11

Of Gods and Demons: Habermas and Practical Reason

'Practical questions', according to Habermas, 'admit of truth';[1] 'just (*richtige*) norms must be capable of being grounded in a similar way to true statements'.[2] Truth, on his view, means 'warranted assertibility': this is shown when participants enter into a discourse and 'a consensus can be realized under conditions that identify it as a justified consensus'.[3] If, he writes,

philosophical ethics and political theory are supposed to disclose the moral core of the general consciousness and to *reconstruct* it as a normative concept of the moral, then they must specify criteria and provide reasons: they must, that is, produce theoretical knowledge.[4]

Thus for Habermas judgements about moral and political questions can be rationally grounded and differences about such questions can be rationally resolved.

His position thereby contrasts with those of noncognitivists, moral

This chapter was first published in 1982. Prof. Habermas has replied to some of the arguments advanced here, in the concluding chapter to *Jürgen Habermas: Critical Debates* (London: Macmillan, 1982).

1 Habermas, *Legitimation Crisis*, trans. T. McCarthy (Boston: Beacon Press, 1975; London: Heinemann, 1976).

2 Id., 'Wahrheitstheorien', in *Wirklichkeit und Reflexion: Walter Schulz zum 60. Geburtstag*, ed. H. Fahrenbach (Pfullingen: Neske, 1973), p. 226. 'I suspect', he adds, 'that the justification of the validity claims contained in the recommendation of norms of action and of evaluation can be just as discursively tested as the justification of the validity claims implied in assertions. Of course, the grounding of just (*richtigen*) commands and evaluations differs in the structure of argumentation from the grounding of true statements. The logical conditions under which a rationally motivated consensus can be attained differ as between practical and theoretical discourse' (ibid, 226–7). Habermas uses the term '*richtig*' in such a way as to imply that there is a truth of the matter as to whether norms, commands, or evaluations are *richtig* or not.

3 Ibid. 239–40.

4 Habermas, *Communication and the Evolution of Society*, trans. T. McCarthy (Boston: Beacon Press, and London: Heinemann, 1979, pp. 202–3.

sceptics, subjectivists, relativists, and pluralists of various kinds, and it relies upon a conception of rationality that is sufficiently comprehensive to allow rational solution to such perennial and inherently controversial questions as: Which rules, laws, distributive arrangements, etc., are just? Is the state's claim to legitimacy valid or empty? What is the scope of legitimate authority? And so on.[5] Thus he seeks to 'vindicate the power of discursively attained, rational consensus against the Weberian pluralism of value systems, gods and demons', speaking disdainfully of the 'empiricist and/or decisionist barriers, which immunize the so-called pluralism of values against the efforts of practical reason'.[6] Habermas's rationalism has, moreover, a distinctively Hegelian dimension. He postulates the possibility of society reaching a stage of transparent self-reflection, among parties who are 'free and equal' and whose discourse has reached a stage where 'the level of justification has become reflective', in the sense that mythological, cosmological, religious and ontological modes of thought have been superseded and 'rational will-formation' can be achieved, free of dogmas and 'ultimate grounds', through ideal mutual self-understanding.[7]

Within the Marxist tradition, Habermas's position is a distinctive one. By and large, Marxists have been dismissive, even contemptuous, of morality, for reasons examined in Part III of this volume, and relatively uninterested in the problem of justifying norms and normative judgements. Among those who considered the question, Engels, Kautsky, and Trotsky saw no need for such justification, though speaking of a 'really human morality which stands above

[5] Habermas, *Communication and the Evolution of Society*, p. 200. On this point see R. J. Bernstein, *The Restructuring of Social and Political Theory* (New York: Harcourt Brace Jovanovich, 1976).

[6] Habermas, *Legitimation Crisis*, p. 107. The passage by Max Weber, to which Habermas refers, is the following: 'What man will take upon himself the attempt to 'refute scientifically' the ethic of the Sermon on the Mount? For instance, the sentence "resist no evil", or the image of turning the other cheek? And yet it is clear, in mundane perspective, that this is an ethic of undignified conduct; one has to choose between the religious dignity which this ethic confers and the dignity of manly conduct which preaches something quite different; "resist evil—lest you be co-responsible for an overpowering evil". According to our ultimate standpoint, the one is the Devil and the other the God, and the individual has to decide which is God for him and which is the Devil. And so it goes throughout all the orders of life' (*From Max Weber: Essays in Sociology*, ed. and trans. H. H. Gerth and C. Wright Mills (London: Routledge & Kegan Paul, 1948), p. 148). For a no less eloquent statement of the same viewpoint see Leszek Kolakowski's 'Ethics without a Moral Code', *Triquarterly* 22 (1971), esp. pp. 172–4.

