions and desires concerning connectedness and community, nowhere is there seen an attempt on

⁴² *Ibid.*

⁴³ *Ibid.*, p. 191.

⁴⁴ *Ibid.*

⁴⁵ *Ibid.*, p. 192.

their part to provide us with a way to measure the depth or strength of commitment, or the quality of commitments.⁴⁶ As Allen Buchanan comments: "Until such a theory of commitment is available, any evidence that liberal societies contain forces that tend to hinder commitments will be incomplete and to that extent indecisive."⁴⁷ While it may well be the case that communitarians are right about the existence of such forces, it seems incumbent upon them to refine their evidence. Perhaps they would retort that this is a practical concern and that they as theorists should not be saddled with any obligation to supply such evidence. Although there may be some merit to this retort, it certainly seems that communitarians would strengthen their position and gain more adherents if they could offer more solid evidence that their deepest motivational concerns are indeed empirically grounded.

State University of New York at Albany

⁴⁶ Allen E. Buchanan, "Assessing the Communitarian Critique of Liberalism," *Ethics*, Vol. 99 (1989), pp. 866-867.

⁴⁷ *Ibid.*