Politics, Philosophy & Economics

The burdens of public justification: constructivism, contractualism, and publicity
Samuel Freeman

The online version of this article can be found at:

Published by:

$SAGE

http://www.sagepublications.com

On behalf of:

The Murphy Institute of Political Economy

Additional services and information for *Politics, Philosophy & Economics* can be found at:

Email Alerts: http://ppe.sagepub.com/cgi/alerts

Subscriptions: http://ppe.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

>> Version of Record - Jan 23, 2007

What is This?

The burdens of public justification: constructivism, contractualism, and publicity

Samuel Freeman

University of Pennsylvania, USA

abstract
The publicity of a moral conception is a central idea in Kantian and contractarian moral theory. Publicity carries the idea of general acceptability of principles through to social relations. Without publicity of its moral principles, the intuitive attractiveness of the contractarian ideal seems diminished. For it means that moral principles cannot serve as principles of practical reasoning and justification among free and equal persons. This article discusses the role of the publicity assumption in Rawls's and Scanlon's contractualism. I contend that a regard for publicity and a moral conception's potential to provide a public basis for justification and agreement account for much of the evolution of Rawls's account of justice after *A Theory of Justice*. I also discuss whether contractualism can provide a basis for justification and general agreement under the social conditions that it endorses. I contend that it cannot, and conclude with a discussion showing why this should not be a problem for contractualism. Despite appearances, contractualism is a distinctive form of contractarianism, substantially different from Rawls's position and the social contract tradition out of which it evolved.

keywords
contractarianism, contractualism, John Rawls, public justification, T.M. Scanlon, justice

The publicity of a moral conception is a central idea in Kantian and contractarian moral theory. Publicity carries the idea of general acceptability of principles through to social relations. When satisfied by a contract conception, publicity shows that it is within human capacities to live together according to principles acceptable to all reasonable persons. Otherwise, general acceptability and agree-

ment is an unobtainable ideal, satisfiable only at the theoretical level of hypothetical moral agents. It might still be important, but without publicity the intuitive attractiveness of the contractarian ideal as an account of justice and other moral principles seems diminished. For among other things, the feasibility and stability of a fully public moral conception means that the moral principles that should structure and regulate social and political relations are both within human reach and are capable of serving reasonable agents as principles of practical reasoning, and public justification and criticism.

Publicity has two significant consequences when satisfied by a moral conception. First, moral agents can know the real reasons for moral constraints and expectations and can apply these reasons to plan their actions and pursuits. Moral grounds for action need not be occluded in individual deliberations. This is crucial to a person being a fully responsible moral agent. It is also essential to self-knowledge, knowing significant social influences on the kind of person one is. For Kantian and Marxian views, full publicity is a precondition of moral and rational autonomy. Second, their publicity gives moral principles a social function: a community of agents can rely on moral principles as a shared basis for discussion, argument, and agreement. They can then assess and criticize actions and institutions using shared criteria, and justify them to one another, when they are justifiable, on the basis of reasons all accept. This connects with the idea of mutual respect in the relations of fully responsible agents. Without the publicity of ultimate principles and other justifying reasons, justification (where it obtains at all) is incomplete. Public discussion is then susceptible to manipulation by emotionally persuasive rhetoric and appeals to biases and interests that are all too familiar in contemporary political life. Even worse, agents may be prone to general illusions about their social life, with no one in a position to discern why moral relations are as they are.

The publicity requirement then raises the question of whether the basic reasons and principles a moral conception advances can be relied upon in individual deliberation and in public argument and justification. It examines whether these reasons and principles are within human reach, and suited to our capacities for understanding, judgment, motivation, and action.

The initial focus in this article is the role of publicity in John Rawls's constructivism. I contend that a regard for publicity and a moral conception's potential to provide a public basis for justification and agreement account for much of the evolution of Rawls's account of justice after *A Theory of Justice*. I contend that a helpful way to understand Rawls's transition to *Political Liberalism* is to see it as a response to the tensions Rawls discovered between his dual aims of the full publicity and public justification of a moral conception of justice.

In the second half of this article, I focus on the role of publicity in T.M. Scanlon's moral contractualism. Contractualism also seeks public principles to provide 'a basis for general agreement'. A natural question is whether it is subject to tensions similar to those confronted by Rawls. I contend that, if we see the

publicity or moral principles as full or complete within a well-ordered contractualist society, contractualism encounters similar problems. This raises a question as to whether contractualism can provide a basis for justification and general agreement under the social conditions that it endorses. If it cannot, is this a problem for contractualism? I suggest ways that the failure of its full publicity should not be regarded as problematic given contractualism's distinct aims: despite appearances, contractualism is a distinctive form of contractarianism, substantially different from Rawls's position and the social contract tradition out of which it evolved.

1. Constructivism: political, moral, and metaphysical

Generally, constructivism in ethics addresses a traditionally metaphysical question regarding the possibility and nature of moral truth or similar standards of correctness ('reasonableness' for Rawls or 'universal validity' in Kant). As opposed to moral realism, constructivism denies that moral statements correspond to antecedent moral facts or to a realm of values that are prior to and independent of practical reasoning. As opposed to moral skepticism, which denies moral truths and objective standards of correctness, constructivism affirms that there are truth conditions for moral statements. Finally, as opposed to relativistic conceptions of moral truth (in cultural relativism, social constructivism, and some forms of communitarianism) moral constructivism affirms a universal conception of moral objectivity and applies fundamental moral principles to all persons capable of understanding moral requirements.

There are different kinds of moral constructivism, but all have in common the idea that moral principles are correct (true, reasonable, universally valid, and so on) when they are the outcome of a deliberative procedure that incorporates all the relevant requirements of correct reasoning.[1] In this regard, constructivism affords priority to the objectivity of judgment in determining moral truth: moral truths are the judgments that would be made by correctly reasoning from an *objective point of view*, which is a deliberative procedure incorporating all relevant requirements of practical reasoning. By contrast, realism reverses this priority and explains objectivity of judgments in terms of their satisfying the conditions needed for ascertaining prior moral truths. In this regard, moral constructivism rejects moral realism's basic claim that morally true or correct judgments represent (in some manner) moral facts or moral principles that are prior to anyone's practical reasoning and judgment. Apart from reasoning in terms of the deliberative procedure for constructing moral principles 'there are no moral facts' Rawls says.[2]

There are different ways to develop constructivism in ethics depending upon the design of constructivist procedures and the requirements deemed relevant to practical reasoning. Hobbesian contractarians incorporate a conception of practical reason as individual utility maximization along with a conception of human

nature as largely self-focused and indifferent to others. John Harsanyi's utilitarian choice procedure relies also upon a utility-maximizing account of practical rationality, which is conjoined with a thin veil of ignorance and a Bayesian equiprobability assumption which insures equal consideration of everyone's interests. T.M. Scanlon's contractualism eschews utility maximization; his contractors have interests to protect, but their primary motivation is to justify themselves to others on terms all can reasonably accept. In Rawls's *Lectures on the History of Moral Philosophy*, he interprets Kant's categorical imperative as a non-contractarian constructivist procedure that assesses the validity of the rules ('maxims') that are implicit in rational actions. Finally, for Rawls himself, among the 'relevant requirements of practical reason'[3] to be incorporated into his constructivist procedure (the original position) are ideal conceptions of persons as free, equal, and reasonable and rational, and of society as well ordered by a public conception of justice that all generally accept. The aim of moral constructivism for Rawls is to show how moral principles of justice are connected with these and other relevant ideas of reason and principles of practical reason. On the assumption that all relevant requirements of practical reasoning have been incorporated into the original position, the moral principles of justice that are unanimously agreed to might then be said to be among the requirements of practical reason.[4]

By incorporating ideals of the person and society, Rawls integrates features of Kant's view into his version of constructivism. But 'Kant's moral constructivism', Rawls says, is only one version. It is a 'comprehensive moral doctrine' with claims about the nature and origin of moral value and the supreme human good. It says that the activity of practical reason makes or creates the order of values and that the order of (moral) values has no other source. This is 'constitutive autonomy'.[5] Kant's thesis is that practical reason *constitutes* the order of values out of itself and its own resources (its principles and ideas).

Rawls's 'political constructivism' by contrast does not imply that practical reason originates morality or that moral principles have no other source.[6] It makes a more modest claim: that democratic principles of justice can be 'represented' or 'displayed' as the outcome of a procedure based in practical reason and a conception of citizens as free and equal moral persons. This is 'doctrinal autonomy'[7] (an extension of what Rawls called earlier 'the independence of moral theory'). Rawls sees doctrinal autonomy as important for the justification of a political conception in a democracy, since it depicts political principles of justice as moral requirements that are not externally imposed upon free and equal citizens. Rather, principles are represented as the product of democratic citizens' act of collective self-government and self-determination on the basis of shared democratic reasons. Political constructivism aims to uncover the conception of justice best associated with what we accept as good political reasons as citizens within a democratic society. There is no (metaphysical) claim that the values of democratic justice *are* self-imposed out of our reason with no other original

8

source. They may be, then again they may be imposed by God or the nature of things or by the effects of circumstance on our emotions or imaginations. Political constructivism leaves open ('brackets') philosophical questions about the real origins of moral principles and their ultimate epistemic status.

Kant's constructivism, by contrast, responds to these philosophical questions, and argues that moral reasons are universally valid, that reason is a self-originating source of moral value, and that we are morally autonomous beings. These aims stem, Rawls says, from Kant's view of 'the role of philosophy as apologia: the defense of reasonable faith . . . [that is] showing the coherence and unity of reason, both theoretical and practical, with itself'.[8] The aim of political constructivism is different: to uncover a *public basis of justification* on questions of political justice among people who have different, conflicting, but still 'reasonable' philosophical, religious, and moral views.[9]

Seemingly, it is also part of Kant's constructivism to achieve a public basis for justification on questions of political justice and for morality generally. For publicity plays a role in Kant's account of justice too, and it is implicit in his idea of a realm of ends. Rawls once said that in Kant's moral constructivism, unlike rational intuitionism, first moral principles are not regarded as justified and true independent of 'the public social role of moral doctrines'.[10] What then are we to make of Rawls's later claim that only political constructivism can seek a public basis of justification? Most striking is his following claim: 'Thus, it is *only* by affirming a constructivist conception – one which is political and not metaphysical – that citizens generally can expect to find principles that all can accept.'[11] The implications of this sweeping claim, if true, is that not just Kant's doctrine, but any traditional moral conception (justice as fairness included) is not feasible so far as it aspires to be the public grounding of substantive moral and political principles.

To understand what leads Rawls to this position and appreciate the structure of political constructivism, it helps to examine the evolution of constructivism in Rawls's theory – starting with its origins in *A Theory of Justice* (discussed below in Section 2), then tracing its path through Kantian constructivism (Section 3)[12] to political constructivism (Section 4). Underlying Rawls's claim regarding the limits of public political philosophy is a tension between two of the stated aims of Kantian constructivism which makes it unworkable. On the one hand, there is what I will call (following Scanlon) a 'contractualist' strand which aims to discover a public basis for justification and agreement acceptable to all reasonable citizens. On the other, there is a Kantian strand which manifests a regard for a political system that best realizes the conditions for moral agency and the autonomy of citizens. To fit these two strands of argument into one view, Rawls modifies Kantian constructivism, giving up certain Kantian aspirations. Consequently, he arrives at political constructivism, which aims to encompass an independent domain of moral reasoning, that is, the 'domain of the political'.