[7] Habermas, *Communication*, pp. 186, 185, 184.

class antagonisms and above any recollections of them',[8] of an unfolding 'general human morality',[9] and of the 'liberating morality of the proletariat'.[10] It was the neo-Kantian Marxists and the Austro-Marxists who explicitly raised the question of such justification, heretically distinguishing between facts and values and offering transcendental arguments purporting to justify the struggle for socialism in terms of universal values. It was, however, Lukács who developed the position that arguably accords most closely with Marx's own view on this subject: that from the privileged standpoint of the proletariat, engaged in the revolutionary transformation of the world, the subject and object of history are united, as the process of history becomes identical with the free development of conscious-ness and, accordingly, moral judgement becomes identical with the self-understanding of the universal class as it destroys the old world to create a new one. On this view the norms and institutions thereby created require no further justification: they are constitutive of the various ascending 'stages' of 'truly human' society.[11]

Habermas makes none of these moves, proposing instead, as an inheritor of the tradition of critical theory, to develop a mode of theorising that is grounded, in Max Horkheimer's phrase, by an interest in the future, and in particular by 'the idea of a reasonable organization of society that will meet the needs of the whole com-munity' as a goal of human activity which is 'immanent in human work, but . . . not correctly grasped by individuals or by the common mind'.[12] However, unlike his predecessors in this tradition, Haber-mas makes a serious attempt to give content and grounding to the key notion of emancipation, seeing it as immanent not in work but in communication. His central idea is that

the *design* of an ideal speech situation is necessarily implied in the structure of potential speech, since all speech, even intentional deception, is oriented toward the idea of truth. This idea can be analyzed with regard to a consensus achieved in unrestrained and universal discourse. Insofar as we master the means for the construction of the ideal speech situation, we can

[8] F. Engels, *Anti-Dühring* (Moscow: Foreign Languages Publishing House, 1959).

[9] Karl Kautsky, *Ethics and the Materialist Conception of History*, trans. J. B. Askew (Chicago, Mass.: Charles H. Kerr & Co., n.d.), p. 160.

[10] L. Trotsky, *Their Morals and Ours*, 4th edn. (New York: Pathfinder Press, 1969), p. 37.

[11] For this interpretation of Lukács and Marx, see Leszek Kolakowski, *Main Currents of Marxism*, 3 vols. (Oxford: Clarendon Press, 1978).

[12] Horkheimer, *Critical Theory* (New York: Seabury Press, 1973), p. 213.

conceive the ideas of truth, freedom and justice, which interpenetrate each other—although of course only as ideas.[13]

His whole theory of universal pragmatics is devoted to establishing these claims and I cannot directly consider them here.[14] I shall concentrate rather on Habermas's claim to have established a vantage-point from which the social world can be critically analysed and from which one can identify ideological deception and normative power—forms of domination whose legitimacy is imposed and which rely on 'contingent and forced consensus', on 'preventing questions that radicalize the value-universalism of bourgeois society from even arising'.[15]

Let us, then, look more closely at Habermas's recent attempt to specify that critical standpoint. A social theory critical of ideology can, he writes,

identify the normative power built into the institutional system of a society only if it starts from the *model of the suppression of generalizable interests* and compares normative structures existing at a given time with the hypothetical state of a system of norms formed, *ceteris paribus*, discursively. Such a counterfactually projected reconstruction . . . can be guided by the question (justified, in my opinion, by considerations from universal pragmatics): how would the members of a social system, at a given stage in the development of productive forces, have collectively and bindingly interpreted their needs (and which norms would they have accepted as justified) if they could and would have decided on organization of social intercourse through discursive will-formation, with adequate knowledge of the limiting conditions and functional imperatives of their society?[16]

Such ideal discourse would be solely concerned with discussing the 'bracketed validity claims' (to intelligibility, truth, rightness, and sincerity) of participants' speech-acts and its sole goal would be to test them, so that

no force except that of the better argument is exercised, and that, as a result, all motives except that of the cooperative search for truth are excluded. If under these conditions a consensus about the recommendation to accept a norm arises argumentatively, that is, on the basis of hypothetically pro-

13 Habermas, 'Toward a Theory of Communicative Competence', *Inquiry*, 13 (1970), p. 372.
14 They are admirably treated in John B. Thompson's essay in J. R. Thompson and D. Held (eds.), *Jürgen Habermas: Critical Debates* (London: Macmillan, 1982).
15 Habermas, *Communication*, pp. 188, 198.
16 Id., *Legitimation Crisis*, p. 113.

posed, alternative justifications, then this consensus expresses a 'rational will'. Since all those affected have, in principle, at least the chance to participate in the practical deliberation, the 'rationality' of the discursively formed will consists in the fact that the reciprocal behavioral expectations raised to normative status afford validity to a *common* interest ascertained *without deception*. The interest is common because the constraint-free consensus permits only what *all* can want; it is free of deception because even the interpretations of needs in which *each individual* must be able to recognize what he wants become the object of discursive will-formation. The discursively formed will may be called 'rational' because the formal properties of discourse and of the deliberative situation sufficiently guarantee that a consensus can arise only through appropriately interpreted, *generalizable* interests, by which I mean needs *that can be communicatively shared*. The limits of a decisionistic treatment of practical questions are overcome as soon as argumentation is expected to test the generaliz*ability* of interests, instead of being resigned to an impenetrable pluralism of apparently ultimate value orientations (or belief-acts or attitudes).[17]

In what follows I shall examine this attempt to establish a vantage-point that purports to yield a rational basis for critical theory in order to see whether it succeeds in providing a determinate notion of emancipation (from ideology, imposed legitimacy, forced consensus, etc.) and whether it fulfils its promise of eliminating a 'decisionistic' treatment of practical questions and avoiding 'an impenetrable pluralism of apparently ultimate value orientations'.

