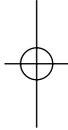# 9

Rawls and Moral Psychology

Thomas Baldwin

In his obituary of John Rawls Ben Rogers remarked that after completing *A Theory of Justice* (*TJ*)[1] Rawls intended to develop his ideas on moral psychology.[2] In the event the debates aroused by *TJ* kept Rawls fully occupied and he never wrote an extended account of the subject. But there are discussions of it throughout his writings and it merits more attention than it has received (for example, the four-volume collection of papers on Rawls edited by Chandran Kukathas contains no papers directly on this theme[3]). My aims here are to elucidate Rawls's conception(s) of moral psychology and then to explore critically some of the complexities and tensions inherent in his uses of it, especially those arising from its roles in his moral and political theories. Some of these questions are focused around the 'problem of stability', the problem of showing how a just society is likely to be stable because it provides a basic framework for the activities of its members which they recognize as congruent with their individual interests; and at the end of this paper I will discuss Rawls's treatment of this issue, both in *TJ* and in his later writings, and, in this context, consider how far the moral psychology Rawls relies on to address this problem has an essential social dimension.

This last question, how far Rawls's moral psychology has a social dimension, may immediately give rise to the reflection that it should be

[1] Rawls (1971); in 1999 Rawls published a revised edition of the book—Rawls (1999*a*). I give page numbers for both editions in the form '(*TJ* x; y)', where x is the page number in the first edition and y is the number in the second edition. As I explain in n. 7, one of the revisions Rawls makes is relevant to the argument of this paper.

[2] See www.guardian.co.uk/obituaries/story/0,3604,848488,00.html.

[3] Kukathas (2002).
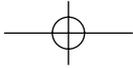
no surprise if his moral psychology turns out to be inherently sociological or political since Rawls was primarily a political philosopher. Yet although Rawls was of course a great political philosopher, his political philosophy was initially founded on a broadly Kantian moral philosophy, and his moral psychology first took shape within this context. It is then a delicate matter, especially in the context of those of his later writings which are primarily contributions to political philosophy, to identify those discussions which are not exclusively directed to questions of political philosophy; but his moral psychology is, I believe, one of the areas less affected by his later emphasis on the political. Hence even for Rawls there is a substantive issue as to how far moral psychology is inherently social or political; and thinking about the way in which this issue plays out for Rawls certainly provides a stimulating way of thinking about the issue itself.

1. WHAT IS MORAL PSYCHOLOGY?

The first issue to be addressed is what it is that Rawls means when he writes of 'moral psychology'. We get an initial answer to this question from chapter 8 of *TJ* where there is a long section (§75) called 'The Principles of Moral Psychology'. Rawls here summarizes and comments on the account of moral development he has presented in the preceding sections, in particular his account of the development of a 'sense of justice', which is a disposition to act in accordance with the principles of justice for their own sake and to feel guilt or shame when one recognizes that one has violated these principles. This gives us one feature of moral psychology, namely that it deals with the development of feelings and judgments whose content is distinctively moral. But a further point is that, according to Rawls, our psychology is itself affected by the moral value of the context in which we grow up and live:

Perhaps the most striking feature of these laws (or tendencies) is that their formulation refers to an institutional setting as being just, and in the last two as being publicly known to be such. The principles of moral psychology have a place for a conception of justice … Thus some view of justice enters into the explanation of the corresponding sentiment; hypotheses about this psychological process incorporate moral notions even if these are understood only as part of the psychological theory. (*TJ* 491; 430)

Rawls recognizes that some theorists will regard this as odd: 'No doubt some prefer that social theories avoid the use of moral notions' (*TJ* 491; 430). But, he holds, this is a mistake: 'The justice or injustice of society's arrangements and men's beliefs about these questions profoundly influence the social

feelings' (*TJ* 492; 431). Thus in his use here of the term 'moral psychology' Rawls implies that in some respects our psychology is inherently 'moral', not only in respect of its content, but also in respect of its dependence upon the justice, and thus the morality, of our society.

In *TJ* Rawls maintains that our 'natural attitudes' bring with them a 'liability' (*TJ* 489; 426) to moral sentiments, such that in the normal course of human development in a reasonably just society a normal human being develops a sense of justice. Hence, as Rawls puts it 'The moral sentiments are a normal part of human life. One cannot do away with them without at the same time dismantling the natural attitudes as well' (*TJ* 489; 428). This account of the matter suggests that moral psychology, understood as the psychology of the moral sentiments, deals with an aspect of the normal development of human beings, and therefore belongs within a comprehensive account of human psychology. But this position seems to be at variance with that affirmed by Rawls in his later writings. In lecture II of *Political Liberalism* (*PL*),[4] Rawls gives the final section (§8) the somewhat puzzling title 'Moral Psychology: Philosophical not Psychological' (*PL* 86).[5] He then begins the section as follows:

1. This completes our sketch of the moral psychology of the person. I stress that it is a moral psychology drawn from the political conception of justice as fairness. It is not a psychology originating in the science of human nature but rather a scheme of concepts and principles for expressing a certain political conception of the person and an ideal of citizenship. (*PL* 86–7)

It seems clear Rawls is now using the expression 'moral psychology' in a rather different way from that in which he had used it in *TJ*, as a way of capturing 'a certain political conception of the person and an ideal of citizenship'. One issue here is the importance of the qualification of this conception of the person as 'political', but I want to set this aside for the moment: one can find largely similar accounts of the conception of the person in Rawls's middle-period lectures 'Kantian Constructivism in Moral Theory' (KC),[6] without any qualification of this conception as political. I should add that in these lectures he does not make much use of the phrase 'moral psychology' to describe this conception of a person, but the phrase does occur at least once with this use (KC 346) and the substance and role here of the conception of a person is much the same as that which it plays in later writings such as *PL* where he routinely describes it as moral psychology.

---

[4] Rawls (1993).

[5] How, one wants to ask, can a 'psychology' not be 'psychological'?

[6] These lectures were originally published in 1980. They are reprinted in Rawls (1999*b*), and page references are to this edition. See pp. 330–3 for the account there of the conception of a person.