2. The Kantian interpretation of justice as fairness

The first indication of 'Kantian constructivism' appears in Section 40 of *A Theory of Justice*, the 'Kantian Interpretation of Justice as Fairness'. After presenting the argument from the original position, Rawls says:

> there is a Kantian interpretation of the conception of justice . . . This interpretation is based upon Kant's notion of autonomy . . . [Kant] begins with the idea that moral principles are the object of rational choice. They define the moral law that men can rationally will to govern their conduct in an ethical commonwealth . . . once we think of moral principles as legislation for [such] a kingdom of ends, it is clear that these principles must not only be acceptable to all but public as well . . . Kant [also] supposes that this moral legislation is to be agreed to under conditions that characterize men as free and equal rational beings. The description of the original position is an attempt to interpret this conception.[13]

By acting from the principles that would be chosen in this position, 'persons express their nature as free and equal rational beings'[14] and are autonomous.[15]

So far as the argument from the original position goes, the Kantian interpretation is just one way, by no means necessary, to make the argument for Rawls's principles of justice. Rawls's main argument in Part I of *A Theory of Justice* relies primarily on our (presumed) considered convictions of fairness and what we deem to be reasonable restrictions on arguments for principles of justice; on this basis, he constructs the original position. The Kantian interpretation has no essential role here; it is attached, as it were, as an afterthought in Section 40.

The Kantian interpretation, however, comes into its own in *A Theory of Justice*, Part III, 'Ends'. Part III aims primarily to show the feasibility and endurance (or, as Rawls says, the 'stability') of justice as fairness. A conception of justice is *inherently stable* when it internally generates its own support. It does so when people living under conditions of a just or well-ordered society come to have a steadfast will to do justice. If it can be shown how persons can develop a willingness to do justice for its own sake under the optimal conditions of a well-ordered society in which justice is implemented,[16] then, Rawls says, 'we might say that humankind has a moral nature'.[17] If this cannot be shown, then a conception of justice is utopian, not realistically possible given human nature and the enduring facts of social life.

The argument Rawls advances for the stability of justice as fairness in *A Theory of Justice* depends upon showing the 'congruence of the right and the good'. The congruence argument aims to show that it is rational, or in each person's interest, to be reasonable. Having the virtue of justice is then an intrinsic *good* for each person moved by justice under these ideal conditions. The central argument Rawls has for this strong claim depends on the Kantian interpretation. Very roughly, the Kantian congruence argument asserts that (on grounds of the Aristotelian principle, a psychological law that says, other things being equal, people prefer activities that engage their higher capacities and skills[18]) it is ration-

10

al for moral persons to want to express their nature as free and equal rational beings. For to do so is to realize fully one's moral and rational capacities for practical reasoning, and thereby be autonomous. Since choice from the original position expresses our nature as free and equal rational beings, the desire to express our nature and be autonomous is the same desire, under a different description, as the desire to act on and from the principles of justice. So it is rational to want to be a just person, and the right and the good are 'congruent'.[19]

Without further comment on this complicated argument, I simply emphasize that its conclusion assumes that everyone in a well-ordered society has sufficient reason to affirm their desire to do justice for the *same* reasons: because it enables each to realize their common nature as moral agents and as morally autonomous rational beings. A well-ordered society is unified in its basic moral beliefs, and its citizens endorse the same (partial) conception of their good as autonomous agents. This purportedly accounts for the stability of a well-ordered society. Problems lie ahead for this argument.

3. Kantian constructivism: Rawls's Dewey Lectures

The argument for congruence is central to the second part of Rawls's argument for the principles of justice in *A Theory of Justice*; it is designed to 'confirm', Rawls says, the principles of justice agreed to in the original position. Congruence does so by arguing that justice as fairness is rational, compatible with both human nature and the human good. In the Dewey Lectures of 1980, 'Kantian Constructivism in Moral Theory', Rawls gives the Kantian interpretation greater prominence in the first part of his argument for the principles of justice by showing how the original position interprets a Kantian conception of the person and society and incorporates principles of practical reasoning.[20]

Two features of Kantian constructivism warrant discussion here and a third a brief mention, since all are important for the development of political constructivism. First, Rawls makes explicit the 'social role of principles of justice'. This says that a condition of the justification of a conception of justice is that it meets a practical requirement, namely, that the conception be capable of serving as a public basis in light of which citizens can justify to one another their common institutions.[21] Second, Rawls expands the publicity requirement on principles of justice. The 'full publicity condition' requires that for a conception of justice to be justified, not only must its principles be publicly known in a well-ordered society, but also its complete justification must be available to all and acknowledged as the basis for political decisions. Third is the feature of constructivism mentioned earlier, common to all versions of it: Rawls says, 'an essential feature of constructivism is that its first principles single out what facts . . . are to count as reasons of justice. Apart from the procedure of constructing these principles, there are no reasons of justice.'[22] Kantian constructivism defines a distinct position regarding the nature of morality and its validity and objectivity. The

11

objectivity of our judgments of justice depends not on the judgments being true to a moral order antecedent to reason, but on their 'reasonableness'. 'Judgments are reasonable and sound if they result in following the correct procedure correctly and rely only on true premises.' Here, 'the correct procedure' is one that incorporates 'all the relevant requirements of practical reasoning'.[23]

Constructivism in philosophy is normally associated with the third feature Rawls mentions; it is a distinct view about objectivity and validity. My focus here, however, is on the first two features Rawls specifies as central to Kantian constructivism. I suggest below that the dual aim of providing a public basis of justification, on the one hand, and the full publicity of a conception of justice, on the other, for all their apparent similarities, ultimately stems from two separate strands of thought in Rawls. By insisting on full publicity, Rawls renders the Kantian interpretation of justice as fairness (and Kantian constructivism) impossible as a shared basis for public justification. Because of the difficulty of combining these two features of Kantian constructivism (full publicity and the social role of public justification) in a feasible moral conception, Rawls relinquishes certain ideas and redefines and develops others that are central to Kantian constructivism and to the Kantian interpretation. The result is political constructivism.

Begin with the first idea mentioned, that is, with the practical task of political philosophy. In the Dewey Lectures, Rawls says that in Kantian constructivism:

> conditions for justifying a conception of justice hold only when a basis is established for political reasoning and understanding within a public culture. The social role of a conception of justice is to enable all members of society to make mutually acceptable to one another their shared institutions and basic arrangements, by citing what are publicly recognized as sufficient reasons, as identified by that conception.[24]

I take Rawls to be saying here that the philosophical justification of a conception of justice (from our point of view, in reflective equilibrium) has as one condition that the conception itself be capable of providing what Rawls later calls 'a public basis for justification' among persons to whom that conception is intended to apply. For Rawls's purposes, this means that a conception of justice is justified only if it is generally acceptable as a shared basis for justification among free and equal persons in a well-ordered society.

The idea of justification at work here is already implicit in *A Theory of Justice*, where Rawls says that justification 'is designed to reconcile by reason [and] proceeds from what all parties to the discussion hold in common'[25]. Taken as a general epistemological claim, one might object that this misses what is essential to justification, namely, that it aims at truth. Equally controversial is Rawls's claim that, reasoning from what is held in common, 'one of the aims of moral philosophy' is to discover and suitably express possible bases of agreement where none seem to exist.[26] Again, a natural objection is that 'Discovering bases of agreement is at most a subordinate aim in moral philosophy, as a means to

12

achieving civility and widespread acceptance of moral principles that are true; but how could it ever take priority over the search for moral truth?'[27]

Rawls does not make these claims regarding justification and moral philosophy because he endorses a general non-realist account of objectivity, justification, and truth for all subjects. Rather, he thinks that the search for reasoned agreement is especially appropriate in morality and matters of justice, for it is a basic *moral* requirement that directly ensues from (his conception of) what morality (or at least justice) is about: 'respect for persons is shown by treating them in ways that they can see to be justified'.[28] Moreover, being prepared to *give* a justification to others, explaining the grounds for one's actions that constrain or adversely affect those others, is part of the duty of mutual respect.[29] For Rawls, respect for others as free and equal persons is so integral to our understanding of what morality (or at least justice) is about that this abstract moral requirement plays a role in defining not simply moral principles of justice, but the justification, objectivity, and validity of these principles. Rawls's constructivism, along with reflective equilibrium, stems not from adherence to a general epistemology; rather his moral epistemology ultimately relies upon commitment to certain fixed, considered moral convictions which taken together support the abstract conviction that justice involves respect for persons as free and as equal.[30] The dependence of moral justification upon moral requirements of justice is an aspect of what Rawls means by 'The Independence of Moral Theory'.[31]

Still, utilitarians and others might object here: 'But why should respect for persons as equals require justifying ourselves *to* others (whatever that might mean)? Equal consideration of their interests is sufficient to respect them as persons?'[32] The contractarian's reply is that persons are beings who have not only interests, but also the capacities to understand and act on reasons. They are then able to deliberate critically about and order interests according to their importance, act rationally to pursue them, and do so normally compatible with relevant moral requirements. It is then appropriate to moral justification not simply that we judiciously take people's interests into account, but that we respect their capacities for practical reasoning about their duties and interests. This requires that we be in a position to cite reasons to others affected by our actions that they, in the relevant capacity (as reasonable persons or as reasonable democratic citizens), can accept. Since all are constrained and affected by demands of justice, respect for persons as equals requires that a moral conception of justice provide an agreed basis for public justification and agreement.[33]

The foregoing discussion of the first feature of Kantian constructivism, the social role of a conception of justice in public justification, goes some way toward clarifying the second feature, the publicity condition on principles of justice. As Rawls already says in *A Theory of Justice*, one primary reason for publicity is that it connects 'with the purpose of moral principles' in establishing possible bases of justification and agreement where none seem to exist.[34] Here it is crucial that public justification and publicity are not the same. Principles that

13

structure social relations may be public and still not provide a public basis for justification when people cannot endorse them (for example, because of the unfairness or oppressiveness of principles). These situations are unstable in the sense that disruptions of the status quo (for example, by sustained economic recession or population growth of discontented classes) can lead to social unrest, rebellion, and (if all goes well enough) the eventual adoption of different principles. Likewise, the fact that people have a shared basis for public justification does not require that the ultimate principles that structure their relations be publicly known to them. One example is Marx's account of ideological consciousness. (Plato's Noble Lie or the Grand Inquisitor's religion of unquestioned authority are further examples.) An ideology provides a shared basis for public justification which obscures from people the real grounds of their social relations. If they were in a position to recognize the principles underlying social and political institutions, social stability would be undermined by the many deep-seated conflicts or injustices that are hidden from awareness. A fitting judgment for ideological public justifications is that, however stable or generally beneficial they might render society, they do not exhibit respect for persons as free, equal, and capable of taking responsibility for their lives and relationships. Social stability and the general welfare instead depend upon citizens' illusions.

One reason for the publicity condition on basic principles of social cooperation is to insure that a society's basis for public justification is not ideologically grounded. Rawls says:

> [Publicity] can be explained as insuring that the process of justification can be perfectly carried through (in the limit so to speak) without untoward consequences. For publicity allows that all can justify their conduct to everyone else (where their conduct is justifiable) without self-defeating or other disturbing consequences.[35]

I take it Rawls means (by 'self-defeating or other disturbing consequences') that the publicity of principles is appropriate for a conception of justice since it enables people to justify their institutions and conduct to one another according to the *real* reasons that structure their relations, and to do so without destabilizing the social system.[36] There is no instability from publicity in a well-ordered society, since the public recognition that society satisfies principles all accept itself fosters citizens' sense of justice, which is the main ground for political stability.[37] Moreover, having a shared public basis in light of which all can justify their conduct and common institutions (on non-ideological grounds true to their conduct and institutions) exhibits respect for persons, and encourages ties of community.[38]

One basis then for Rawls's publicity condition is that it satisfies the practical aim of moral philosophy: that principles of justice have a social role of providing a public basis for justification and agreement on grounds that insure a stable basis for cooperation on terms of mutual respect. This 'contractualist' strand (so-called because of its resemblance to Scanlon's contractualism) is a central aspect of

14

Kantian constructivism (and of Kant's moral constructivism on Rawls's earlier account).