There is, to begin with, much to be said for the practical implications of Habermas's general approach. He maintains, correctly, that any serious social analysis—and certainly a Marxist or critical theorist—must address the question: are social norms which claim legitimacy genuinely accepted by those who follow and internalize them, or do they merely stabilize relations of power? His central insight is that, in asking this (exceedingly complex) question, one must reason counterfactually and engage in a complicated thought-experiment—albeit guided by general theoretical considerations and relevant empirical evidence—in order to determine whether forms of power, manipulation, mystification, etc., are at work, shaping and deflecting the beliefs and preferences of actors in such a way as to preclude them from thinking and acting as they would otherwise autonomously do. In order to determine this, Habermas argues, we need to postulate undistorted communication among all affected by the norms in

17 Ibid. 107–8.

question, in which they articulate their needs and in which they both form and discover their interests and the norms that they can rationally accept as binding. Thus he writes that

only communicative ethics guarantees the generality of admissible norms and the autonomy of acting subjects, solely through the discursive redeemability of the validity claims with which norms appear. That is, generality is guaranteed in that the only norms that may claim generality are those on which everyone affected agrees (or would agree) without constraint if they enter into (or were to enter into) a process of discursive will-formation.[18]

Habermas here reveals his firm commitment to the view (fundamentally at odds with any Leninist or Lukácsian assumption of privileged access in the imputation of interests) that people are the sole judges of their own interests, which are formed and discovered through dialogue on the part of all concerned—a political commitment to opening up public, democratic processes which Habermas has elsewhere described as 'the conversation of citizens'.[19] Contrary to the assumptions of technocrats, he has portrayed 'the depoliticization of the mass of the population and the decline of the public realm as a political institution' as 'components of a system of domination that tends to exclude practical questions from public discussion': the 'enlightenment of political will', he argues, 'can become effective only within the communication of citizens'.[20] His general approach leads him, therefore, to see democracy, not as any particular institutional form, but as a 'self-controlled learning process', in which the problem is to find 'arrangements which can ground the presumption that the basic institutions of society and the basic political institutions would meet with the unforced agreement of all those involved, if they could participate, as free and equal, in discursive will-formation'.[21]

But, aside from these admirable political implications of his approach, it remains to examine exactly *how* he proposes to establish a determinate rational basis for social criticism, an 'Archimedean point', as John Rawls might say, 'for assessing the social system' that 'is not at the mercy, so to speak, of existing wants and interests'.[22] How does Habermas propose to specify which norms and claims to

[18] Habermas, *Legitimation Crisis*, p. 89.
[19] Id., *Towards a Rational Society: Student Protest, Science and Politics*, trans. J.J. Shapiro (Boston: Beacon Press, 1970; London: Heinemann, 1971).
[20] Ibid.
[21] Id., *Communication*, p. 186.
[22] Rawls, *A Theory of Justice* (Oxford: Clarendon Press, 1972), p. 261.

legitimacy are capable of securing the rational consent of social actors—actors to a greater or lesser extent caught up in structures of 'normative power', 'illegitimate domination' that 'meets with consent', and 'contingent and forced consensus'?[23]

Habermas offers as his answer the 'model of the suppression of generalisable interests'—a hypothesis or 'counterfactually projected reconstruction' specifying of all the actors in question which of their interests are 'generalisable', how they would interpret their needs, and which norms they would accept as justified under conditions of unconstrained communication. It is true that Habermas observes that the 'social scientist can only hypothetically project this ascription of interests; indeed a direct confirmation of this hypothesis would be possible only in the form of a practical discourse among the very individuals or groups involved'. It can, he maintains, be indirectly confirmed (it would perhaps be better to say supported) through 'empirical indicators of suppressed interests'. Following Claus Offe (who likewise seeks a 'critical standard' for identifying structures that perpetuate 'suppressed, that is latent, claims and needs'), Habermas mentions specifically (1) the existence of an observed discrepancy between legal norms and actual legal practices; (2) codified rules which systematically exclude claims from the political agenda (claims which thus express suppressed interests); (3) the existence of a discrepancy between claims that are made and the level at which they are politically allowed satisfaction; and (4) comparative evidence, drawn from different political systems, which indicates, *ceteris paribus*, which possibilities are actualized when putatively repressive structures are absent or removed.[24] But before we can know whether the hypothesis of suppressed generalizable interests is either directly confirmed or indirectly well supported, we must first know what the hypothesis *is*. The prior issue is to establish whether such a hypothesis can be intelligibly advanced and what precisely it amounts to.

Two deep and connected difficulties are raised by Habermas's proposal. First, just how counterfactual is the hypothesis in question? Just how much of the real past and present are we to alter in the proposed thought-experiment? More particularly, to what extent, if

[23] Habermas, *Legitimation Crisis*, p. 111; Id., *Communication*, pp. 202, 188.
[24] Id., *Legitimation Crisis*, p. 114.

at all, is the identity of the agents preserved across the transition between the actual and the imagined possible world?

There appear to be three possible answers to this last question, to each of which Habermas seems to be drawn, each of which is unsatisfactory, and none of which suits his purposes. The first is that the *actual* agents are to be imagined in conditions of undistorted communication (thus he speaks of the 'interests that would have to find expression among those involved if they [sic] were to enter into practical discourse').[25] The second is that *typical* or *representative* actors are to be imagined in such conditions (thus he speaks of 'representatively simulated discourse' between groups with conflicting interests and also describes the discoursing partners as 'the members of a social system, at a given stage in the development of productive forces').[26] The third is that ideally rational, theoretically defined agents are to be imagined in such conditions (thus he speaks of the 'ideal speech situation', the participants in which must presumably be ideally constituted, having reached a level of cognitive and moral development that will enable rational consensus to be reached). The first answer suggests a counterfactual of the form 'Under (counterfactual) conditions C, the real-life actors A_1 would agree on X.' The second suggests a counterfactual of the form 'Under conditions C, typical actors A_2 (defined by, say, their membership in conflict groups or their incumbency of roles) would agree on X.' The third suggests a counterfactual of the form 'Under conditions C, ideally rational actors A_3 (defined by a theory of rationality) would agree on X.'