What, then, is the role and substance of this later conception of moral psychology? As the passage above from *PL* indicates, it is intended to capture the conception of a person which is central to moral and political theory. Indeed Rawls recognizes that different moral theories will bring with them different moral psychologies and he especially contrasts the 'sparse' moral psychology implicit in the 'rational intuitionism' of moral realists such as Moore and Ross (*PL* 92) with that which is central to his own neo-Kantian constructivism. Central to this latter psychology is the attribution to persons of two 'moral powers', a capacity for a sense of justice and a capacity for a conception of the good (*PL* 81). The first of these was central to the account of moral psychology in *TJ*, but the second, the capacity for a conception of the good, is not part of that account at all. It comprises both the fact that a person has certain values and final ends, things which they care about or aim at for their own sake and which bring with them a more or less explicit way of thinking about their relationships with others and the world, and also the fact that they have the ability to revise these values and ends in the light of new evidence or other reasons (*PL* 19). Rawls adds that, in addition to having these moral powers, a person should be conceived to have further dispositions which are 'aspects of their being reasonable and having this form of moral sensibility' (*PL* 81). These further dispositions include a 'readiness to propose and abide by fair terms of cooperation' provided one has reasonable assurance that others will also do their part, and a tendency to develop trust and confidence in others as the success of cooperative arrangements is sustained (*PL* 86).

Much of this latter material was in fact present in the account of the development of a sense of justice in *TJ*, so it is primarily the emphasis on the capacity for a conception of the good which marks a substantive addition to the content of his early account of our moral psychology.[7] This difference between *TJ* and his later writings shows that the fundamental difference between the two conceptions of moral psychology lies in their role in Rawls's presentation of his moral philosophy. As I have indicated, its role in his work from 'Kantian Constructivism in Moral Theory' onwards belongs to the description of the capacities and dispositions whose possession by persons is essential to the articulation of a moral theory, be it Rawls's own Kantian constructivism, Ross's rational intuitionism, or Mill's utilitarianism. Rawls

---

[7] In the revised edition of *TJ* Rawls does signal the importance of this capacity. In § 82 ('The Grounds for the Priority of Liberty') he describes the motivations of the parties in the hypothetical original position and remarks: 'The parties conceive of themselves as free persons who can revise and alter their final ends and who give priority to preserving their liberty in this respect' (*TJ* 475). Like most of § 82, however, this passage was not present in the first edition (as Rawls acknowledges in his 'Preface to the Revised Edition', p. xiii).

does of course also provide descriptions of these capacities and dispositions in *TJ*, especially in the first part of the book (e.g. § 25). But none of this material is here described as moral psychology; instead in *TJ* the 'principles of moral psychology' are the psychological laws which govern the development of our moral sentiments and explain the possession of a sense of justice. These principles are introduced specifically in order to help with the problem of stability (*TJ* 453; 397) because, Rawls thinks, if it is part of normal human psychology that moral sentiments such as a sense of justice develop within the context of life in a just society, then a normal person who is a citizen of a just society should be motivated to fulfil the requirements of justice, with the result that such a society should be reasonably stable. As we shall see below, Rawls thinks that there is more to be said on this matter: but on the face of it this 'stabilizing' role of moral psychology is quite different from its foundational role in *PL*. An easy way to bring out the difference here is to take the case of Rational Intuitionism. According to Rawls the sparse moral psychology implicit in Rational Intuitionism is primarily one which ascribes to persons a capacity for knowledge of moral principles and a capacity for motivation by this knowledge (*PL* 92). It is obvious that this moral psychology does little to show that it is in a person's interest to act in accordance with this motivation; but it was that task which was to be assisted by moral psychology in its stabilizing role.

In the light of this discussion I want to return to Rawls's characterization of moral psychology in *PL* as 'Philosophical not Psychological' (see the passage quoted above, *PL* 86). The contrast he draws here is one between moral psychology conceived as 'a scheme of concepts and principles for expressing a certain political conception of the person and an ideal of citizenship' and 'a psychology originating in the science of human nature'. Plainly the first of these conceptions identifies the foundational role of moral psychology in moral and political philosophy; whereas the second, which is disavowed, concerns an approach to the psychology of the moral sentiments which originates in a 'science of human nature'. This contrast is overtly drawn in terms of origins: philosophical versus scientific. But the question to which this contrast gives rise is whether a further contrast is implied, intentionally or not, between an a priori philosophical moral psychology and an empirical scientific psychology. That would seem to threaten an untenable dualism, reminiscent of Kant's distinction between the noumenal and empirical selves which Rawls would like to think he has discarded (*TJ* 256–7; 226). In fact it is clear that Rawls's contrast between a philosophical moral psychology and a scientific one is not intended to be exclusive: the contrast is fundamentally one of rationale, and Rawls explicitly affirms that it is a condition of any acceptable philosophical moral psychology that it
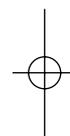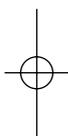
be consistent with our natural capacities: 'Human nature and its natural psychology are permissive: they may limit the viable conceptions of persons and ideals of citizenships and the moral psychologies that may support them, but do not dictate the ones we must adopt' (*PL* 87).

This shows that Rawls is not a metaphysical dualist; however, it still leaves open a question as to the relationship between our 'natural psychology' and the favoured philosophical moral psychology which it 'permits' but does not 'dictate'. The principles of moral psychology propounded in *TJ* are supposed to be psychological laws which concern the development of moral sentiments and capacities such as a sense of justice which belong within Rawls's favoured philosophical moral psychology. So the picture we get there is one of an intimate explanatory relationship between natural and moral psychology. Admittedly, natural psychology does not by itself 'dictate' moral psychology so conceived, since the development of moral sentiments is contingent upon the moral character of the relationships and society in which the individuals concerned grow up and live. Nonetheless, given, to use Rawls's own phrase, 'The Connection between Moral and Natural Attitudes' (the title of § 74 of *TJ*), it follows that a complete understanding of human life, a true 'science of human nature' as one might put it, has to make room for our moral sentiments; for (to repeat a passage quoted earlier) 'The moral sentiments are a normal part of human life. One cannot do away with them without at the same time dismantling the natural attitudes as well' (*TJ* 489; 428). Hence Rawls's early work encourages the prospect of a unified explanatory approach to human psychology which embraces both natural and moral psychology.