Now turn to a second strand of thought that anchors Rawls's publicity condition. In *A Theory of Justice*, Rawls introduces this condition, saying that 'publicity . . . arises naturally from a contractarian standpoint'.[39] The thought here seems to be that, if we all agree to something, then presumably we explicitly recognize that we all accept it. But everyone's agreeing to principles in Rawls's original position does not have to mean that all will publicly recognize what is agreed to: some people might forget or their representatives in the original position might choose that principles not be public. Here Rawls invokes the Kantian interpretation of the social contract to justify publicity. On the Kantian interpretation, the original position is viewed as 'a procedural interpretation of Kant's conception of autonomy and the categorical imperative'.[40] In addition, Rawls says:

> The publicity condition is clearly implicit in Kant's doctrine of the categorical imperative insofar as it requires us to act in accordance with principles that one would be willing as a rational being to enact as law for a kingdom of ends . . . an ethical commonwealth, as it were, which has such moral principles for its public charter.[41]

It is the Kantian interpretation of the contract situation, if not the idea of a contract per se, that accounts for the publicity condition on this interpretation.

In *A Theory of Justice*, Rawls is not altogether clear about how thorough publicity is to be; he refers only to the publicity of principles of justice. Later, in 'Kantian Constructivism in Moral Theory', Rawls distinguishes three levels of publicity: first, the publicity of principles of justice; second, the publicity of the general beliefs in the light of which first principles of justice can be accepted ('that is, the theory of human nature and of social institutions generally'[42]); and, third, the publicity of the complete justification of the public conception of justice as it would be presented on its own terms. All three levels, Rawls contends, are exemplified in a well-ordered society. This is the 'full publicity condition'.

The reasons Rawls cites for full publicity are revealing; they invoke the Kantian interpretation. Full publicity, Rawls says, 'ensures . . . that free and equal persons are in a position to know and to accept the background social influences that shape their conception of themselves as persons, as well as their character and conception of their good. Being in this position is a precondition of freedom; it means that nothing is or need be hidden.'[43] Later Rawls says, 'the realization of the full publicity condition provides the social milieu within which the notion of full autonomy can be understood and within which its ideal of the person can elicit an effective desire to be that kind of person'.[44]

On the Kantian interpretation, full publicity is called for because of its 'educative role' in leading citizens to self-knowledge of their potential full (moral and rational) autonomy and to the realization of their essential good as free and equal moral persons. Full publicity then fosters the full autonomy of citizens and there-

15

with the congruence of justice with their rational conceptions of the good. Here it bears emphasizing that Rawls's reasons for full publicity do not depend on the social role of a conception of justice in providing a public basis for justification and agreement. His focus, rather, is on the educative role of full publicity in promoting the free moral agency and full autonomy of equal citizens.

To sum up, publicity of principles of justice serves two roles in Kantian constructivism which parallel two of that doctrine's essential features. First, it makes possible the social role of a conception of justice by providing a public basis for justification and agreement that all reasonable persons can accept. The public availability of the conception of justice that society realizes is supposed to allow the public justification of social institutions to be completely carried through without destabilizing consequences. This is necessary for mutual respect among free and equal persons.[45] Second, publicity, when full, educates citizens to the ideal of the person (as free, equal, reasonable, and rational) underlying society's conception of justice, thereby enabling citizens to form an effective desire to be this sort of person, and so realize citizens' essential good as morally autonomous beings. Full publicity then makes *congruence* possible, which is the primary basis for stability in *A Theory of Justice*.

At this point the tension becomes apparent between two aspects of Kantian constructivism: (1) the construction of moral principles from a conception of the person as morally autonomous and (2) the public social role of principles. Already in 'Kantian Constructivism in Moral Theory' Rawls says of the 'public role that a conception of justice has in a well-ordered society' that, 'Because its principles are to serve as a shared point of view among citizens with opposing religious, philosophical and moral convictions, as well as diverse conceptions of the good, this point of view needs to be appropriately impartial among those differences.' So, the kinds of reasons that can be invoked to justify and apply the conception of justice must be limited to 'part of the truth, and not the whole', more specifically, to that part implicit in 'our present commonly based and shared beliefs'. 'Justice as fairness tries to construct a conception of justice that takes deep and unresolvable differences on matters of fundamental significance as a permanent condition of human life.'[46]

But surely, one might reply, there will be deep and unresolvable differences in a well-ordered society of justice as fairness about the nature of morality and moral justification, and even about the justification and truth of the shared, public conception of justice that regulates society? So, too, we might expect that there are fundamental differences in people's beliefs regarding the intrinsic good of autonomy and about whether justice is the supreme human good.[47] Because of the 'burdens of judgment' (discussed below) and the diversity of philosophical, religious, and moral views in a well-ordered society of justice as fairness, many reasonable and conscientious people who accept Rawls's principles of justice are unable to accept their congruence with their rational good on grounds of autonomy. But this undercuts the argument for the stability of a well-ordered society.

16

So we have the peculiar situation that there will be reasonable and rational citizens[48] in a well-ordered society of justice as fairness who can endorse the public principles of justice and the institutions they authorize, yet who, because of their conception of the good, cannot accept its public justification. Significantly, this problem is not peculiar to those with religious beliefs. It is typical of anyone (perhaps even a majority) who accepts liberal principles of justice on philosophical and ethical grounds that conflict with the Kantian interpretation of justice as fairness. The idea of a well-ordered liberal society in which everyone accepts the same liberal principles of justice does not guarantee an absence of significant numbers who might accept these liberal principles on grounds of utilitarian, perfectionist, pluralist, or even morally skeptical views. Indeed, given such basic liberties as freedom of thought, conscience, and association, it is predictable that these and other philosophical doctrines will gain adherents. But given the requirement of full publicity, this means that the public conception of justice incorporates conclusions starkly at odds with some citizens' most deeply held conscientious beliefs. As a result, the full publicity of the shared conception of justice, far from insuring a shared basis for justification and stabilizing the social system, is likely to have destabilizing consequences. It can adversely affect the sense of self-respect of those who hold nonconforming (non-Kantian) philosophical and ethical views, and undermine their allegiance and support for just institutions.

It seems then that Kantian constructivism, by aiming to incorporate the Kantian interpretation into the public justification of justice as fairness, fixes on a goal that cannot be realized. Two main features of Kantian constructivism (and of Kant's moral constructivism as Rawls defined it[49]) are at odds: (1) the conception of the person (as reasonable, rational, free, and equal) as morally autonomous and the original source of moral principles of justice and (2) the public social role of principles of justice. These two central features of constructivism cannot be combined into one view without untoward consequences, therewith destabilizing the conception of justice.

Does this mean that constructivism is unworkable as a moral or political theory providing a public basis for justification and agreement? It may be that Kantian constructivism is unworkable. My sense is that Rawls came to think so, for he no longer alluded to Kantian constructivism, and appeared to regard it as a transitory stage in the development of his thought. But not all forms of constructivism are impracticable so long as certain things are relinquished. In political constructivism, Rawls relinquishes, first, the Kantian congruence argument for stability: that moral autonomy, hence justice as fairness, is a rational and intrinsic good for each person in a well-ordered society. Congruence on these grounds seems to be unacceptable to, if not untrue of, many reasonable persons in a well-ordered society of justice as fairness; this includes not just liberal Catholics and other liberal faiths, but liberal utilitarians, perfectionists, pluralists, and many others.

Second, Rawls gives up the part of Kantian constructivism that holds (1) that

17

there is no moral order apart from the activity of practical reason[50] and (2) that moral objectivity and the validity of judgments are to be understood only in terms of correct judgment from a suitably constructed social point of view.[51] As part of the public conception of justice, these philosophical claims, even if true, cannot gain widespread agreement since they conflict with other not unreasonable views about the nature of morality, and moral objectivity and validity.

Third, the Kantian conception of the person (as free, equal, reasonable, and rational) is retained, but not without modification. It is no longer presented as part of a 'partially comprehensive' moral doctrine that grounds morality in the conditions of moral agency and the powers of practical reason. Rather, it forms the basis for a 'freestanding' (or 'self-standing') political conception of justice that 'brackets' philosophical questions of the original source of moral principles, their objectivity, and how we ultimately can know them. This conception of the person is now said to be implicit in our awareness of ourselves, not as moral agents and in whatever we do, but in the more circumscribed role we occupy as democratic citizens. In thinking about ourselves politically, we regard ourselves as free and equal citizens; as such, we rationally should have a higher-order interest in developing the capacities or 'moral powers' that enable us to engage in social cooperation.[52] In political constructivism, there is no longer the controversial claim that these powers constitute our 'nature' as moral beings or that they are necessary conditions of moral agency or that realizing these powers is a supreme human good. The moral powers, rather, are seen as essential to our gaining the benefits of social cooperation and pursuing our conception of a good life. Without them, we could not take advantage of the rights or comply with the duties of democratic citizens, or achieve our most important aims. It is then rational for democratic citizens who aim to cooperate with others, whatever their reasonable conceptions of the good, to take an interest in the development and exercise of their capacities for rationality and justice.

4. Political constructivism

The problem with Kantian constructivism is that it undermines general agreement on principles of justice by providing a public justification for these principles that is not generally acceptable in the face of the 'reasonable pluralism' of comprehensive moral, religious, and philosophical positions that would gain support in a well-ordered society of justice as fairness. Political constructivism seeks to preserve the general acceptability among reasonable persons of liberal egalitarian principles of justice. It does so by arguing that for 'public political purposes' democratic citizens are to conceive of principles of justice as 'constructed' from fundamental ideas implicit in democratic culture. It 'represents' these fundamental ideas within a 'procedure of construction' (the original position) and thereby shows how the principles derived from this procedure follow from basic features of democratic reasoning (including shared norms of practical

18

reason, together with a conception of citizens as free and equal with the moral powers necessary for social cooperation).

First, this representation is *possible* since (1), unlike the Kantian interpretation and Kantian constructivism, political constructivism does not endorse 'constitutive autonomy' and make the philosophical claim that moral principles of justice *originate* in the capacities or activities of practical reason with no other source. The political conception brackets questions of the nature of morality and the truth of moral principles, and is 'freestanding' of metaphysics. The most that Rawls claims for these principles is that they are 'reasonable', insofar as they satisfy 'all the relevant requirements of practical reason'[53], and stand in reflective equilibrium with our considered moral convictions of justice. These, presumably, are requirements that any moral conception must meet if it is to meet more exacting criteria of moral truth or validity.[54] Here it is noteworthy that in *The Law of Peoples* Rawls further deflates his claims regarding the reasonableness of principles. He says he does not mean to rely on any Kantian assumptions in referring to 'the requirements of practical reason',[55] and is not trying to deduce principles of right and justice from a conception of practical reason. 'Practical reason as such is simply reasoning about what to do, or reasoning about what institutions or policies are reasonable, decent, rational, and why.'[56] Reasonable principles, it now appears, are those that would be affirmed by reasonable persons upon due reflection and (apparently) in reflective equilibrium with democratic ideas and their considered moral convictions. (2) Political constructivism also brackets assumptions regarding the intrinsic good and is 'freestanding' of any comprehensive ethical claims or doctrines. The stability of Kantian constructivism depends on a sufficient number of people endorsing the *same* comprehensive doctrine, and willing justice for the right reasons, namely, to realize their nature as morally autonomous. By contrast, the stability of political constructivism depends on the much weaker requirement that people endorse and will liberal justice for the particular (and often conflicting) reasons implicit in the reasonable comprehensive doctrines they hold. An 'overlapping consensus' among these doctrines (their convergence on the same conception of justice) is, Rawls argues, now sufficient for the stability of a well-ordered society.