The problem with the first answer is that no reason is given for supposing that the actual agents would, under the conditions supposed (that is, where there is 'a symmetrical distribution of chances to select and employ speech acts and an equal opportunity to assume dialogue roles'),[27] reach the required rational consensus. Indeed, there is surely every reason to suppose that *they* would not, since *they* would continue to exhibit all kinds of traits conducive to 'distorted communication'—prejudices, limitations of vision and imagination, deference to authority, fears, vanities, self-doubts, and so on. Doubtless many of these traits will be the outcome of relations of domination and exploitation within the family and the larger society, but

25 Habermas, *Legitimation Crisis*, p. 114.   26 Ibid. 117, 113.
27 T. McCarthy, *The Critical Theory of Jürgen Habermas* (London: Hutchinson, 1978), p. 306.

they will be sufficiently integral to, and internalized within, the personalities of actual people to make it implausible to suppose their eradication without those people becoming different people. Real participants in ideal communication would hold fast to their conceptions of their needs and interests and the norms they accept, to the extent that these conceptions are integral to their very identity.

The problem with the second answer is that the agents in question are constructs of a social theory which specifies relevant conflict groups or roles and their associated needs and interests, and the norms that regulate them. But, in that case, the counterfactual hypothesis will be a direct entailment of the theory and will do no more than spell out what, for the postulated conflict groups and role incumbents, are, according to the theory, negotiable or shared interests, true of self-interpreted needs and generally acceptable norms. In other words, this solution does no more than employ the counterfactual to spell out the implications of a theory accepted on other grounds, and its plausibility as a rational foundation for critical theorizing is no stronger than those grounds themselves.

The problem with the third answer is a familiar one to students of Kant and Rawls. If we are asked to imagine what ideally rational agents would do under the posited conditions, the whole argument turns on the nature of those agents and the constraints set by the conditions. If these together are such that the appropriate answers are necessarily reached, then the counterfactual hypothesis emerges as vindicated but only because it has been so formulated that it must do so. Ideally rational people in an ideal speech situation cannot but reach a rational consensus.

None of these answers suits Habermas's purposes, since his aim is to identify suppressed interests, 'contingent and forced consensus', and illegitimate domination, which requires him to specify and render plausible the counterfactual hypothesis that the individuals and groups in question would, under specified alternative conditions, acknowledge *their* true interests, *their* real and unforced consensus, and the claims to legitimacy *they* genuinely accept. John Stuart Mill reasoned thus when comparing the higher and lower pleasures, deferring to the judgement of the person who had experienced both: Habermas offers a version of this argument transposed to a collective and communicative level. And Habermas is quite right to observe that Marx reasons in this way, making the crucial, further assumption that 'the consciousness of justified and, at the same time,

suppressed interests' is a 'sufficient motive for conflict' to realize those interests.[28] The problem for Habermas is that none of his suggested answers to the question posed gives a plausible rendering of the counterfactual hypothesis required.

A further possible answer might be suggested: that under the imagined conditions of ideal communication, actual actors would be so transformed as to become capable of the requisite rational consensus. But this proposal encounters two severe objections. First, there is every reason to suppose that this would not be so, if socialization processes and relations of economic and political power remain unchanged: to suppose the contrary could be described as a rationalist illusion, typical of the eighteenth century, inconceivable to a Marxist. And second, even if it were so, the greater the change from the actual agents to the ideally rational agents, capable of reaching the requisite consensus, the less relevant would be the deliberations of the latter to the purpose at hand—which is to establish how the actual participants would think and feel, were alleged structures of domination to be overthrown.

This first difficulty I have been considering can be summed up in the question: *Who* are the participants in the unconstrained discourse that is held to offer the possibility of rational consensus? The second difficulty, to which I now turn, can be summed up in the question: *What* are they supposed to agree about?

In the first place, they are to interpret their needs 'collectively and bindingly' and in such a way that 'each individual must be able to recognize what he wants' and this becomes 'the object of discursive will-formation'. This seems to amount to a strong insistence on the self-ascription of desires and needs by autonomous agents and their acceptance of responsibility for the consequences of such self-ascriptions. Moreover, Habermas stresses the public and discursive nature of such self-ascriptions.[29] Second, in the light of such public, discursively reached self-ascriptions of desires and needs, they count certain norms as 'justified' and 'legitimate' if they fulfil 'commonly accepted needs'.[30] Another way that Habermas expresses this is to

28 Habermas, *Legitimation Crisis*, p. 114.
29 This is one of many interesting points of convergence with Rawls, who also stresses the importance of publicity: see Habermas, *A Theory of Justice*, pp. 177 ff.
30 Id., 'Wahrheitstheorien', p. 245.

say that the norms must 'express' and 'regulate' 'generalizable interests'. By 'generalizable interests' he means '*needs that can be communicatively shared*'. Thus rationally justified norms are those which 'express' and 'regulate' 'generalizable interests', that is needs that are publicly self-ascribed by autonomous and responsible participants in unconstrained discourse.