On the face of it, this prospect is not sustained in Rawls's later writings, where he seems primarily concerned to put a distance between his philosophical moral psychology, the psychological assumptions inherent in his moral philosophy, and natural psychology, the empirical science of human nature, even if the latter has to 'permit' the former. But the issue is not clear: for Rawls remained concerned to provide a solution to the problem of stability and unless there are some substantive connections between the demands of morality, and thus our moral psychology, on the one hand, and our 'natural' psychology, on the other hand, the problem of stability will remain unsolved. For stability requires that, under normal circumstances, it is in our interest to be moral and our interests are rooted in our natural psychology. One thing that complicates discussion of this issue, however, is Rawls's development and refinement of his constructivist approach to moral and political philosophy. For if one takes it that that anything that merits the description 'natural' is to be discovered, not constructed, whereas moral principles are constructed and not discovered, it is going to

be difficult to make substantive connections between the moral and the natural. Hence to take this issue forward, it is necessary to consider Rawls's constructivist approach to moral and political philosophy and the role of moral psychology in this context.

2. CONSTRUCTIVISM AND MORAL PHILOSOPHY

Rawls's approach to metaethics is set out in detail in his 1975 paper 'The Independence of Moral Theory' (IMT).[8] He here develops the brief remarks in *TJ* in which, alluding specifically to Quine, he rejects the application within moral philosophy of a methodology based on the analytic/synthetic distinction which would imply giving priority to questions of definition over substantive issues of principle (*TJ* 51, 578–9; 44–5, 506–7). If anything, Rawls urges, the priority runs in the other direction: just as the advances in logical theory and the theory of meaning due to the work of Frege, Russell, and others have profoundly transformed the philosophy of logic and language, in moral philosophy, he suggests, something similar may occur: 'Once the substantive content of moral conceptions is better understood, a similar transformation may occur. It is possible that convincing answers to questions of the meaning and justification of moral judgments can be found in no other way' (*TJ* 52; 45). So insofar as Rawls has a metaethical perspective in *TJ*, it is a bottom–up rather than a top–down approach that he favours, and this thesis is explicitly affirmed in 'The Independence of Moral Theory':

A relation of methodological priority does not hold, I believe, between the theory of meaning, epistemology, and the philosophy of mind on the one hand and moral philosophy on the other. To the contrary: a central part of moral philosophy is what I have called moral theory; it consists in the comparative study of moral conceptions, which is, in large part, independent. (IMT 301)

An important element of this bottom–up approach is a willingness to engage with psychology: for psychology lies at the heart of moral theory since such a theory aims to provide

a deeper understanding of the structure of the moral conceptions and of their connections with human sensibility … We must not turn away from this task because much of it may appear to belong to psychology or social theory and not to philosophy. For the fact is that others are not prompted by philosophical inclination to pursue moral theory; yet this motivation is essential, for without it the inquiry has the wrong focus. (IMT 302)

---

[8] Reprinted in Rawls (1999*b*); page references are to this edition.

Thus moral theory has to include a psychological inquiry ('without it the inquiry has the wrong focus'); and so far from the resulting moral psychology being dependent upon philosophy of mind, the dependence runs, if anything, in the other direction. Although by and large abstract philosophical debates about mind and body do not intersect with moral theory, where there are connections, as between natural attitudes and moral sentiments, philosophy of mind has to accommodate itself to moral psychology, to our having a psychology which is not wholly value-free.

Rawls's doctrine of 'Kantian constructivism' in moral theory is to be understood in the light of this bottom–up approach to moral philosophy. The position is not an application of any more general metaphysical or epistemological doctrine concerning truth:

A constructivist view does not require an idealist or a verificationist, as opposed to a realist, account of truth. Whatever the nature of truth in the case of general beliefs about human nature and how society works, a constructivist moral doctrine requires a distinct procedure of construction to identify the first principles of justice. (KC 351–2)

Instead Rawls's constructivism is grounded in his moral theory. The central claim of this theory is that morality is a way of achieving autonomy, a life which combines respect for individual freedom, especially our status as 'self-authenticating sources of valid claims' (*PL* 72), with recognition of our essential dependence upon others who have equal status, a dependence which is not merely practical but such that we can normally realize our own conception of the good only through co-operative activities with others ('the self is realised in the activities of many selves'– *TJ* 565; 495). This involvement with others necessitates compliance with principles for social cooperation, and these principles count as moral principles only insofar as they can be viewed as principles which we and others would choose to impose upon ourselves because there are reasons for them which respect 'our status as free and equal moral persons' as Rawls frequently puts it (e.g. *PL* 19). Thus by internalizing the fact of our essential dependence upon others we recognize the social requirements of this interdependence as moral principles whose application to us is not a limitation of our autonomy, but a condition of it. So, according to this way of thinking, the substance of morality is not constituted by a set of moral facts which are available to be discovered in the social world; instead it is to be thought of as 'constructed' through agreements made in accordance with an idealized procedure for regulating social cooperation which represents both the equal status of the parties involved and the reasons which favour or oppose different policies. In *TJ* Rawls famously describes this procedure in terms of a hypothetical contract to be agreed under the conditions of an 'original position', but I

shall not discuss this matter here. My interest lies in the underlying moral psychology and its connection with his general constructivist approach.

The account above implies that the connection is very close. For, as Rawls puts it, Kantian constructivism is the doctrine that the principles which define the requirements of morality are to be 'viewed as specified by a procedure of construction … the form and structure of which mirrors both of our two powers of practical reason',[9] that is, our capacity for a conception of the good and our sense of justice. Equally, our essential dependence upon others is not just a matter of the need for practical cooperation under conditions of potential relative scarcity; instead it expresses a deep psychological truth about the conditions for personal self-realization, namely that it is dependent upon mutual interaction and appreciation by others. Thus it is our moral psychology, combined with the recognition that others with the same moral psychology have equal moral status, which sets the constraints within which the construction of morality is conceived. But it is important to note that neither our moral psychology itself nor its status is here thought of as constructed in the same way. Instead our 'two powers of practical reason' owe their status as moral powers to the ways in which they express the requirements of practical reason: the capacity for a conception of the good expresses our rationality, since it is the ability 'to form, to revise, and rationally to pursue a conception of one's rational advantage or good' (*PL* 19); and a sense of justice expresses our willingness to be reasonable, since it is the capacity 'to understand, to apply, and to act from the public conception of justice which characterizes the fair terms of social cooperation' (*PL* 19). So Rawls's constructivism in moral and political theory is founded upon a moral psychology whose status is not constructed at all but is explicated in terms of its role as the expression of practical reason.