Second, political constructivism's representation is also *suitable* for a democracy because (1) it enables members of a democratic society, for public political purposes, to conceive of their political relations as the product of their free, joint political deliberations and agreement. Their political relations are not publicly regarded as imposed by external authority, their natural inclinations or circumstances, or any source other than the reasons they employ as democratic citizens and their conceptions of themselves as free and equal citizens. This is 'doctrinal autonomy'. Democratic justice is publicly conceived as the product of citizens' political (not their moral) autonomy.[57] Moreover, (2) the public conception of justice has its origin in ideas that are part of democratic political culture, ideas that are normally relied on by citizens in the course of public debate and argu-

ment. The public conception originates then in *public political reason*, in practical reason as applied to the central ideas of a democratic culture.

Third, and finally, political constructivism's representation is *necessary* (so Rawls seems to say) for purposes of public political justification, because any other source publicly assigned to moral principles of justice (1) will not fit with the fundamental ideas of the political culture or (2) invokes metaphysical or comprehensive moral claims regarding the origins of democratic justice (such as 'Our democratic Constitution is the product of God's will' or 'Democratic justice originates in the individuality or moral autonomy of its citizens') or (3) involves some other controversial claim that is not publicly acceptable among reasonable citizens (for example, 'Democracy is sheer historical accident and is morally arbitrary').

So Rawls claims:

> Thus, it is *only* by affirming a constructivist conception – one which is political and not metaphysical – that citizens generally can expect to find principles that all can accept. This they can do without denying the deeper aspects of their reasonable comprehensive doctrines. Given their differences, citizens cannot fulfill in *any other way* their conception-dependent desire to have a shared political life on terms acceptable to others as free and equal.[58]

What is striking about this claim is how much it rules out. It means that any moral conception that seeks to provide a public basis for justification in a liberal democracy must severely limit the kinds of reasons it invokes in justifying and applying principles of justice. It cannot make metaphysical, epistemological, or even moral claims that are not endorsable by widely different reasonable doctrines, but must restrict itself to 'the political values of public reason' and an epistemology of common sense. Liberalism, Rawls now believes, can tolerate several nonpublic justifications of its principles. But if liberal principles are to be generally shared, uniformly applied, and are to serve as a basis for public reasoning, they must also be grounded in a shared justification that is freestanding of, but still compatible with, the wide range of reasonable nonpublic views. This follows from a moral requirement of respect for persons, now expressed through Rawls's criteria of reciprocity and liberal principle of legitimacy;[59] it is also required for political stability.

To sum up, Rawls contends that respect for persons as free and equal citizens requires not just (1) that restrictions on their conduct have a public justification on terms that they can accept in their capacity as citizens (a shared basis for public justification), but also (2) that the public justification itself makes fully public the conception of justice upon which social and political cooperation is based (full publicity). Full publicity then requires (3) that there must be in the public political domain itself a complete conception of justice, including a justification of its principles in terms of reasons that all can endorse in their capacity as citizens. This public conception must then be 'freestanding' and based in

20

public reasons that exclude 'part of the truth'. So whatever the ultimate origins of justice might be (whether in moral autonomy, aggregate utility, God's commands, and so on), they cannot be incorporated into the public justification of matters of political justice.

5. Kant and full publicity

In the second half of this article, I consider some general consequences of Rawls's claim that only political constructivism can provide principles that reasonable persons can accept. What bearing does this have for two similarly motivated constructivist views: Kant's moral constructivism (Section 5) and Scanlon's moral contractualism (Sections 6–8)? Can either provide a basis for public justification that satisfies the full publicity condition in societies that satisfy each position's moral requirements?

Kant thought the validity of laws of justice (and probably other moral rules) were subject to 'the capacity for publicity'.[60] For laws to be just, they must be publicly acceptable among fully reasonable and rational beings. Kant apparently assumed the reasons for laws must be publicly acceptable too. Not just any reason is publicly acceptable in Kant's view. People frequently accept and agree on laws because they meet their conflicting personal or group interests; such compromises are frequently struck in democratic legislation. Kant rejected such a 'coalition of private wills' as contrary to the 'public will' represented through the 'Original Contract'. Behind this is the thought that to be publicly acceptable in Kant's sense, a law must be endorsed not by reason in its private use, but by 'the free public use of reason'. For Kant, this means reason in its full critical use, not constrained by authority, custom, or one's social role or position. It is the complete freedom 'to impart to the public all one's carefully considered and well-intentioned thoughts'[61] and 'an unrestricted freedom to use one's own rational capacities and to speak one's own mind'.[62, 63]

This unrestricted freedom critically to exercise our rational capacities, and express in the political forum our enlightened understanding on all (politically relevant) issues, underpins Kant's account of the free use of public reason. Suppose we see the free use of public reason as a condition upon the universal acceptance of principles or valid laws. Then Kant's standard for the justification of laws, and perhaps even moral norms, would be their universal acceptability for the *same* reasons, following critical public deliberation on the basis of *all relevant reasons*. Call this the 'broad view' of public reason. (Habermas's discourse ethic seems to endorse this broad view.)

This broad view is not Rawls's account of public (political) reason. For purposes of publicly justifying and applying political principles and laws, Rawls restricts the domain of relevant reasons to shared political ideas and values implicit in democratic culture. Rawls's restricted view of public reason stems from his account of the 'burdens of judgment':[64] people differ in their capacities,

21

experiences, upbringing, knowledge, circumstances, and so on. Moreover, crucial concepts are vague, and evidence is often conflicting. For these and other reasons, Rawls sees it as inevitable that, in a free democratic society, even under the ideal circumstances of a well-ordered society, there are always going to be differences among reasonable and rational persons in their philosophical and religious beliefs and in their ethical and even moral convictions. No amount of discussion or ideal discourse can eliminate these disparities.[65] Public reason is simply not amenable to wide-ranging philosophical or ethical agreement any more than it is to religious agreement. The 'free use of public reason' in Kant's sense, rather than providing a basis for 'public reason and justification' in Rawls's sense, is one of the deepest bases for reasonable disagreement, inevitably resulting in reasonable pluralism. For this reason, to require of principles of justice, or any other moral principles, that they satisfy Kant's broad view of public reason (to be acceptable to all for the same reasons on the basis of all relevant reasons) will not result in an agreement on moral principles that satisfies the full publicity condition, even under ideal conditions. If moral principles are ever to be universally acceptable, inevitably they will be finally endorsed by reasonable persons for different reasons.

There are two ways to interpret the implications of the burdens of judgment and reasonable pluralism for Kant's moral constructivism. First, it might be argued that Kant's view incorporates conditions that cannot be realized.[66] The only way to satisfy Kant's requirement (that to be justified, laws [or moral rules] must be publicly acceptable to all reasonable persons on the basis of reasons all can accept) is by severely circumscribing the scope of public reason and justification to the domain of the political. But that would make Kant's view a kind of political constructivism, which would compromise the regulative role of autonomy in moral and political life (surely not Kant's original intention).

Alternatively, one might reply, on Kant's behalf, that the fact of reasonable pluralism and the burdens of judgment do not accurately characterize a realm of ends and Kant's test of publicity only applies under those ideal conditions. By definition, reasonable and rational beings in a realm of ends all endorse and will moral principles for the same reasons: out of a good will and for the sake of moral autonomy. The question then arises of whether this is a realistic assumption, and, if not, whether this has any bearing on the justification of moral principles in Kant's view.

Must a realm of ends (a world in which public moral principles are universally endorsed by all for the same moral reasons [of autonomy]) be a realistic possibility in order for the universal principles that would be willed there to be justified? Alternatively, is the realm of ends a hypothetical construct that tests the validity of moral principles, regardless of their feasibility under the conditions of human life? I have no definite answer. But one way to understand Kant's own dictum that 'ought implies can' is that moral principles cannot apply to us unless they are compatible with human nature and describe a feasible social world. Can we

22

then have a duty to act only on those maxims which we can will as universal public law for and in a realm of ends if this hypothetical situation is not realizable on earth? Perhaps so. Less clearly, does it make sense to say with Kant that we also have a duty to *promote* and seek to realize a realm of ends if this ideal situation is not realizable given the enduring circumstances of social life? On the aforementioned interpretation of 'ought implies can' (namely, that a moral conception describe a feasible social world), it is difficult to see how we could have a duty to realize such an unrealizable situation.[67] Admittedly, this interpretation may be too strict for the purposes of a moral conception of justice, as opposed to a political conception of justice. I will consider this in connection with contractualism.

6. Contractualism and the publicity condition

Return to the claim that only political constructivism can provide principles reasonably acceptable to all reasonable persons. What implications does this have, if true, for T.M. Scanlon's contractualism? Insofar as the aim of contractualism is to give a purely philosophical account of the subject matter of morality, the possibility of moral validity and objectivity, and the possibility of moral knowledge and the nature of moral reasons, there appears to be no problem. Regarded as an account of these theoretical issues, contractualism does not aim to provide a basis for public justification. It is addressed primarily to philosophers and others interested in philosophical issues.

But what about contractualism as a moral conception that aims to provide reasonable persons with a framework for reasoning about principles of right and wrong? Contractualism as a moral conception differs from both Kant's moral constructivism and Rawls's Kantian constructivism in that it does not advance a regulative ideal of autonomy as essential to the human good. Nor does moral contractualism offer a decision procedure for deciding what is right to do in particular situations. It more modestly provides an account of the moral reasons that all reasonable persons have for avoiding wrong and doing right, and a mode of deliberation (reasoning in terms of the principles that could not be reasonably rejected) through which these reasons are applied to determine the moral rules for everyday life.[68] Unlike Rawls's and Kant's constructivist procedures, contractualism provides no method of application or priority rules for moral rules and principles. (In this regard, it resembles what Rawls calls 'intuitionism'.) Is moral contractualism characterized in this more modest way realistically possible as a public basis for justification among reasonable individuals?

For a moral conception to be realistically possible in this sense requires more than simply metaphysical possibility (which any moral conception can satisfy). I understand the feasibility and stability of a moral conception as follows:

1. The motives for behaving morally, as interpreted by a moral conception, are not only something we are capable of, but are normal dispositions of reason-

able persons in any well-ordered society that realizes the principles of that moral conception.

2. This moral motivation is generally compatible with reasonable persons' good and is understood as such by all who comprehend and endorse the moral conception. Thus, in acting from their desire to do what is right, moral agents do not defeat their most deeply held purposes or distort their characters.

3. There are no insurmountable coordination problems or social dilemmas in the interactions of sincere and conscientious moral agents acting from moral principles.

4. In a well-ordered society that conforms to the rules and institutions required by the moral conception, reasonable agents accept the public moral conception as an accurate account of morality and the real source of their moral rights and duties, and deliberate and justify their actions by reference to its terms.

5. Public knowledge of the moral conception and its full justification does not have the untoward consequences of undermining moral agents' self-respect, their confidence in their self-worth, and confidence in the worth of their permissible life plans.

Some of these conditions are achievable by a contractualist moral conception. As item 1 requires, the contractualist motivation (the willingness to be able to justify one's actions and institutions on grounds others who are similarly motivated could not reasonably reject) is plausible as an account of moral motivation in a well-ordered contractualist society.[69] Moreover, regarding item 2, this motive is likely compatible with reasonable persons' conceptions of their good since, being reasonable, they do want to justify themselves to others and can be expected to adopt a conception of the good that has an important place for this motive. Regarding item 3, contractualism does not seem to imply any social dilemmas or coordination problems that cannot be dealt with on its terms. Problems arise, however, with items 4 and 5, which assume the (full) publicity of moral contractualism. Is it likely in a well-ordered contractualist society, wherein contractualist moral rules are generally accepted, that reasonable people generally also will accept contractualism itself as an adequate account of the nature and origin of their moral duties and, moreover, will endorse contractualist moral principles *because* these principles could not reasonably be rejected as a basis for informed, unforced agreement? Alternatively, do we have the same problems Rawls encountered with reasonable people (Kantians and intuitionists, utilitarians and perfectionists, people of religious faith, cultural relativists, and others) who have different views of the subject matter of morality and the origin of their moral duties and who endorse shared moral principles ultimately for very different reasons?