As yet this is a cloudy formula, and it is worth pressing on to see if clarification issues in a determinate solution. How, then, are we to recognize norms which express and regulate 'generalizable interests'? Habermas addresses this question. How, he asks, is it possible to separate by argument generalizable interests from 'those that are and remain particular'? His answer is that 'the only principle in which practical reason expresses itself' is 'the principle of universalization'.[31] He speaks of this principle as a 'bridge principle' comparable with the principle of induction:

Induction serves as a bridge principle for justifying the logically discontinuous passage from a finite number of singular statements (data) to a universal statement (hypothesis). Universalization serves as a bridge principle for justifying the passage to a norm from descriptive indications (concerning the consequences and side effects of the application of norms for the fulfilment of commonly accepted needs).[32]

The rationally justified norms to be agreed are, then, universalized. But what does this mean? There is a good case for distinguishing three forms or stages of universalization.[33] At the first stage, the principle of universalization simply dictates that the rules or norms guiding and controlling human conduct make no essential reference to proper names or indexical terms: that is to say, purely numerical differences are treated as irrelevant, so that what is right (or wrong) for you or your group is right (or wrong) for me and my group, unless there are morally relevant differences between us or our situation (with no restriction as to what can count as a morally relevant difference). At the second stage, the principle dictates that the rules or norms must be subject to a further test, namely putting oneself in the other person's or group's place. This test will allow through only those norms or rules which the actors are prepared to go on applying, no matter how they might change in respect of their mental and

31 Id., *Legitimation Crisis*, p. 108.
32 Id., 'Wahrheitstheorien', p. 245.
33 In what follows, I am indebted to the discussion of this question in J. L. Mackie, *Ethics: Inventing Right and Wrong* (Harmondsworth: Penguin, 1977), chap. 4.

physical qualities, resources and social status; on this test, 'differences can be fairly regarded as relevant if they look relevant from whichever side you consider them'.[34] At the third stage, a further, more stringent test is applied, namely that account is taken of rival and alternative desires, tastes, preferences, ideals, and values. Now one must simultaneously take account of conflicting points of view, and seek maxims that will be acceptable from all viewpoints. This test will allow through only those norms and rules which give equal weight, in some sense, to all interests, those interests being determined by rival preferences, values, and ideals.

The test at the first stage allows in all kinds of rules that are clearly partisan and unfair (e.g. all blacks should have an inferior education). Similarly the test at the second stage allows in universal rules that favour some preferences, values, and ideals as against others (e.g. a puritan code of conduct). We might perhaps assume, therefore, that Habermas is thinking of the third stage of universalization.

But the problem with the third stage of universalization is that the test is so severe that it is not clear that any rules or norms will pass, and it is certainly very far from guaranteed that unconstrained discourse between the parties will yield action-guiding principles of this sort. There would appear to be two possible approaches to this problem. One is the highly ambitious strategy of seeking to derive, through reflection on features inherent in the human condition and general facts about society, a determinate set of principles that will pass the test. This is John Rawls's approach. The 'circumstances of justice', which are 'the normal conditions under which human cooperation is both possible and necessary', obtain wherever 'mutually disinterested persons put forward conflicting claims to the division of social advantages under conditions of moderate scarcity'. Knowing this, and accepting that 'individuals not only have different plans of life but there exists a diversity of religious and philosophical belief, and of political and social doctrines', the parties in Rawls's 'original position', operating behind a veil of ignorance (about their own natural assets, future positions, and plans of life), have the task of unanimously settling on 'principles of justice that are genuinely neutral as between alternative plans of life'.[35] These principles, for

34 Ibid. 91–2.
35 Rawls, *A Theory of Justice* pp. 226–30. Rawls even claims that 'to see our place in society from the perspective of this position is to see it *sub specie aeternitatis*: it is to regard the human situation not only from all social but also from all temporal points

Rawls, 'not only specify the terms of cooperation between persons but they define a pact of reconciliation between diverse religious and moral beliefs, and the forms of culture to which they belong';[36] they are 'principles of accommodation between different moralities',[37] intended to secure 'social cooperation among equals for mutual advantage'.[38] Two crucial and telling objections that have been made against this approach, in the Rawlsian form, are, first, that his principles are not uniquely derivable from his hypothesized choice situation, and, second, that the 'original position' and its abstract, theoretically defined inhabitants have been so artificially constructed as to yield a predetermined solution, which has therefore no independently compelling qualification for the title '*the* principles of justice'.

A less ambitious approach to applying the test at this third stage of universalization is that of Mackie, who argues that the test is too severe: given radically divergent preferences and values and the obstinate moral disagreements arising therefrom, 'we must lower our sights a little, and look not for principles which can be wholeheartedly endorsed from every point of view, but for ones which represent an acceptable compromise between the different actual points of view'.[39] On this view of morality as compromise, there will no longer be any reason to suppose that a definitive or uniquely determinable set of rules or maxims can be arrived at through general rational argument: rather, there will be contingently different principles and rules, depending on the actual circumstances, divergences and possibilities of agreement. These must be 'invented' and 'reinvented' anew in the recurrent quest for mutual accommodation.