In his later writings Rawls often compares his constructivist doctrine with the position of the rational intuitionist who holds that fundamental moral principles are discovered through a capacity for intuitive insight, and it is worth looking at the way in which Rawls makes this comparison. The familiar objections to rational intuitionism, such as Mackie's 'argument from queerness',[10] appeal to general metaphysical and epistemological considerations; in the light of Rawls's metaethical stance, however, it is not surprising that this is not the way in which he argues against the rational intuitionist. Instead his argument is rooted in his moral theory, in the importance of framing a conception of morality whereby the practice of morality can be seen to be a way of achieving autonomy, the expression of one's nature as 'free and equal', and not as a way of fulfilling requirements

---

[9] Rawls (2000: 237).   [10] Mackie (1977: 38 ff).

which, because they are external to one's own reasons for action, are 'heteronomous'. Rawls puts the point in the following way:

Yet it suffices for heteronomy that these [first] principles obtain in virtue of relations among objects the nature of which is not affected or determined by the conception of the person. Kant's idea of autonomy requires that there exist no such order of given objects determining the first principles of right and justice among free and equal moral persons. Heteronomy obtains not only when first principles are fixed by the special psychological constitution of human nature, as in Hume, but also when they are fixed by an order of universals or concepts grasped by rational intuition, as in Plato's realm of forms. (KC 345)

Thus Rawls takes it that rational intuitionism is essentially a secular version of a Divine Command theory and that moral demands cannot be in this way altogether independent of us: morality can secure its authority over us only by answering to our nature as free rational beings. This criticism is associated with further points. Although Rawls agrees with the intuitionist that morality aims to be objective, he rejects the intuitionist's inference that moral principles purport to be true since he takes it that truth and facts are inseparable and he denies that there are any moral facts because 'the idea of constructing facts seems incoherent' (*PL* 122). This point is associated with a disagreement concerning the relationship between morality and reason: whereas the intuitionist takes it that moral judgment is an exercise of theoretical reason, for the constructivist it is based upon practical reason, as expressed through our capacities for rationality and reasonableness. As a result, there is an important difference with respect to moral psychology: for the intuitionist, moral theory implies only that our psychology includes a capacity for intuitive moral insight and for motivation by the knowledge thus acquired; but because constructivism relies on the capacities through which practical reason is expressed as a basis for the construction of moral principles it involves the richer moral psychology exemplified by Rawls's account of our two fundamental moral powers, our capacity for a conception of the good and our sense of justice (*PL* 93).

In the passage quoted above Rawls also criticized Hume's moral theory ('Heteronomy obtains not only when first principles are fixed by the special psychological constitution of human nature, as in Hume'), and it is worth setting Rawls's criticisms of Hume alongside his discussion of rational intuitionism, since there is a sense in which he conceives of his Kantian constructivism as a mean between these two positions. As we have just seen, Rawls complains that his position is a kind of heteronomy. What Rawls seems to have in mind here is that Hume treats our moral sentiments as just special cases of more general sentiments whose function and application can be understood without reference to any moral concepts. Rawls sums

up his complaint as follows: 'What is distinctive of his view is that it seems to be purely psychological and to lack altogether what some writers think of as the ideas of practical reason and its authority.'[11] I shall not pursue the question of the justice of this verdict; what interests me here is the way in which Rawls damns Hume's moral theory for being too dependent on psychology while equally insisting, as against the Rational Intuitionist, on the importance of moral psychology within his Kantian moral theory. The key issue here seems to be one of reduction. For Rawls what is fundamental to moral philosophy is a conception of us as persons with the two fundamental moral powers, a capacity for a conception of the good and a sense of justice, which are in different ways exercises in practical reason. Moral theory is then supposed to show how putative moral principles are justified by explaining how they belong to the normative framework which would be adopted by a community of persons with these powers. Since the notion of a moral person with these powers is here taken to be fundamental, our moral psychology is here taken not to be derivable from more general psychological capacities and dispositions in the way that, according to Rawls, Hume seeks to achieve. Nonetheless, as we have seen earlier, Rawls himself acknowledges that our moral psychology is connected to our 'natural attitudes' in such a way that, where social conditions are appropriate, natural attitudes develop into moral sentiments. So there is a delicate issue as to how this relationship is understood—how, at the individual level there can be a relation of non-reductive dependence of moral psychology upon natural psychology. I shall come back to this in the next section; but before closing this section of the paper there is one further issue to be addressed.

Suppose a contemporary ethical non-cognitivist were to agree with Rawls in rejecting the thesis that moral attitudes are reducible to non-moral ones: would that commit such a person to being a Rawlsian Kantian? Surely not! For such a non-cognitivist, while agreeing with Rawls that moral judgments express our fundamental and irreducible moral psychology, would deny that there is any objectivity to be constructed on this basis. The disagreement here would be centred on what Rawls referred to, in the passage quoted above, as 'the ideas of practical reason and its authority'. The implication of 'Kantian Constructivism in Moral Theory' is that where the conception of a person includes the exercise of practical reason an objective morality can indeed be constructed on this basis. For Rawls the main part of this construction is, of course, 'justice as fairness'; but at that time, and earlier, Rawls took the view that this was just part of a broader construction of 'rightness as fairness' (see especially *TJ* § 18). The non-cognitivist will, in turn, have

---

[11] Rawls (2000: 50).

reasons for questioning this construction, and it is then clear enough where the main locus of this disagreement lies—namely on whether, setting aside doubts about the details of Rawls's moral and political constructions, Rawls made it plausible to hold that some such construction can be shown to be capable of objectivity, in the following sense: 'If, on the other hand, such a construction does yield the first principles of a conception of justice that matches more accurately than other views our considered convictions in general and wide reflective equilibrium, then constructivism would seem to provide a suitable basis for objectivity' (KC 354).

In *A Theory of Justice* Rawls appears confident that objectivity in this sense in attainable (*TJ* 517; 453). But in 'Kantian Constructivism' Rawls is more doubtful:

Of course, this is conjecture, intended only to indicate that constructivism is compatible with there being, in fact, only one most reasonable conception of justice, and therefore that constructivism is compatible with objectivism in this sense. However, constructivism does not presuppose that this is the case, and it may turn out that, for us, there exists no reasonable and workable conception of justice at all. This would mean that the practical task of political philosophy is doomed to failure. (KC 355–6)

On this issue, notoriously, Rawls's doubts continued to grow, at least with respect to the construction of a demonstrably objective system of morality. For he came to think that once one considers dispassionately the variety of ethical systems, including the major religions, one should acknowledge that the resolution of fundamental questions of value is underdetermined by reasonable considerations and accept the 'fact of reasonable pluralism' (*PL* 60–1). Hence in his later writings he accepts that the objectivity of morality in general is doubtful—a matter of faith rather than reasonable belief. Nonetheless Rawls retained the view that within the sphere of public political morality agreement concerning principles of justice among citizens who are both rational and reasonable is attainable; and with this he takes it that objectivity with respect to these principles is defensible (KC 115–16).