It may be that these feasibility conditions, while appropriate for a political conception of justice, are too strong for the purposes of a general moral conception. To begin with, political conceptions, while also moral, sanction the use of co-

ercive power to enforce norms of justice. So there is a special need that political principles be publicly justifiable to free and equal persons according to reasons all reasonable persons can accept, both for moral reasons of fairness and mutual respect and for pragmatic reasons of political stability. There is some suggestion of this line of thought in later Rawls.[70] Still this does not decide the issue, since the application of coercive sanctions is just *one* of the reasons Rawls gives for full publicity.

Another reason full publicity may not be appropriate for contractualism is that moral conceptions may have different aims than purely political conceptions of justice. Utilitarians, for example, would say they have no overriding ambition to provide a public basis for justification and deliberation even under the ideal circumstances of a well-ordered utilitarian society. They aspire rather (on one interpretation) to universal principles that satisfy conditions of systematicity, and which are capable of providing a rational moral choice under all possible conditions. That a morality be completely rational in this way is seen by some utilitarians as a condition of moral truth. If, as Sidgwick maintained, none but the 'enlightened few' are able fully to appreciate utilitarianism and its bases, then the best way to realize its requirements may be to maintain it as a nonpublic 'esoteric morality'.[71]

Might contractualism countenance a similar position and reject the publicity of contractualist reasons as a condition of justification of its demands? I admit to puzzlement about this issue. In the remainder of this section, I will state some reasons why contractualism might seem committed to its own publicity. If nothing else, this should help clarify what is meant by contractualism's requirements of reasonable justifiability, particularly the question, justifiable *to whom*?

One aim of full publicity is to insure that the moral considerations that structure reasonable persons' social and political relations are available to them as principles of practical reasoning and justification. For contractualism (unlike utilitarianism) the aim of providing reasonable persons a public basis for justification among themselves does not seem incidental. Contractualism represents morality as a 'system of co-deliberation'.[72] Co-deliberation among whom? In the first instance, contractualism involves hypothetical co-deliberation among artificial agents who are morally motivated, adequately and accurately informed, and who aim to agree and act on principles no one similarly motivated 'could reasonably reject as a basis for informed, unforced general agreement'.[73] This phrase suggests a second level of co-deliberation and agreement, namely, that involved in reasoning in terms of contractualist principles themselves, to arrive at general agreement on their requirements. Once again the question is, who are the parties to this second 'informed, unforced general agreement'? *To whom* are contractualist justifications of principles to be directed? One might say here that this second level of co-deliberation aspires to justifications to and general agreement among the fully informed hypothetical parties who would agree to moral principles in the first place. But they are just artificial deliberative agents and not realistically possible

25

people. It may indeed be a truism that whatever principles fully informed morally motivated hypothetical moral agents can agree to will be principles that *they* could not reasonably reject as a basis for their *own* agreement. Moreover, it would seem true of most any moral conception that hypothetical reasonable persons with full information can be defined in such a way as to agree to that moral conception's principles (for example, the principle of utility) as a basis for *their* reasonable agreement, and that when the time comes to apply the principle they, as fully informed and morally motivated utilitarian agents, will accept its application to themselves. Why otherwise would they agree to the principle in the first place? In that event, agreement itself would seem to add little if anything to the justification of the principles already agreed to. It is this reading of contractualism that leads some philosophers to think that contractualism can be used to justify utilitarianism, or that it can be used to argue for most anything. One way to escape these objections is for contractualism to say that the principles that hypothetical co-deliberators deliberate about are not for artificial people like themselves, but for the general agreement of historically situated moral agents who are reasonable and are morally motivated to justify their actions to one another on terms no one similarly motivated could reasonably reject.

To interpret contractualism in this way means that to be justified, moral principles and contractualist reasons for them must be capable of providing a basis for co-deliberation for, justification to, and agreement among historically situated reasonable persons in an ideal, but nonetheless feasible, social world. (Ideal because each is assumed to be reasonable and morally motivated to act on moral principles for their own sake; feasible because such a world takes into account human nature.) These ideal, but still feasible, conditions are the circumstances that fully informed artificial contractualist agents deliberate about, and which the moral principles they agree upon are designed to apply to in the first instance. Hypothetical contractualist agents with no false beliefs are co-deliberating about principles for a public system of co-deliberation among reasonable persons with diverse conceptions of the good, like themselves, but who have a human nature and who live under real-world conditions (in which people do not have, for example, the full and accurate information that artificial agents do).

Several considerations seem to speak in favor of this reading of contractualism. First, Scanlon himself, in his initial paper on contractualism, 'Contractualism and Utilitarianism', criticizes Rawls's original position on grounds that the choice of rational self-interested agents behind a veil of ignorance misrepresents what is involved in a person's acceptance of principles on grounds that no one could reasonably reject whatever position he or she turns out to occupy *in society*. Moreover, Scanlon contends that rational choice from the original position is not even necessary to Rawls's contractualist argument, since Rawls already makes a number of arguments (from the strains of commitment, stability and publicity, and self-respect)[74] for his principles from the perspective of what would be acceptable to reasonable individuals in a well-ordered society:

> The [contractualist] connection with self-respect, and with the Kantian formula, is preserved by the requirement that principles of justice be ones which *no member of the [well-ordered] society* could reasonably reject. This connection is weakened when we shift to the idea of a choice which advances the interests of a single rational individual for whom the various individual lives in a society are just so many different possibilities.[75]

Second, in Rawls, the idea of justification *to* a person is defined by reference to the 'practical task of political philosophy', which in turn consists in its 'social role' or 'public role' in providing a basis for public justification. It is the idea of public justification to members of a well-ordered society that provides the framework for Rawls's idea of justification to a person:

> The social role of a conception of justice is to enable all members of society to make mutually acceptable to one another their shared institutions and basic arrangements, by citing what are publicly recognized as sufficient reasons, as identified by that conception. To succeed in doing this, a conception must specify admissible social institutions . . . so that they can be *justified to all citizens*, whatever their social position or more particular interests.[76]

This suggests that the idea of justification to a person is defined by reference to the idea of a well-ordered society. It is because people inevitably have different conceptions of the good and different comprehensive views that contractarians see a need for the idea of 'justification to a person'. This idea becomes something quite different taken out of the context of a well-ordered society and applied to artificial people with full knowledge and an absence of conflicting moral beliefs. Why is there any longer a need for the idea of 'justification to' persons if we are no longer driven by the practical task of a moral conception to provide a basis for public justification, itself a response to real-world differences in conceptions of the good and moral, philosophical, and religious beliefs due to the 'subjective circumstances of justice'?

Third, for contractualism to reject publicity as a condition of the validity of moral norms has the consequence that first moral principles need not serve as principles of practical reasoning and public justification in any feasible social world. It may seem peculiar for contractualism to accept this, given that its primary focus is not individual welfare, but the justifiability of acts and institutions to reasonable persons out of respect for them and their capacities for practical reasoning and judgment. Here one might contend that contractualism indeed can serve as a basis for public justification among reasonable people who accept contractualism. But this resembles Sidgwick's position that a moral principle need only be public among the 'enlightened few'. Any moral conception can say that artificial or real-world agents who already accept that very conception can justify themselves to one another on terms others accepting the same moral conception can reasonably accept. To understand contractualism this way does not

27

make clear how it differs from other moral conceptions and how it is distinctly about justification *to* others.

Finally, as Scanlon contends, the desire to justify one's actions and institutions is quite strong in people, and they will go to great lengths involving heavy sacrifices to avoid admitting their unjustifiability.[77] One might reply, 'Of course, but this fact can be accommodated by most any moral view.' So, for a committed utilitarian to want to be justified might be described as their wanting to act on principles that would be approved by an ideal impartial observer. Also, a Thomist will want to act in ways that are justifiable according to divinely ordained natural laws. But contractualism says that part of the willingness to justify our actions and institutions is to be willing to give reasons *to others* which they, in some capacity, could not reasonably reject. What is distinctive about contractualism is that *other reasonable persons' acceptance*, in light of their primary aims and ambitions, of our actions and institutions is incorporated into contractualism's account of justification and moral motivation. Moral justification is regarded as *justification to* others, conceived as *reasonable persons* with different aims and interests and conceptions of their good. Hence, our desire to justify ourselves is the desire to be able to give reasons to others that *they* could not reasonably reject.[78] Now, in wanting to justify ourselves to others our aim does not seem to be to justify ourselves to artificial contractualist agents (or to ideal observers or any other philosophical construct). We want to justify ourselves to the real people our actions and institutions affect, on the assumption that they are reasonable and rational, are open to understanding and appreciating moral reasons, and are capable of endorsing the justification of our actions when they are justifiable. It is true that in order to justify ourselves to others, contractualism says that we are to cite the reasons and principles that it would be unreasonable for fully informed artificial contractualist agents to reject. But we are not justifying ourselves to these artificial agents, but to people whom our actions affect, on the assumption that they are reasonable and reasonably well informed.

Now for our audience to be in a position to accept these kinds of considerations as reasons, and for us to achieve our aim of justifying ourselves to others, our audience ultimately has to accept contractualist reasons as providing adequate justification for moral principles. If they reject contractualism and the reasons it provides (suppose they are committed utilitarians or endorse a doctrine of divinely ordained natural law), then the reasons we give will not be recognized and accepted as sufficient justification. So to justify ourselves to others on grounds they could not reasonably reject would seem to require the publicity of contractualism so that its reasons may serve as the shared basis for justification and agreement.[79]

I have interpreted contractualism as a system of co-deliberation among hypothetical persons who reason about the moral principles that will serve as a public basis for justification and agreement among reasonable persons like themselves who are in ideal, but feasible, real-world conditions. The question is whether con-

28

tractualism can meet the full publicity condition? Is there a realistically possible world in which reasonable persons accept contractualist moral principles for the reasons contractualism provides (because these principles could not be reasonably rejected among fully informed, morally motivated, hypothetical co-deliberators)? I assume that reasonable pluralism (the endorsement of different reasonable philosophical, ethical, and religious doctrines) is characteristic of any society that satisfies the demands of a contractualist conception of justice. Is there a realistic social world in which all reasonable persons accept not just the same moral norms, but also accept contractualist reasons[80] as sufficient justification of their moral duties, whatever their reasonable conception of the good may be? Assuming the burdens of judgment apply under even the best of circumstances, it would have to be a social world in which different comprehensive doctrines with different metaphysical and evaluative convictions nonetheless agree in their accounts of the ultimate (contractualist) reasons for moral duties. But that rules out not only most religions' accounts of morality, but also all philosophical accounts that are non-contractualist. It does not seem at all likely that reasonable persons, freely exercising their capacities for practical reasoning under optimal social conditions, will converge on contractualism (or any other doctrine) as a (partially) comprehensive moral view.

It may be that the appropriate reply by contractualism to this problem is to deny that the feasibility of general agreement on a moral conception with full publicity (as defined) is a condition of the justification of a moral conception. For aside from contractualist norms of justice (discussed below), other contractualist norms do not involve the use of coercive force, and there is not the same need as there is for the stability of a conception of justice. Why is it such a problem for the justification of contractualism that some reasonable persons in the world are unable to accept a contractualist justification for non-coercive contractualist norms?