Neither of these approaches to universalization would seem to suit Habermas's purposes. Rawls's quest for an Archimedean point

of view' (ibid. 587). Ronald Dworkin, similarly, has argued that the constitutive political morality of liberalism rests on the idea that 'government must be neutral on what might be called the question of the good life', that 'political decisions must be, as far as is possible, independent of any particular conception of the good life, or of what gives value to life' ('Liberalism', in *Public and Private Morality*, ed. S. Hampshire (Cambridge: Cambridge Univ. Press, 1978), p. 127).

36 Rawls, *A Theory of Justice*, p. 221.
37 Id., 'Fairness to Goodness', *Philosophical Review* 84 (1975), p. 539.
38 Id., *A Theory of Justice*, p. 14.
39 Mackie, *Ethics*, p. 93.

eludes him, for the reasons suggested and others,[40] while Mackie's position does not yield a uniquely rationally justified outcome, only contingent compromises. But in any case, Habermas appears to reject a basic assumption shared by both these approaches, namely that morality is a means of solving the problem posed by the conflicts generated by limited resources and limited sympathies, with a view to securing mutually beneficial co-operation. To the contrary, he explicitly distinguishes 'norms that may claim generality' from 'compromise',[41] which he describes as a case of 'normative power' involving a 'normed adjustment between particular interests' that 'takes place under conditions of a balance of power between the parties involved'. Although such compromise may be 'indirectly justifiable' through discourse, it is plainly seen as distinct from and contrasting with the case of 'interests that permit of a rational will': a compromise can be justified only where 'a balance of power among the parties involved and the non-generalizability of the negotiated interests exist'.[42]

Nor indeed is it surprising that Habermas should take this line, since the entire Marxist tradition is committed to denying that morality is a response to an inherent limitation of resources and sympathies in the human condition—a view classically expressed by Hume when he wrote that 'It is only from the selfishness and confined generosity of man, along with the scanty provision nature has made for his wants, that justice derives its origin'.[43] On the contrary, Marx and subsequent Marxists see both limitations as historically contingent and socially generated and look forward to overcoming both in a unified society of abundance.

What, then, if all this is so, does Habermas understand by the principle of universalization as expressed in the formula 'general-izability of interests'? In truth, I find it difficult to say. A clue may lie

[40] For an argument in support of this claim, see the present author's *Essays in Social Theory* (London: Macmillan, and New York: Columbia Univ. Press, 1977), chap. 10.

[41] Habermas, *Legitimation Crisis*, p. 89.

[42] Ibid. 112.

[43] D. Hume, *A Treatise of Human Nature* (1739), ed. L. A. Selby-Bigge (Oxford: Clarendon Press, 1951), p. 495, bk. III, pt. II, sect. II. For an interesting discussion of 'morality as compromise', see Arthur Kuflic, 'Morality and Compromise', in J. R. Pennock and J. W. Chapman (eds.), *Compromise in Ethics, Law and Politics, Nomos XXI*. This being the *Yearbook of the American Society for Political and Legal Philosophy* (New York: New York Univ. Press, 1979).

in his attack on the Rawlsian notion of 'primary goods' as supposedly 'neutral means for attaining an indefinite multiplicity of concrete ends selected according to values'. He argues that this picture is misleading, that the Rawlsian primary goods are not compatible with all forms of life that could be chosen, but involve 'clearly circumscribed "opportunity structures"' and imply a particular underlying form of life, of private commodity production and exchange relations, and familial, occupational, and civil privatism. In the light of this critique, he argues for the possibility that the 'pursuit of happiness' might one day mean something different—for example, not accumulating material objects of which one disposes privately, but bringing about social relations in which mutually predominates and satisfaction does not mean the triumph of one over the repressed needs of the other.[44]

The idea seems to be that there will be an endogenous change of preferences on the part of social actors (induced by ideal discourse?) such that preferences, tastes, values, ideals, plans of life, etc. will to some large degree (to what degree?) be unified and no longer conflict. On this interpretation, the principle of universalization would require that norms and rules pass the test that they should embody common aims, regulate shared activities and lead to common and shared satisfactions ('needs that can be communicatively shared'). But why should one suppose such a moral change to be either possible, necessary, or desirable? It is true that *Legitimation Crisis* is, in part, addressed to showing such a change to be on the historical agenda in late capitalist societies and there are, of course, good arguments in favour of a more unified, less conflictual, privatized and consumerist form of life than is predominant in contemporary capitalist societies; but I cannot find in Habermas's writings any argument for the thesis that such a form of life (which is, in any case, barely even sketched, except in the most abstract possible manner) is either an appropriate interpretation of the principles of universalization or uniquely capable of rational justification.

I turn now to the question of whether Habermas has justified his claim that his approach overcomes 'the limits of a decisionistic treatment of practical questions as soon as argumentation is expected to test the generalizability of interests'.[45] I cannot see that it

[44] Habermas, *Communication*, pp. 188–9.

[45] Id., *Legitimation Crisis*, p. 108.

does so. For the principle of universalization, at all the three stages I have discussed, requires, at each stage, a *decision* whether or not to let one's actions and choices be guided only by maxims and norms which pass the test in question. As Mackie has remarked, 'the universalisability of moral judgements . . . does not impose any rational constraint on choices of action or defensible patterns of behaviour'.[46] At every stage it is not merely logically possible but in fact quite common for people to opt out of, or not opt into, these ways of reasoning. And this argument applies, with equal force, to Habermas's own (putative) version of the universalization principle.

I conclude from all this that (1) these arguments do not succeed in establishing a determinate, uniquely rational basis for critical theory; (2) they do not dispense with the role of decision in moral and political thinking; and (3) they do not therefore disprove the thesis of value pluralism.