3. MORAL DEVELOPMENT

Having explored the foundational role of moral psychology in the context of Rawls's constructivist moral philosophy, I want to return to Rawls's early account of the 'principles of moral psychology', in particular his account of the development of a sense of justice. Rawls himself does not show how these two aspects of his conception of moral psychology fit together; but it

is important to consider this matter since, as I mentioned in the previous section, it is not clear how Rawls can maintain that the sense of justice is the product of a normal process of development from natural attitudes without slipping into the Humean 'psychological naturalism' he rejects.

In chapter 8 of *A Theory of Justice* Rawls argues that our sense of justice is the natural outcome of a deep tendency to reciprocity in our psychological constitution. Very briefly, Rawls's account runs as follows: initially, within the family, a young child develops the self-confidence which gives it a capacity for affection and friendship through growing up in an environment in which it feels secure in the love and care of its parents. This then helps the growing child to form friendly relationships outside the family and through these relationships the child develops a capacity for trust and responsibility as it is itself treated in these ways by others. Finally as a young adult it internalizes the requirements of justice as adherence to general principles through the experience of being treated with respect and fairness by others with whom it has no special friendship or relationship. This whole approach suggests that justice itself is a kind of reciprocity; for it is by thinking of justice in this way that one can see how the limited reciprocity of the first two stages becomes a reciprocal disposition which is a 'sense of justice' when it is extended to apply to relationships with just anyone—typically fellow citizens, but in principle strangers as well. Rawls's most famous early paper was 'Justice as Fairness' (1958) and throughout his life he used the phrase 'justice as fairness' to describe his conception of justice. But in his 1971 paper 'Justice as Reciprocity' he in fact distinguishes justice from fairness and argues that their common foundation is reciprocity, which is therefore more fundamental to justice than fairness itself: 'It is this requirement of the possibility of mutual acknowledgment of principles by free and equal persons who have not [*sic*] authority over one another which makes the concept of reciprocity fundamental to both justice and fairness.'[12]

The first issue to be addressed concerning this three-stage development is Rawls's observation, mentioned earlier, that the account 'refers to an institutional setting as being just', so that 'some view of justice enters into the explanation of the corresponding sentiment'—the sense of justice (*TJ* 491; 430). This claim has to be understood in the context of his constructivism, so that the explanatory role here of justice is not that of a distinctive fact which structures the context in which this development takes place. For, according to Rawls, there are no such moral facts. Instead Rawls's position must be that this development is accomplished in the

---

[12] 'Justice as Reciprocity' as reprinted in Rawls (1999*b*) 209. I am indebted to Patricia Greenspan for directing me to this important paper.

context of relationships informed by ties of personal affection and loyalty through which we come to see the point of fair practices which underpin the relationships from which we benefit, such as growing up within a family, sharing a house with a group of friends, or living as a citizen in a well-ordered society. Hence for Rawls it is by thinking of moral principles as principles setting terms for cooperation which would be agreed by free and equal persons that we should understand the role of social practices and institutions which incorporate these principles as providing the context for the psychological developments which issue in a sense of justice.

In *A Theory of Justice* Rawls suggests that we should think of these developments as based upon transformations in the kinds of desire that we have:

The three laws describe how our system of desires comes to have new final ends as we acquire affective ties. These changes are to be distinguished from our forming derivative desires … [Instead the laws] characterize transformations of our patterns of final ends that arise from our recognizing the manner in which institutions and the actions of others affect our good. (*TJ* 494; 432)

In his later writings Rawls writes of our capacity for 'principle-dependent' and 'conception-dependent' desires, as opposed to ordinary 'object-dependent' desires (*PL* 82 ff.), and this offers a more detailed way of thinking about his account of moral development. Object-dependent desires are desires whose objects are personal goods whose characterization as such relies on no moral or other normative principle; by contrast specifying the objects of principle-dependent desires such as fidelity involves moral principles, and similarly specifying the objects of conception-dependent desires involves moral ideals such as citizenship. Can one match these three types of desire with the three stages? The match is easy to see at stages two and three: the growing child who becomes trustworthy through shared activities with friends can be thought of as someone who begins to develop principle-dependent desires; and for Rawls the final development of a sense of justice is accomplished as one identifies oneself as a member of a potentially well-ordered society in which one can aspire to the ideal of citizenship. The first stage, however, is not to be thought of as that in which one becomes susceptible to object-dependent desires, since these are all too prevalent anyway. Instead what is important at this first stage is, I think, the development of a capacity to care about others, to make their good one's own good, since this is a prerequisite of the capacity for friendship which enters into Rawls's second stage. Initially, of course, the others in question are those who care about one themselves, most notably the members of one's family. So, at least for the purpose of completing the match between Rawls's three-stage account of moral development and his three-way hierarchy of
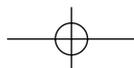
desires, Rawls's hierarchy of desires needs to be augmented by inserting a category of 'relationship-dependent desires' between the object-dependent and the principle-dependent desires. Object-dependent desires come for free, and do not mark a significant stage of moral development; and the first stage of development requires, not principle-dependent desires, but those which involve concern for others to whom the subject, typically a child, is connected by an affectionate relationship such as the love between members of a family.

In setting things out in this way I have been trying to explore the way in which a Rawlsian moral psychology might be thought to work. The sketch of moral development above does, I think, meet the twin requirements of neither tacitly drawing on rational intuitionism, the intuitive appreciation of moral truths as such, nor turning out to be a form of reductive naturalism which derives moral sentiments from non-moral natural attitudes. For although the account of moral development involves a hierarchy of desires, starting from non-moral object-dependent desires, the progression is achieved though transformations in which the subject's motivational set is thought of as enhanced as a result of including concerns, principles and ideals which the subject recognizes as informing relationships, practices and institutions that are essential to his own good as he grasps a broader sense of his own identity as a member of groups in which his own good is dependent upon that of others and vice-versa. Thus these transformations exemplify Rawls's belief that the moral psychology appropriate to Kantian constructivism is founded upon the exercise of practical reason, understood as a rational appreciation of one's own good and a willingness to cooperate with others on a reasonable basis.