Consider the two versions of this reply just mentioned. (1) Some reasonable people are always going to have false or misguided beliefs about the foundations of morality and the nature of moral reasons, but this is not a problem. So long as a social world is practically possible in which reasonable people accept and endorse contractualist moral norms, what does it matter to the justification of these norms that many endorse them for non-contractualist reasons? An overlapping consensus on contractualist norms is sufficient. (2) Some reasonable people are always going to have false beliefs about moral norms themselves, but so what? Even if a social world is not feasible in which reasonable people generally accept contractualist norms, why should this matter to the validity of these norms?

The more optimistic reply, (1), seems plausible. A characteristic feature of reasonable pluralist societies is that different religious and ethical views, while they differ on some moral questions (on distributive justice or the morality of abortion and euthanasia, for example), still accept many of the same moral rules

29

and principles needed for social life (about honesty, fidelity to promises and commitments, respect for others' persons and possessions, and so on). Suppose a reasonable pluralist society in which reasonable people all endorse contractualist moral rules and principles (including a liberal political conception of justice) on the basis of different ethical, philosophical, and religious views. How likely is such an overlapping consensus on contractualist norms to endure when members of society lack a shared basis for deliberation and argument about how to justify and apply these norms, or how to understand their scope and limits?[81] Even assuming an initial convergence among reasonable doctrines, the application of shared contractualist norms by many non-contractualist means is no doubt going to result in disagreement, which eventually may affect the evolution of moral norms among different groups and the range of norms agreed to. One worry is that, without a shared basis for moral discussion and general agreement, such a society is liable soon to become the society described by (2), one that lacks widespread agreement on contractualist norms themselves.

Response (2) (that the feasibility of a contractualist moral conception does not matter to its justification) raises complicated issues of a different sort which I cannot adequately address here.[82] In general, there is something to be said for the idea that the burdens of judgment, so far as they depict the inevitable human propensity to affirm different values and make different (and often erroneous) judgments, should not matter to the validity of moral norms, but should be taken into account in their application to real-world conditions.[83] Still, it seems that in considering the relative strengths of competing moral conceptions, it often does matter whether they are feasible as public moral conceptions, and that reasonable people be able to rely on these conceptions to deliberate about and justify their conduct to others. For in order reasonably to expect of people that they comply with morality's requirements, it is important that morality be compatible with human nature ('within human reach') and also conducive to, or at least not inconsistent with, the human good. In these regards, some of the most common objections to utilitarianism are that it is self-defeating as a generally accepted moral conception, and that a well-ordered utilitarian society in which everyone knows and accepts the principle of utility and its requirements is not feasible. This prospect led Sidgwick to dispense with publicity and advocate utilitarianism as an esoteric morality.[84]

But contractualism would not have the same problems as utilitarianism were it publicly recognized as the basis of moral relations. For what makes utilitarianism self-defeating is that it sacrifices the essential good of some individuals for the sake of greater aggregate utility; hence public knowledge of utilitarianism leads those who are especially disadvantaged to resent their sacrifices for the benefit of the more advantaged, and therewith reject utilitarian requirements. Contractualism's demands do not impose similar disadvantages on people. If public knowledge of contractualism would lead some to reject contractualist norms, it would be because they refuse the contractualist justification of norms which they accept

30

on other philosophical or religious grounds. This implies that a well-ordered contractualist society is unfeasible, but not that contractualism is unreasonable in the way that utilitarianism is.

7. Why contractualism's failure of full publicity is not a serious problem

I began this discussion with the question of whether there is a feasible social world in which contractualism can provide a basis for public justification among reasonable persons that meets the full publicity condition. We have seen that it cannot. I conclude with some suggestions why this should not pose a serious problem for contractualism.

While it may be that moral contractualism cannot impose a full publicity condition as a requirement on the reasonableness of moral principles generally, this does not mean that contractualism cannot endorse full publicity as a requirement for liberal principles of justice. For assuming political liberalism is correct, (1) a contractualist moral conception surely can and would require for its own contractualist reasons a liberal political conception of justice; moreover, (2) contractualism also can endorse the strictures of public reason in the public political domain. Within contractualism this means that it would be unreasonable to reject: (1) that to be correct (reasonable or true), a political conception of justice must be publicly knowable and affirmable among reasonable persons; (2) that this conception should provide 'content' to public reason and therewith a basis for public political justification; and (3) that the political conception of justice itself may not incorporate into the political domain exclusively contractualist reasons (or any other comprehensive reasons) regarding the nature and foundations of morality, but should only rely upon public reasons acceptable to reasonable democratic citizens.

This implies that, even if a well-ordered contractualist society may not be feasible, a society in which all aspire and seek to justify themselves in matters of *justice* to others on grounds that others cannot reasonably reject is *still* feasible. Such a world is feasible since not just contractualism, but all reasonable comprehensive doctrines, are in an overlapping consensus on the public liberal conception of justice, and thus can agree on the criteria of reciprocity, political legitimacy, and the strictures of public reason. Unlike contractualism, however, other reasonable doctrines accept political principles and requirements not because they regard them as intrinsically reasonable as contractualism does, but because these requirements are thought to further the foundational requirements of one or another reasonable view (for Kantians, acting from liberal political principles of justice realizes moral autonomy, while for utilitarians, these principles and requirements of public reason are believed to promote aggregate utility, and so on).

What we have then, in effect if not by design, is the real possibility of a shared

31

contractualist morality of justice without widespread contractualist convictions regarding the foundations of morality. In a well-ordered society described by Rawls's political liberalism, people act *as if* contractualism were the true comprehensive moral doctrine, at least with respect to justice, without many (most?) believing it to be so. This is not true of any other comprehensive moral view. For if you ask any non-contractualist why he or she ultimately believes that people ought to provide a public justification for constraints of justice that others can reasonably accept in their capacity as democratic citizens, they will cite the particular moral or religious or metaphysical reasons peculiar to their own comprehensive views. But the contractualist will be able to say that, fundamentally, this is just what morality is all about. This results from the fact that the contractualist account of the nature of interpersonal morality is basically the same standard as Rawls's criteria of reciprocity and principle of political legitimacy.[85]

Moreover, because of this coincidence of the criteria endorsed by contractualism and political liberalism, even moral duties that are not based in justice are likely to be publicly discussed using contractualist standards among people with different comprehensive views. For example, consider what will transpire in a dispute over the scope of a duty not to lie among people of different moral and religious persuasions (Kantians, Catholics, and utilitarians). They do not share in common the moral foundations or comprehensive reasons for resolving this dispute. But they already have at hand a shared political standard and are in the habit of resolving disputes of justice according to public reasons and rules that reasonable persons could not reasonably reject in their capacity as democratic citizens. Since they have no other shared basis for discussion and deliberation, it is only to be expected that they would gravitate toward and even appeal to the more general contractualist standard of reasonable rejectability of rules and reasons that no reasonable person *as such* could reasonably reject. But here, unlike political disputes of justice, they may disagree over what more is involved in being a reasonable person – each will insist that it involves endorsing their moral conception's particular interpretation of practical reason and its implications. Moreover, unlike the political case, appealing to the characteristics and interests of free and equal democratic citizens will not likely be adequate or even relevant to resolving many moral disputes. But at least the contractualist standard provides a shared basis for public discussion and argument that conscientious people already accept in political disputes of justice. While not endorsed as fundamental by other comprehensive views, the contractualist standard should be unobjectionable in other moral disputes and can at least provide a basis for argument and discussion (if not always agreement) among people of different comprehensive views. The important point is that by complying with the requirements of this standard in moral disputes in addition to those of justice, people respect and mutually recognize one another as reasonable persons. In these respects, contractualism can provide a public basis for moral, not just political, justification among reasonable persons with different comprehensive views, even if it cannot

be generally accepted by all as the foundation for interpersonal morality. It satisfies at least a partial, if not the full, publicity condition.

There is another reason that its inability to satisfy a full publicity condition should not pose a threat to contractualism. A primary reason for the stability and publicity requirement for Rawls in *A Theory of Justice* is to show that humans 'have a moral nature', and that they can reason from basic moral principles without self-defeating consequences that lead people to reject those very principles. For Rawls, humans can be said to 'have a moral nature' so long as (1) they have a regulative sense of justice and can understand, apply, and act on and from principles of justice, and (2) they regard justice as rational and essential to their good. Regarding (1), in *A Theory of Justice*, Chapter 8, Rawls demonstrates how people in a well-ordered society normally acquire a sense of justice, and shows that acting for the sake of justice is within the capacities of human moral psychology. The Kantian congruence argument in Chapter 9 was supposed to establish (2). The problem with congruence was that, even if it is true that justice is a highest-order good since it realizes the capacities for human agency and moral autonomy, due to the burdens of judgment many reasonable people in a well-ordered society refuse to recognize justice in this manner. This problem is resolvable, according to *Political Liberalism*, so long as an overlapping consensus eventuates in a well-ordered liberal society; then, all reasonable people still can regard justice as essential to their good and thus rational, but not because it realizes their capacities for agency and their moral autonomy.

If we assume that contractualism is true (independent of the full publicity condition) and that it endorses political liberalism and a liberal political conception of justice, then they are true also. The fact that there will always be reasonable persons who, because of the burdens of judgment, have false beliefs about the true contractualist origins of justice and morality does not show that humans do not have a moral nature or that morality is not compatible with the human good, in the sense required by (1) and (2). It shows, rather, that even the best of people (those with a steadfast will to do what is right and just for moral reasons) can be mistaken about the nature and requirements of moral reasons. Some just and conscientious people believe that the origin of justice and other moral requirements lies in aggregate utility; others, in God's commands or in moral agency and autonomy or in cultural conventions, and so on. They all have (according to contractualism) mistaken philosophical beliefs. Ultimately, they accept different reasons in justification of their own and others' morally justifiable conduct, though they agree on principles of justice and even their public political justification. But their mistaken beliefs about justice's comprehensive philosophical foundations surely are not a reason for denying humanity's capacity for justice and its moral nature (even if there are grounds for denying these persons' capacity for autonomy). Nor should their mistaken beliefs alone lead to the instability of a well-ordered liberal society, so long as there is an overlapping consensus on liberal political conceptions. If this is true, then it would seem that the failure of

33

the full publicity of a comprehensive contractualist moral conception should not jeopardize the stability with full publicity of the liberal political conception of justice that contractualism itself endorses as true.

I cannot think of any significant remaining problems stemming from contractualism's inability to satisfy a full publicity condition. Granted, the burdens of judgment still imply that a second aim that Rawls had for full publicity[86] is not achievable by contractualism: the condition for moral autonomy (that people not have false beliefs about their social and moral relations). But as I understand Scanlon, this Kantian and idealist aspiration is not an aspect of his position. What is important is mutual recognition and respect, which are achievable only so long as people can justify their conduct to one another according to rules that they cannot reasonably reject. This standard can be met adequately, if not entirely, in a society that is well ordered by a political conception of justice that satisfies the criteria of reciprocity and liberal legitimacy.

8. The main difference between Rawls's and Scanlon's contract doctrines

One of Rawls's primary ambitions in the full publicity condition is to provide a generally acceptable basis for public justification for a democratic society by reference to which democratic citizens may *justify to one another, as free and equal moral persons*, the arrangement and workings of their basic social and political institutions. This aspiration was announced in 'Kantian Constructivism in Moral Theory' (1980), well before Rawls recognized the problems that led to political liberalism. One way to understand Scanlon's early article 'Contractualism and Utilitarianism' (1982) is that he aimed to extend a similar idea of public justification to reasonable persons *as such* in a well-ordered contractualist society to apply to morality generally. But no comprehensive moral conception, including the Kantian interpretation of justice as fairness and contractualism, could meet the full publicity condition in any feasible well-ordered society because of inevitable differences in comprehensive views.