Habermas has, however, advanced a further set of considerations which are intended to provide an 'empirical' and a 'systematic' basis for the objectivity and universality of standards of rationality as he conceives them. These considerations are intended to provide further grounding for his positive claim to establish a determinate rational basis for critical theory and thereby his rejection of 'decisionism' and Weber's 'rationally irresoluble pluralism of competing value systems and beliefs'.[47] I refer to his application of Piaget's and Kohlberg's cognitive developmental psychology.

I cannot here discuss in detail Habermas's intriguing attempt to use these ideas to try to reconstruct historical materialism by incorporating within it a theory of the development of normative structures. I shall refer only to the gist of his argument, in so far as it concerns the question at issue here. His basic thought is that, since cognitive developmental psychology is 'well corroborated and . . . has reconstructed ontogenetic stages of moral consciousness', it should be possible to reconstruct these stages logically, that is, 'by concepts of a systematically ordered sequence of norm systems and behavioral controls', and at the highest stage to identify a corresponding 'universal morality', which can be traced back to fun-

46 Mackie, *Ethics*, p. 99.
47 Habermas, *Legitimation Crisis*, p. 100.

damental norms of rational speech'.[48] With extraordinary boldness, he traces homologies between ego development on the one hand, and 'the social evolution of moral and legal representations', of 'world views', and 'the historical constitution of collective identities' on the other.[49] Just as the ego develops from the symbiotic, through the egocentric and the sociocentric-objectivistic to the universalistic stages of development, and from pre-conventional through conventional to post-conventional patterns of problem-solving,[50] so, Habermas argues, it should be possible to develop a 'communication theory' that would analyse 'the symbolic structures that underlie law and morality, an intersubjectively constituted world, and the identities of persons and collectivities', that would show normative structures to have their own 'internal history' which has a 'direction of development' that can be characterized by such concepts as 'universalization and individualization, decentration, autonomization, and becoming reflective'.[51] In short, the 'stages of law and morality, of ego demarcations and world views, of individual and collective identity formations' can be seen as stages in progress towards increasing rationality, measured by 'the expansion of the domain of consensual action together with the re-establishment of undistorted communication'.[52] Such progress can be seen as a sequence of developmental stages of moral consciousness, corresponding to stages of development in interaction competence. At the final, post-conventional stage, 'systems of norms lose their quasi-natural validity; they require justification from universalistic points of view'.[53]

Do these considerations help to render the 'universalistic' perspective of the highest stage, and the principles and judgements it delivers, determinate, uniquely rational and objective, and do they

48 Id., *Zur Rekonstruktion des Historischen Materialismus* (Frankfurt: Suhrkamp, 1976), p. 205 ('History and Evolution', trans. D. J. Parent, *Telos* 39 (Spring 1979), 8); id., *Legitimation Crisis*, p. 95.
49 Id., *Communication*, pp. 99, 102, 110.
50 Kohlberg's six stages of moral consciousness are: (1) punishment-obedience orientation, involving maximization of pleasure through obedience; (2) instrumental hedonism, involving maximization of pleasure through exchange of equivalents; (3) 'good-boy/nice-girl' orientation, involving concrete morality of gratifying interactions; (4) law-and-order orientation, involving concrete morality of a customary system of norms; (5) social-contractual legalism, involving civil liberties and public welfare; and (6) ethical-principles orientation, involving moral freedom.
51 Habermas, *Communication*, pp. 116–17.
52 Ibid. 120.
53 Ibid. 156.

bypass the need for a decision to adopt it? It is notable that Kohlberg has himself argued a similar case to that of Habermas. He has claimed to 'have successfully defined the ethically optimal end-point of moral development', that 'there is a sense in which we can characterize moral differences between groups and individuals [which are themselves to be understood as differences in stage or developmental status] as being more or less adequate morally'.[54] For Kohlberg, the later stages of development are 'cognitively and ethically higher or more adequate', the final sixth stage being 'a more universalistic, moral orientation, which defines moral obligations in terms of what alternatively may be conceived as (a) the principle of justice, (b) the principle of role-taking, or (c) the principle of respect for personality', this stage being 'the most adequate exemplification of the moral'.[55] In this sense, Kohlberg argues for the 'superiority of Stage 6 judgements of duties and rights (or of justice) over other systems of judgements of duties and rights'.[56] Habermas, similarly, argues for both an empirical and a systematic 'superiority' of universal morality.[57]

But Kohlberg in fact established only that there are recurrent sequences of stages of preferred modes of moral reasoning employed by growing children in different contexts and cultures.[58] The claim that this repeated sequence is a development towards greater adequacy only makes sense if one applies specific criteria of adequacy, criteria that are themselves disputable and indeed disputed among moral thinkers, philosophical and otherwise. In Kohlberg's theory such criteria are already built into his 'scientific' theory. The claimed superiority of the later stages over the earlier and the last stage over the rest is necessitated neither by the observed data, nor by the fact that the later modes of reasoning may logically presuppose earlier modes, nor by the fact that they may imply cognitive superiority or require more elaborate conceptual thinking. There is no inherent compulsion in Kohlberg's claim that Stage 6 reasoning is ethically

54 L. Kohlberg, 'From Is to Ought: How to Commit the Naturalistic Fallacy and Get Away with it in the Study of Moral Development', in *Cognitive Development and Epistemology*, ed. T. Mischel (New York: Academic Press, 1971), pp. 153, 176.