4. THE PROBLEM OF STABILITY

So far, then, so good for this sketch of the way a Rawlsian moral psychology might work. But as I indicated at the start an important further consideration comes from the role that this psychology is supposed to fulfil in Rawls's theory of justice by contributing to a solution to the problem of stability. This is the problem of showing that a state whose political institutions are just will be reasonably stable in the sense that such a state need not rely primarily on coercion to ensure its citizens obey the law and support the state's institutions because, for the most part, these institutions and laws enjoy the support of the citizens anyway. The contribution of moral psychology to the solution of this problem was supposed to be that it would show how the requirements of justice are broadly congruent with the interests of individual citizens, so that, as Rawls put it, 'being a good

person is a good thing for that person' (*TJ* 397; 349); and this requires, as one might put it, that the moral psychology which makes a person 'a good person' be congruent with the natural, normal, psychology which identifies what is 'a good thing for that person'.

Rawls's first thought, in his 1963 paper 'The Sense of Justice', was that merely by showing that the development of a sense of justice is the normal psychological outcome of life within a society whose institutions are broadly just one shows that such a society is stable.[13] For there is a virtuous circle whereby institutions and moral sentiments reinforce each other. In *TJ* Rawls starts by repeating this line of thought in chapter 8, arguing that the moral psychology of his conception of justice as fairness is more conducive to stability than, say, the psychology that one would associate with a utilitarian conception of justice (*TJ* § 76 'The Problem of Relative Stability'). But he does not treat this comparison as the end of the matter, since only a few pages later, at the start of chapter 9, he says that only now is he in a position to deal properly with the task of showing that justice and goodness are congruent (*TJ* 513; 450). As he acknowledged later, commenting on the part of *A Theory of Justice* which includes these chapters, it is not clear what is going on here: 'Throughout Part III too many connections are left for the reader to make, so that one may be left in doubt as to the point of much of chs. 8 and 9'.[14] It appears nonetheless that in *A Theory of Justice* Rawls believed that he could not simply rely on the normal development of a sense of justice to vindicate the congruence thesis for the reason that this account of the development of a sense of justice was primarily causal and not normative: it showed how one would expect a sense of justice to be inculcated among those growing up in a just society; but it did not thereby show that it was good for them to have this motivation which might be just a 'neurotic compulsion' (*TJ* 514; 451). And without a demonstration of this, the problem of stability was not fully resolved, since if it remained an open question whether acting justly was in general good for one, one could not reliably expect people to obey the laws of a just state despite their having a sense of justice.

I think that Rawls was right about this; but that Rawls's own favoured account of the congruence of individual good and morality is unpersuasive. He argues that because 'the desire to act justly and the desire to express our nature as free moral persons turn out to specify what is practically speaking the same desire' (*TJ* 572; 501) fulfilment of this desire promotes one's own good. This thesis assumes the dominance of his Kantian moral psychology ('the desire to express our nature as free moral persons') among

---

[13] 'The Sense of Justice', in Rawls (1999*b*: 105).
[14] 'Justice as Fairness: Political not Metaphysical', in Rawls (1999*b*: 414 n. 33).

the ends which determine one's individual good. Yet the practical doubt which motivates the search for congruence can be readily redirected at the question as to whether the ends intrinsic to this Kantian moral psychology are indeed central for one's own individual good. In chapter 7 of *A Theory of Justice* Rawls had set out a complex account of goodness, which includes an account of a person's good as 'the successful execution of a rational plan for life' (*TJ* 433; 380).[15] So the congruence thesis is the thesis that the successful execution of a rational plan for life normally requires one to act in accordance with the principles of justice. Rawls seems to have thought that the fact that this latter requirement is tantamount to acting in accordance with the motivations specified by his Kantian moral psychology as 'expressing one's nature as a free person' *ipso facto* shows that it contributes to the successful execution of one's rational plan for life. But as it stands this is not persuasive: for why should the aim of 'expressing one's nature as a free person' be given a central position in one's individual plan for life? The introduction here of what in *A Theory of Justice* he calls the 'Kantian interpretation' (§ 40) of moral psychology does not by itself suffice to make a connection between adherence to the principles of justice and individual good. If Rawls were to introduce a Kantian dualism of noumenal and empirical selves, it would be plausible to hold that for the noumenal self individual good and the moral life are inseparable; but Rawls explicitly rejects any such dualism, and, anyway, it would leave the problem completely insoluble as far as empirical selves are concerned, which is where it matters most.

Notoriously, there is a sense in which Rawls himself came to agree about this. For in the 1992 'Introduction' to *Political Liberalism* (esp. pp. xvi–xvii) he explains that he himself came to see that his Kantian argument for the congruence of goodness and justice was unsatisfactory as a general solution to the problem of stability since his Kantian moral theory was but one of several reasonable comprehensive moral theories. As a result, he inferred, no one moral theory can be employed in a political philosophy which aims to provide arguments that will be persuasive for all reasonable citizens, and it had been a mistake to rely on the Kantian theory alone to solve the problem of stability. This last move is not unchallengeable: one might think that insofar as the problem is not solved by explaining how a just society nurtures a sense of justice among its citizens, the problem of stability is essentially theoretical and thus that one cannot expect to avoid drawing on one's moral theory to resolve it—even while recognizing that there are other 'reasonable moral theories' which will promote different solutions.

---

[15] So Rawls combines a broadly naturalistic account of goodness with his constructivist account of morality.

But Rawls thinks that a different line of thought is available: if one can show that there is a conception of justice which expresses the reasonable political aspirations of adherents of different moral theories who acknowledge the fact of reasonable pluralism, one will thereby be in a position to show that a state which realizes this conception should be stable, since 'it can win its support by addressing each citizen's reason, as explained within its own framework' (*PL* 143).

To follow this line of thought would take me well away from moral psychology. So having noted how the problem of stability continues in Rawls's later writings, I want to return to the problem of stability as he left it in *A Theory of Justice*. For in his discussion there he introduced some suggestive ideas which remain largely unexploited by him but which, I think, can be used not only to provide a better solution to the problem of stability than the approach he favoured but also to enrich his account of moral psychology.