Rawls responds to this problem by revising the kind of public justification a conception of justice is to provide. A conception of justice can be generally accepted and meet the full publicity condition only if it is a *freestanding political conception* that provides a basis for *public political justification*. Public political justifications are justifications *to citizens* in a well-ordered society in their capacity *as free and equal citizens*, and on the basis of *public reasons* (as opposed to nonpublic or comprehensive reasons) and political values they also accept as citizens.[87]

At this point, Scanlon's position is in need of clarification. *To whom* are contractualist justifications now to be addressed if they cannot be accepted by all reasonable members of a well-ordered contractualist society? There are two alternatives, either of which suggests that contractualism's aspirations must differ

from Rawls's contractarianism in certain important, perhaps fundamental, respects. One alternative is that contractualist justifications are to be addressed to those reasonable persons in a well-ordered society who already accept contractualism. That is, one way to preserve a full publicity condition is to truncate the class of reasonable persons to whom public justifications are addressed. The problem with this reading is that real-world contractualists, even if they are attuned to contractualist reasons and their moral beliefs are ostensibly in order, still have other false philosophical, religious, and evaluative beliefs (regarding personal identity or the personhood of fetuses and embryos, for example). It is hard to see how these comprehensive doctrines would not influence and, when false, taint their moral convictions and what they regard as unreasonable to reject, in spite of their avowal of contractualism. (For example, two contractualists, one of whom believes that fetuses have souls, will not agree on the moral permissibility of abortion, even though they might accept its political permissibility.) The benefit of Rawls's idea of a freestanding political conception as a basis for public justification and his account of public reason is that they insulate his political conception of justice from its advocates' comprehensive reasons and beliefs. But this alternative may not be open to Scanlon's contractualist position: as a comprehensive moral doctrine, it cannot be made freestanding of comprehensive philosophical views.

Another, more promising alternative is for contractualism to forgo any requirement of justification to reasonable members of a well-ordered society. Contractualist justifications can then be understood as justifications *to* the hypothetical contractualist agents who are fully informed and have different interests and particular aims which provide them with 'personal reasons', but who do not have false moral, religious, and philosophical beliefs that prevent them from agreement on contractualist principles or contractualist justifications of those principles.[88]

Assuming this second alternative, the main difference between Rawls's and Scanlon's view might be clarified this way: justice as fairness, as a politically liberal view, is designed to provide a public conception of justice for a feasible social world (a well-ordered democratic society) where everyone accepts the same political conception of justice (or at least a liberal political conception) which serves them as a basis for public justification of the myriad uses of coercive political power. Public justifications are justifications *to* realistically possible persons in two capacities they all occupy: first, in their capacity as free and equal citizens who are reasonable and rational and, second, in their capacity as reasonable and rational persons who have a reasonable (albeit often untrue) comprehensive doctrine and a conception of their good. By contrast, contractualist justifications are justifications *to* hypothetical moral agents who are adequately informed and unencumbered by false or misguided comprehensive philosophical, moral, and religious doctrines, and who are morally motivated to justify themselves in ways that are sensitive to the legitimate interests and 'per-

35

sonal reasons' that contractualist agents have. In this regard, contractualism tailors reasonable moral principles to the *pluralism of values*, but not to the pluralism of reasonable comprehensive (philosophical and religious) doctrines and differences in moral beliefs. Why one, but not the others? My conjecture is that contractualism is geared to respond mainly to the multiple objective goods that hypothetical contractualist agents freely pursue ('objective' since these agents are not encumbered by false beliefs leading them to affirm false values), but where these objective goods are nonetheless the source of potentially conflicting 'personal reasons' since the goods agents pursue are freely chosen. This would mean that, unlike political liberalism, contractualism does not seek to accommodate Rawls's 'subjective circumstances of justice' or 'burdens of judgment', which include the plurality of reasonable conceptions of the good (some of which may contain false values) and diverse comprehensive philosophical, religious, and moral conceptions (some of which contain false beliefs) that exist in any feasible social world. If so, then Scanlon's contractualism, unlike Rawls's, is not a response to permanent facts about human nature and society, but to the plurality and potential conflict of objective values. This marks a fundamental difference between the two contract views, and belies the common assumption that Scanlon's contractualism is an extension of Rawls's contractarian account of justice to morality more generally. They are fundamentally different applications of the contractarian idea of reasonable agreement.

I end with an observation about Rawls's understanding of contractarianism soon after *Political Liberalism*. There are two things Rawls said (in conversation) in this connection. First, he thought that the main ideas of social contract doctrine worked well in the context of addressing problems of justice, but that it was difficult to extend the social contractarian framework to other areas in moral philosophy. Whether this amounted to a retraction of his suggestion in *A Theory of Justice* of 'rightness as fairness', a comprehensive contractarian moral conception, was not clear. Second, Rawls always referred to justice as fairness as a 'contractarian' position. He was opposed to others' use of Scanlon's term 'contractualism' as a generic term used to refer to justice as fairness and other non-Hobbesian contract conceptions of justice and their efforts to reserve the term 'contractarianism' for Hobbesian views. He said contractualism was Scanlon's own position, and that it was original, distinctive, and in important respects quite different from what he was trying to do. I conjecture that part of what he had in mind is that contractualism is not designed to meet three conditions that play such a fundamental role in Rawls's contractarianism: the full publicity of a political conception of justice, the stability of a well-ordered society, and the social role of a political conception in providing a public basis for justification.

36

notes

1. As Rawls says of constructivism in moral philosophy and mathematics: 'In both cases the idea is to formulate a procedural representation in which . . . all the relevant criteria of correct reasoning – mathematical, moral, political – are incorporated and open to view. Judgments are reasonable and sound if they result in following the correct procedure correctly and rely only on true premises.' See John Rawls, *Political Liberalism*, expanded edn (New York: Columbia University Press, 2004 [1993]), p.102.
2. See John Rawls, 'Kantian Constructivism in Moral Theory', *Journal of Philosophy* 77 (1980): 515–72 at p. 519; also in Rawls, *Collected Papers*, edited by Samuel Freeman (Cambridge, MA: Harvard University Press, 1999), p. 307. Scanlon agrees in T.M. Scanlon, 'Contractualism and Utilitarianism', in *Utilitarianism and Beyond*, edited by A.K. Sen and Bernard Williams (Cambridge: Cambridge University Press, 1982), pp. 103–28.
3. Rawls, *Political Liberalism*, p. 90.
4. Other ethical constructivists could agree with this statement, so far as it goes. Gauthier, for example, also seeks to show that morality is simply an extension of practical rationality, conceived as individual utility maximization.
5. Rawls, *Political Liberalism*, p. 99.
6. 'Political Constructivism accepts [Kant's] view that the principles of practical reason originate, if we insist on saying they originate anywhere, in our moral consciousness as informed by practical reason. They derive from nowhere else . . . Still, accepting this, it is a separate question whether the principles of practical reason constitute the order of values.' See ibid., p. 100.
7. Ibid., p. 99.
8. Ibid., p. 101.
9. Compare ibid., p. 100.
10. John Rawls, 'Themes in Kant's Moral Philosophy', in *Collected Papers*, edited by Samuel Freeman (Cambridge, MA: Harvard University Press, 1999), pp. 511–3.
11. Rawls, *Political Liberalism*, p. 97.
12. As set forth in John Rawls's Dewey Lectures. See John Rawls, 'Kantian Constructivism in Moral Theory', *Journal of Philosophy* 77 (1980): 515–72.
13. John Rawls, *A Theory of Justice* (Cambridge, MA: Belknap Press, 1971), pp. 251–2.
14. Ibid., p. 252.
15. 'Kant held, I believe, that a person is acting autonomously when the principles of his action are chosen by him as the most adequate possible expression of his nature

37

as a free and equal rational being.' See ibid., p. 252. '[T]o express one's nature as a being of a particular kind is to act on the principles that would be chosen if this nature were the decisive determining element'. See ibid., p. 253. The original position, Rawls goes on to contend, is a 'procedural interpretation' of our nature as free and equal rational persons, or, as Rawls says, 'the original position interprets the point of view of noumenal selves, of what it means to be a free and equal rational being'. See ibid., pp. 255–6. 'The principles regulative of the kingdom of ends are those that would be chosen in this position, and the description of this situation enables us to explain the sense in which acting from these principles expresses our nature as free and equal rational persons.' See ibid., p. 256. Rawls's analogue to Kant's realm of ends is the idea of a well-ordered society.

16. Rawls defines a 'well-ordered society' as one in which everyone accepts the same conception of justice, this is publicly recognized, and society realizes the terms of this conception. See ibid., pp. 453–62.

17. Ibid., p. 580.

18. On the Aristotelian principle, see ibid., pp. 424–33.

19. This argument is mainly found ibid., Section 86, pp. 567–77. For a reconstruction of the congruence argument, see Samuel Freeman, 'Congruence and the Good of Justice', in *The Cambridge Companion to Rawls*, edited by Samuel Freeman (New York: Cambridge University Press, 2003); *Justice and the Social Contract* (New York: Oxford University Press, 2006), Ch. 6.

20. Of Kantian constructivism Rawls says, 'The leading idea is to establish a suitable connection between a particular conception of the person and first principles of justice, by means of a procedure of construction.' 'Kantian Constructivism in Moral Theory', *Journal of Philosophy*, p. 516.

21. 'A conception of justice is framed to meet the practical requirements of social life and to yield a public basis in the light of which citizens can justify to one another their common institutions.' See Rawls, 'Kantian Constructivism in Moral Theory', *Journal of Philosophy*, p. 561.

22. Ibid., p. 565.

23. Rawls, *Political Liberalism*, p. 102, p. 90.

24. Rawls, 'Kantian Constructivism in Moral Theory', *Journal of Philosophy*, p. 517.

25. Rawls, *A Theory of Justice*, p. 580.

26. Rawls, *A Theory of Justice*, p. 582.

27. Compare Philip Pettit's criticism of contractualism in his review of T.M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1999). See Philip Pettit, 'Two Construals of Scanlon's Contractualism', *Journal of Philosophy* 97 (2000): 148–64. Pettit tries to incorporate contractualism into a consequentialist position, and contends that contractualism is not about rightness and wrongness, but about civility and uncivility.

28. Rawls, *A Theory of Justice*, p. 586.

29. Rawls says that the natural duty of mutual respect requires giving reasons for our actions whenever they constrain or materially affect others. See ibid., pp. 337–8. Compare also ibid., pp. 102–3, 178–9, where Rawls says that the conception of reciprocity expressed by the difference principle is that 'The social order can be justified to everyone, and in particular to those who are least favored; and in this sense it is egalitarian.' That we have a duty to justify our actions and institutions on

38

the basis of public reasons supplies the basis for what Rawls will later call in *Political Liberalism* the 'liberal principle of legitimacy' and the 'duty of civility'. See Rawls, *Political Liberalism*, pp. 216–7 notes 136, 393.

30. Compare 'Reflective equilibrium . . . is a notion characteristic of the study of principles which govern actions shaped by self-examination. Moral philosophy is Socratic.' See Rawls, *A Theory of Justice*, pp. 48–9.

31. This is the title of Rawls's Presidential Address, APA Eastern Division, 1974. See John Rawls, *Collected Papers*, edited by Samuel Freeman (Cambridge, MA: Harvard University Press, 1999), Ch. 15.

32. R.M. Hare claims utilitarianism is based in a 'right to equal concern and respect . . . a precept which leads straight to Bentham's formula and to utilitarianism itself'. See R.M. Hare, 'Reply to Mackie', in *Utility and Rights*, edited by R.G. Frey (Minneapolis: University of Minnesota Press, 1984), p. 107.

33. A similar line of reasoning underlies Scanlon's contractualism, which is why I have called this the 'contractualist' strand in Rawls's argument for a public basis for justification. See Scanlon, *What We Owe to Each Other*, p. 162, where Scanlon refers to the contractualist moral relation as one of 'mutual recognition'.