5. SELF-RESPECT AND MUTUAL RESPECT

The line of thought I have in mind is that which Rawls introduces in chapter 9 of *A Theory of Justice*, starting with the conception of a 'social union' which he takes from Humboldt (*TJ* 523–5; 459–60: esp. n. 4). A social union is a collective institution whose members cooperate in a type of joint activity in order to achieve valuable ends which they cannot bring about without such cooperation. Rawls gives the example of an orchestra as a social union of this kind: for it is only within an orchestra which brings together musicians of many different kinds that the individual musicians can take part in performing great orchestral works. In cases of this kind, he writes, 'persons need one another since it is only in active cooperation with others that one's powers reach fruition. Only in a social union is the individual complete' (*TJ* 524–5; 460 n. 4). The existence of social unions shows us something important about the way in which individuals with different abilities need to collaborate with each other in order to achieve valuable ends, and Rawls infers that a just state can itself be regarded as a social union, a 'social union of social unions' (*TJ* 527; 462). This involves more than just the familiar thesis that political cooperation is essential for the achievement of individual goods: instead, if the social union model is to be applicable there has to be some collective good comparable to the performance of a symphony which is not available without the collective participation of the citizens who are members of this social union, the just state, and which is central to the ends of each individual citizen. Rawls suggests that the collective good is just 'the public realization of

justice' which meets the requirement of providing a distinctive form of self-fulfilment because 'the collective activity of justice is the pre-eminent form of human flourishing' (*TJ* 529; 463).

This suggestion is not plausible. There are indefinitely many forms of human flourishing, arising from the great variety of individual plans for life, and although 'the collective activity of justice' is an ingredient in many worthwhile ends, this is no reason to give it pre-eminence as an end itself; indeed it is questionable whether it really makes sense to regard 'the collective activity of justice' as a form of human flourishing. What is going wrong here is that the social union model for the state does not really work for Rawls: the Rawlsian liberal state is not comparable to an orchestra, an institution whose members rely on each other's complementary activities to accomplish an essentially collective goal such as the performance of a symphony. Rawls's suggestion that 'the public realization of justice' counts as such an end, for example, is unpersuasive. For although individual citizens are of course required to be just, this is not a collective activity on their part and it is primarily the responsibility of public authorities to maintain justice in general. If Rawls were to hold with the communitarians that the state is a collective association with some dominant goal that supposedly meets the requirement of providing self-fulfilment for all citizens, such as the establishment of a classless society, he could use the social union model for the purposes of his congruence thesis. But, of course, that is exactly not the way in which Rawls conceives of his liberal state.[16]

Yet one should not for this reason dismiss altogether all of the themes that enter into Rawls's discussion of the idea of a social union, in particular the suggestion that

the members of a community participate in one another's nature: we appreciate what others do as things which we might have done which they do for us, and what we do is similarly done for them. Since the self is realized in the activities of many selves, relations of justice that would be assented to by all are best fitted to express the nature of each. (*TJ* 565; 495)

Rawls's line of thought here is reminiscent of the kind of reciprocity that came up earlier in connection with his account of our moral development, the three-stage development of a sense of justice via the place of love in 'the morality of authority' and that of trust in 'the morality of association'. Rawls never connects this conception of reciprocity that is central to his early moral psychology with his later discussion of congruence; but I want to propose that there are connections to be made here which enable one to fill out both his moral theory and his moral psychology. The place to start

---

[16] Rawls's rejection of civic humanism is especially notable in this context: see *PL* 206.

is with the good which is for Rawls of primary importance: self-respect. He writes

It is clear then why self-respect is a primary good. Without it nothing may seem worth doing, or if some things have value for us, we lack the will to strive for them. All desire and activity becomes empty and vain, and we sink into apathy and cynicism. Therefore the parties in the original position would wish to avoid at almost any cost the social conditions that undermine self-respect. (*TJ* 440; 386)

Suppose we now apply to self-respect the developmental approach involving reciprocity from Rawls's moral psychology, so that self-respect is held to be dependent upon respect by others whom one respects oneself. It now becomes easy to argue for the congruence of justice and individual good. The argument starts from Rawls's thesis about the value of self-respect:

(i) Any rational plan for life will acknowledge that self-respect is a primary good.

Add my proposal about the dependence of self-respect on respect by others whom one respects oneself:

(ii) The achievement of self-respect is dependent upon reciprocal relationships of mutual respect.

Now add a Rawlsian thesis about justice as reciprocity:

(iii) The conception of justice as reciprocity is the conception of principles whose institutional realisation would affirm the mutual respect of citizens for each other.

It does now follow that

(iv) Any rational plan for life will bring with it a 'conception-based desire' to living in accordance with justice, at least in a well-ordered society.

The crucial claim of this argument is (ii), that self-respect is dependent upon reciprocal relationships of mutual respect. Rawls himself endorses a thought of this kind when he writes that our sense of our own worth is supported by 'finding our person and deeds appreciated and confirmed by others who are likewise esteemed' (*TJ* 440; 386); and he goes on to infer something comparable to thesis (iii) from this, to the effect that the realization of his principles of justice provides 'background conditions' which ensure that 'in public life citizens respect one another's ends' (*TJ* 442; 388). But although he here (*TJ* 442; 388) intimates that he will return to this thesis in his subsequent discussion of the idea of a social union, in that context he does not in fact make any significant use of it. Instead he advances the idea of the state as a social union of social unions, a proposal

which, as I have explained, he cannot adequately substantiate. Yet although the argument from (i) to (iv) is not manifest in *A Theory of Justice* (or elsewhere), it would, I think, be congenial to Rawls.

In thinking about the crucial claim (ii) it is important first to clarify what self-respect amounts to. As the passages cited above indicate, in *A Theory of Justice* Rawls treated self-respect and self-esteem as interchangeable, but this is readily seen not to be correct when one thinks of the difference between behaviour which shows a lack of esteem for someone's work and that which shows a lack of respect for them.[17] To treat someone with a lack of respect is, I take it, to fail to acknowledge their status as (in Rawls's words) a 'self-authenticating source of valid claims', whereas a lack of esteem for someone simply expresses the judgment that their life and work is not especially valuable. In this sense, therefore, self-respect is consciousness of oneself as a self-authenticating source of valid claims, as someone who merits treatment with respect by others; whereas self-esteem is the judgment that one's life includes valuable achievements that are worthy of esteem by others. I take it that both self-respect and self-esteem are important goods. Rawls's description (quoted above) of the situation of someone who lacks self-respect in fact applies best to the case of someone who lacks self-esteem: for someone who lacks self-respect is not so much someone who thinks that nothing is worth doing as someone who thinks that he is worthless, someone whose interests count for nothing. No doubt these two conditions are closely associated: self-esteem, I think, presupposes self-respect, though the converse implication need not obtain (someone who is excessively modest lacks self-esteem but not necessarily self-respect). But it is important for the purposes of the current argument to distinguish them; for (iii) is only plausible when interpreted as a claim about self-respect, properly understood. It is not a requirement of justice that people should esteem each other's life and work.

The key issue, therefore, is whether (ii) is also plausible when self-respect is interpreted as consciousness of oneself as a source of valid claims on others. Where the dependence affirmed by (ii) is understood as a case of normal reciprocal psychological development, of the kind characterized by Rawls in his description in *A Theory of Justice* of the 'principles of moral psychology', (ii) certainly looks to be plausible; indeed it is surely integral to the moral development Rawls describes. So, understood in this way, (ii) is as robust as the rest of Rawls's early moral psychology. But, as we have seen, Rawls hoped that the congruence thesis could be established in a way which did not just rely on the normal course of human psychological

---

[17] See David Sachs, 'How to Distinguish Self-Respect from Self-Esteem', *Philosophy and Public Affairs* 10/4 (autumn, 1981), 346–60.
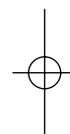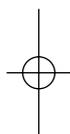
development. So the question is whether there is a way of strengthening the form of dependence in (ii).

One can envisage a stronger way of interpreting (ii), as affirming that self-respect constitutively requires mutually interpenetrating attitudes of respect such that one recognizes that one is respected by others whom one respects oneself. To take this view of self-respect would be to model it on Hegel's famous thesis concerning self-consciousness, that self-consciousness is dependent upon the consciousness of one by others of whom one is oneself conscious: 'Self-consciousness exists in and for itself when, and by the fact that, it so exists for another; that is, it exists only in being acknowledged' (*Phenomenology of Spirit* § 178). where this 'acknowledgment' takes the form of mutual 'recognition: 'They *recognize* themselves as *mutually recognizing* each other (*Phenomenology of Spirit* § 184). Just what Hegel's conception of self-consciousness amounts to is notoriously obscure and disputed, and I shall not attempt to elucidate it; what matters for us is whether this thesis holds for self-respect.[18] On the face of it it is vulnerable to counter-examples: a good case to think about is that of Olaudah Equiano, the remarkable slave whose autobiography *The Interesting Narrative of the Life of Olaudah Equiano* clearly shows how he maintained his self-respect in the face of a failure of recognition by others who bought and sold him as a slave. This case shows that a straightforward Hegelian interpretation of (ii) is too strong; and of course without (ii), the route via (iii) to (iv), the congruence thesis, is broken. But there is a way around this, by taking it that self-respect is to be understood precisely in such a way that (ii) is true of it—that is, by taking it to be the kind of publicly affirmed self-respect in which one's sense of one's own worth is confirmed and strengthened through recognition by others whom one respects. Olaudah Equiano did not enjoy this kind of self-respect until he was able to buy his way out of slavery and work with others for the abolition of slavery; but there is every reason to think that this change in his self-consciousness was a change of great value to him. For once his situation had changed his own sense of himself as a 'self-authenticating source of valid claims' was at last confirmed by the recognition of the validity of these claims by others whose similar status he himself recognized. So even though his initial form of private self-respect was of great value to him, the primary social good in this area is the kind of publicly confirmed self-respect which satisfies condition (ii);

---

[18] In the original version of this paper I tried to use the position presented by Axel Honneth in *The Struggle for Recognition* (Honneth 1995) to develop this line of thought since in some respects his position resembles that advanced by Rawls. But discussion with Carla Bagnoli has persuaded me that it is both unnecessary and confusing to introduce Honneth's position, interesting though it is.

hence proposition (i) holds for this form of self-respect, and since (iii) is plainly also in play, the route to (iv) and the congruence of justice and individual good, is secured.

It may be felt that there is a trick here, in that self-respect has been just defined to be a condition which depends on social recognition. One response to this is to observe that as long as this form of self-respect is agreed to be a primary social good it does not matter that there is another form of self-respect which is not in the same way dependent on recognition. But there is a deeper point here. I have characterized self-respect in Rawlsian terms as consciousness of one's freedom since he takes it that freedom is primarily a matter being a self-authenticating source of valid claims (*PL* 72). For Rawls this freedom depends on one's moral powers, primarily the capacity to be guided by one's conception of the good and to revise this conception in the light of evidence. So freedom is an implication of the conception of a person that is characteristic of Kantian moral psychology, as the case of Olaudah Equiano indicates, since he certainly possessed the relevant moral powers even when he was a slave. The Hegelian move is then to suggest that the consciousness of freedom that comes with self-respect takes us beyond moral psychology because it involves recognition by others. Equiano's case shows that the necessity for this transformation is questionable: in his case self-respect did not, initially, involve recognition by others. But it is important to note that the 'claims' whose self-authenticating validity is affirmed in the attribution of freedom to a person are claims directed at others, with the presumption that their validity is to be recognized by them. So even within Rawls's conception of freedom there is a presumption of recognition by others; and it is this presumption which is then made explicit in the Hegelian account of self-respect as a consciousness of freedom which is dependent upon recognition by others. As we have seen, this suggestion needs qualification; there can be a form of self-respect which is not dependent on actual recognition by others. But since this form of self-respect still makes a claim to recognition and respect by them, there is every reason to think that it is better to enjoy the socially confirmed form of self-respect than the private consciousness of freedom which was Equiano's lot for most of his life. Hence the priority given to the socially confirmed form of self-respect is not a dialectical trick, but is inherent in the conception of self-respect itself.
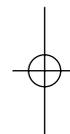
REFERENCES

Honneth, A. (1995) *The Struggle for Recognition* (Cambridge: Polity).

Kukathas, C. (ed.) (2002) *John Rawls (Critical Assessments)* 4 vols. (London: Routledge).

Mackie, J. L. (1977) *Ethics: Inventing Right and Wrong* (Harmondsworth: Penguin).
Rawls, J. (1971) *A Theory of Justice* (Cambridge, MA: Harvard University Press).
—— (1993) *Political Liberalism* (New York: Columbia University Press).
—— (1999*a*) *A Theory of Justice*, rev. edn (Oxford: Oxford University Press).
—— (1999*b*) *Collected Papers*, ed. S. Freeman (Cambridge, MA: Harvard University Press).
—— (2000) *Lectures on the History of Moral Philosophy*, ed. B. Herman (Cambridge, MA: Harvard University Press).
—— (2001) *Justice as Fairness* (Cambridge, MA: Harvard University Press).