34. Rawls, *A Theory of Justice*, p. 582.

35. Ibid.

36. The publicity of principles of justice partially defines a well-ordered society; therein, all reasonable persons accept the same principles of justice and also everyone knows that everyone else accepts the same principles. See note 16 above. Compare ibid., pp. 133, 177.

37. Ibid., p. 177.

38. Rawls goes on to say that, in view of the 'role of moral principles in establishing the ties of community', publicity is a natural condition if we 'take seriously the idea . . . of society as a social union of [social] unions'. See ibid., p. 582.

39. See ibid., p. 133. Compare ibid., p. 175.

40. Ibid., p. 256.

41. Ibid., p. 133.

42. Rawls, 'Kantian Constructivism in Moral Theory', *Journal of Philosophy*, p. 537.

43. Ibid., p. 539.

44. Ibid., p. 553.

45. Compare Rawls, *A Theory of Justice*, p. 572.

46. Rawls, 'Kantian Constructivism in Moral Theory', *Journal of Philosophy*, p. 542.

47. For example, the liberal Catholic in a well-ordered society accepts the principles and natural duties of justice as fairness, but sees them as natural laws willed by God in order to enable rational beings to achieve the *summum bonum*, the contemplation and glorification of God. According to the liberal Thomist, natural laws of justice are divinely ordained and are knowable, some as self-evident truths, by the natural light of reason. This denies a basic assumption of Kantian constructivism: that moral principles are constructed from principles of practical reason. Rejecting this, the liberal Thomist also rejects the intrinsic good of autonomy that underpins the congruence argument. This argument depends on showing that the sense of justice is the same as the desire to realize our nature as free and equal rational beings. But the liberal Thomist denies this identification. For the sense of justice is best regarded as a desire to conform to God's natural laws, not as a desire to express our

39

nature as the author of these laws. The idea of moral autonomy is regarded as a conceit of human reason.

48. 'Reasonable persons' or citizens according to Rawls have an effective sense of justice and desire to abide by fair terms of cooperation. They recognize a duty of mutual respect and want to justify their actions and institutions to others on grounds all can reasonably accept. They also recognize the burdens of judgment and accept their consequences. See Rawls, *Political Liberalism*, pp. 48–9, 49 notes 1, 2. Rawls also says that reasonable people 'are willing to govern their conduct by a principle from which they and others can reason in common; and reasonable people take into account the consequences of their actions on others' well-being'. See ibid., p. 49 note 1).

49. Rawls, 'Themes in Kant's Moral Philosophy', p. 512.

50. Rawls, 'Kantian Constructivism in Moral Theory', *Journal of Philosophy*, p. 568.

51. Ibid., p. 570.

52. The two moral powers are (1) a capacity for a sense of justice, to understand, apply, and act from the public conception of justice, including a willingness to act in relation to others on terms they can reasonably accept, and (2) a capacity to form, revise, and pursue rationally a conception of the good. See Rawls, *Political Liberalism*, pp. 19, 103–4.

53. *Political Liberalism*, p. 90; see also pp. 102, 114.

54. That is 'must meet', unless one assumes that correct reasoning from our considered moral convictions obscures access to moral truth. Some have made this criticism of Rawls. On a related note, whether Rawls's principles are 'reasonable' in his sense depends, of course, on whether his argument works. Here controversy may arise with utilitarians, such as Sidgwick, who contend that among the requirements, or 'philosophical intuitions', of practical reason is the principle of impartial benevolence: that we ought to pursue the good impartially construed. For utilitarians, this is about as fixed a considered moral conviction as there can be, so much so that they would sooner give up the reflective equilibrium of most of their remaining considered moral convictions rather than compromise this fundamental intuition. Insofar as the requirement of impartial benevolence depends, as it does in Sidgwick, on a claim about an epistemology of self-evidence, it would not be admissible within political liberalism.

55. John Rawls, *The Law of Peoples* (Cambridge, MA: Harvard University Press, 1999), pp. 86–7.

56. Ibid., p. 87.

57. Rawls, *Political Liberalism*, pp. 97–8.

58. Ibid.

59. Ibid., pp. 137, 217, 393. The liberal principle of legitimacy says: 'The exercise of political power is legitimate only when it is exercised in fundamental cases in accordance with a constitution, the essentials of which all reasonable citizens as free and equal might reasonably be expected to endorse.' See ibid., p. 393. In 'The Idea of Public Reason Revisited', Rawls adds that 'the reasons we would offer for our political actions' must be such that 'we also reasonably think that other citizens might also reasonably accept those reasons'. See John Rawls, 'The Idea of Public Reason Revisited', in *Collected Papers*, edited by Samuel Freeman (Cambridge, MA: Harvard University Press, 1999), p. 578.

40

60. 'All actions that affect the rights of other men are wrong if their maxims are not consistent with publicity.' See Immanuel Kant, *Perpetual Peace*, in *On History*, edited and translated by Lewis White Beck (Indianapolis: Bobbs-Merrill, 1963), p. 381.
61. Ibid., p. 38.
62. Ibid.
63. Kant evidently thought that, in a world where everyone exercises their reason in this way, all would come to discern and agree only on laws that respect persons as autonomous beings: 'The criterion of everything that can be agreed upon as law by a people lies in this question: Can a people impose such a law upon itself?' 'An Answer to the Question: What is Enlightenment', in *Perpetual Peace and Other Essays*, trans. Ted Humphrey (Indianapolis, IN: Hackett, 1983), p. 44. Kant is clear that the ground for the self-imposition of law by a people is the free public use of their reason.
64. Rawls, *Political Liberalism*, pp. 54–8.
65. The burdens of judgment are not a skeptical denial that philosophical and ethical claims admit of truth or validity. Nor do they imply a philosophical claim regarding the limits of reason, as in Kant's metaphysics. The idea of the burdens of judgment, rather, is an empirical conjecture about the effects of free institutions and the limits on our capacities for agreement in our reasoned and conscientious judgments. See ibid.
66. Rawls clearly thinks Kant's moral constructivism cannot meet conditions Rawls imposes on a political conception of justice: since in Kant's view, 'the ideal of autonomy has a regulative role for all of life . . . it is not suitable to provide a public basis of justification'. See ibid., p. 99.
67. Then again, Thomas Hill reminded me that if Kant's duty to promote the realm of ends is nothing more than the duty to act on and from the moral law by applying the categorical imperative procedure, then the former duty imposes no extra requirements on us.
68. See T.M. Scanlon, 'The Aims and Authority of Moral Theory', *Oxford Journal of Legal Studies* 12 (1992): 12. See also Scanlon, 'Contractualism and Utilitarianism', p. 127, where he says his expectations for moral argument are more modest than Rawls's.
69. For Scanlon's original account of moral motivation, which is expressed in terms of a 'desire to justify' oneself, see Scanlon, 'Contractualism and Utilitarianism', pp. 115–19. In his later book, *What We Owe to Each Other*, he no longer sees desire as an independent feature of motivation, and now refers to 'the aim of justifiability to others'. See Scanlon, *What We Owe to Each Other*, p. 157.
70. 'It is fitting, then, that the fair terms of social cooperation between citizens as free and equal should meet the requirements of full publicity. For if the basic structure relies on coercive sanctions, however rarely and scrupulously applied, the grounds of its institutions should stand up to public scrutiny.' See Rawls, *Political Liberalism*, p. 68. Compare here Rawls's liberal principle of political legitimacy, which says the exercise of political power is appropriate only if exercised in accordance with a constitution whose essentials all citizens may reasonably be expected to accept. See ibid., pp. 137, 217, 393. The reason he gives for this requirement is that political power is coercive, and in a democracy is to be

41

exercised by the body of free and equal citizens themselves. See ibid., pp. 216–7. Moreover, Rawls suggests that full publicity 'may be less compelling for comprehensive doctrines generally; [this] is left open as a separate question.' See ibid., p. 67.

71. Henry Sidgwick, *Methods of Ethics*, 7th edn (Indianapolis: Hackett, 1981), pp. 489–90.

72. Scanlon, *What We Owe to Each Other*, p. 268.

73. Scanlon, 'Contractualism and Utilitarianism', p. 110.

74. All these arguments are from Rawls, *A Theory of Justice*, Section 29.

75. Scanlon, 'Contractualism and Utilitarianism', p. 128. In *What We Owe to Each Other*, Scanlon says things that might suggest a publicity condition, though not as clearly as earlier. He says that 'in deciding whether a principle could reasonably be rejected . . . an assessment of the rejectability of a principle must take into account the consequences of its acceptance in general, not merely in a particular case that we may be concerned with'. See Scanlon, *What We Owe to Each Other*, p. 204. Moreover, 'we must take into account . . . implications (for both agents and others) of having agents be licensed and directed to think in the way that that principle requires'. See ibid., p. 203. These statements, however, might be understood as referring to hypothetical contractualist agents, and not the members of a well-ordered contractualist society.

76. Rawls, 'Kantian Constructivism in Moral Theory', *Collected Papers*, p. 305, emphasis added.

77. Scanlon, 'Contractualism and Utilitarianism', p. 117.

78. Ibid., p. 116.

79. I am grateful to Gideon Rosen for helping me to clarify the basic idea underlying this paragraph.

80. That is, all accept moral norms because these norms could not be reasonably rejected (by fully informed and fully reasonable persons) as a basis for informed, unforced general agreement.

81. In this respect, moral contractualism would differ from political constructivism, a distinctive feature of which is that it offers a 'freestanding' basis for public justification.

82. In this connection, I mentioned earlier in this section the peculiarity of contractualism's requirement of justification *to* persons when it does not apply to real-world agents, but only to hypothetical contractualist agents with full knowledge and no false beliefs. See below, in Section 7, where I revisit this issue.

83. This seems to be Habermas's reply to Rawls, in Jürgen Habermas, 'Reconciliation Through the Public Use of Reason: Remarks on John Rawls's Political Liberalism', *Journal of Philosophy* 92 (1995): 109–31. See also G.A. Cohen, 'Facts and Principles', *Philosophy and Public Affairs* 31 (2003): 211–45.

84. See Sidgwick, *Methods of Ethics*, pp. 489–90.

85. The criterion of reciprocity says that 'for terms [of cooperation] to be fair terms, those offering them must reasonably think that those citizens to whom such terms are offered might also reasonably accept them'. See Rawls, *Political Liberalism*, p. xliv. Of reasonable persons and reciprocity Rawls says: 'Those norms they view as reasonable for everyone to accept and therefore as justifiable to them . . . That [society's] fair terms be reasonable for all to accept is part of its idea of reciprocity.'

See ibid., pp. 49–50. For the principle of legitimacy, which Rawls says is based on the criterion of reciprocity, see Rawls, 'The Idea of Public Reason Revisited', p. 578, note 57.

86. See Section 3 above.

87. There are three levels of justification in *Political Liberalism* as follows. (1) *Public political justification*: it is the role of a *political* conception of justice to provide not a full public justification, but a *public political justification*, which is a justification *to* free and equal citizens in a well-ordered society in their capacity *as citizens* and on the basis of *public reasons*. (2) *Full justification*: by contrast, 'full justifications' of the conception of justice are not public, but are to be provided by each reasonable comprehensive doctrine in terms of its own comprehensive reasons. (This implies, of course, that not all 'full justifications' can be entirely true.) Finally, (3) *Public justification* of the political conception exists, when there is *public knowledge* that (a) each reasonable comprehensive doctrine accepts, and there is an overlapping consensus on, the political conception of justice and its political justification, and (b) each reasonable doctrine has fully justified the political conception in terms of its own comprehensive reasons. See John Rawls, 'Reply to Habermas', Lecture IX, in *Political Liberalism*, expanded edn (New York: Columbia University Press, 2004 [1993]).

88. There may be some problem here in separating a person's personal reasons from the comprehensive moral and philosophical views that lead them to have beliefs about values and the nature of morality.

43