55 Ibid. 208, 218.

56 Ibid. 214–15.

57 Habermas, *Legitimation Crisis*, p. 95.

58 See the critique of Kohlberg by William Alston in *Cognitive Development and Epistemology*. Also see P. R. Dasen, 'Cross-cultural Piagetian Research: A Summary', *Cross Cultural Psychology* 3 (1972), pp. 23–40.

optimal: as Alston has observed, 'many moral philosophers who are surely at least as conceptually sophisticated as Kohlberg's Stage 6 subjects take positions in moral philosophy that reflect Stages 4 or 5'.[59] It is a claim that many can and will dispute, and it cannot be established by stipulatively defining 'moral' so that Stage 6 becomes 'the most adequate exemplification of the moral'. And the same arguments apply, *pari passu*, to Habermas's claim, *contra* Kohlberg, that there is an ultimate Stage 7 at which 'the principle of justification of norms is no longer the monologically applicable principle of generalizability but the communally followed *procedure* of redeeming normative validity claims discursively'.[60]

Indeed, these arguments can only be strengthened by considering the very disagreement between Kohlberg (and by implication Rawls[61]) on the one hand, and Habermas on the other, as to the nature of the highest stage. Kohlberg has argued that the highest stage of moral development implies the notion of justice as 'reversibility', a kind of 'ideal role-taking' which involves 'differentiating the self's perspective from the other's and co-ordinating the two so that the perspective from the other's view influences one's own perspective in a reciprocal fashion'.[62] This amounts to a moral decision procedure, which Kohlberg calls 'moral musical chairs', that is, 'going round the circle of perspectives involved in a moral dilemma to test one's claims of right or duty until only the equilibrated or reversible claims survive'.[63] (Kohlberg further claims that his decision procedure yields the same solutions as Rawls's idea of decision in the 'original position'.) By contrast, as we have seen, Habermas maintains that Kohlberg's approach is 'monological' and fails to attain 'the level of a universal ethic of speech' where self-

59 Alston, in *Cognitive Development and Epistemology*, p. 275.

60 Habermas, *Communication*, p. 90.

61 Rawls claims support from the ideas of Piaget and Kohlberg, in seeking to indicate 'the major steps whereby a person would acquire an understanding of an attachment to the principles of justice as he grows up in this particular form of a well-ordered society' (*A Theory of Justice*, p. 461; see sects. 69–72). On the other hand, *contra* Kohlberg, Rawls is clear that the claimed superiority of his theory of justice 'is a philosophical question and cannot, I believe, be established by the psychological theory of development alone' (ibid. 462). Kohlberg, reciprocally, sees Rawls's theory as consonant with his own.

62 Kohlberg, 'Justice as Reversibility', in *Philosophy, Politics and Society*, 5th ser., ed. P. Laslett and J. Fishkin (Oxford: Blackwell, 1979), p. 266.

63 Ibid. 262.

ascriptions of interests become 'the object of practical discourse'.[64] If my interpretation of Habermas's understanding of universalization, set out above, is correct, he assumes that such practical discourse will lead to an endogenous change of preferences and perspectives on the part of the communicators such that shared needs and 'consensual action'[65] will predominate. Yet how can one rationally resolve this difference between two such rational men? How can one conclude which is rationally and morally superior—Kohlberg's Stage 6 or Habermas's Stage 7, and the norms and judgements they respectively generate—other than by *deciding* between them? Of course, one might at this point maintain that the very distinction between 'decision' and 'rational argument' is misleading here. After all, would not any such decision be based on *reasons*: do we not decide on the basis of *rational grounds*? But neither Habermas nor Kohlberg (nor Rawls) has shown that there is a neutral or objective standpoint from which the reasons grounding alternative decisions can themselves be assessed as more or less rational.

So I conclude that this line of argument also fails to establish the critical vantage-point that Habermas seeks and that we have not yet escaped the Weberian gods and demons.

64 Habermas, *Communication*, p. 90.
65 Habermas, *Communication*, p. 110.

12

C. B. Macpherson and the Real and Ideal Worlds of Democracy

Brough Macpherson's democratic theory strikes a distinctive note. Resolutely Anglo-Saxon in its range of reference and its crisp, clear, analytic style, it unites a Marxist-inspired critique of 'capitalist market society with its class-division'[1] and of the underlying market assumptions of the justifying theory of liberal democracy with the constructive 'liberal' aim of 'retrieving' from that theory the 'notion of a democratic society as one that provides equally for the self-development of all the members of a political community'.[2] His motivating animus is against possessive individualism—'this perverse, artificial, and temporary concept of man', inherited from 'classical liberal individualism', as 'essentially a consumer of utilities, an infinite desirer and infinite appropriator' whose over-riding motivation is 'to maximise the flow of satisfactions, or utilities, to himself from society'.[3] His positive commitment, by contrast, is to a 'co-operative and creative individualism' which rescues 'the human-ist side of Mill's liberalism (the side based on his idea of man as essentially an exerter and developer of his human capacities) from the possessive individualist side (based on the Benthamite concept of man as essentially consumer and appropriator)'.[4] Thus he places himself among 'those who accept and would promote the normative values that were read into the liberal-democratic society and state by

This chapter was first published in 1979.

1 Macpherson, *The Life and Times of Liberal Democracy* (Oxford: Clarendon Press, 1977), p. 21.
2 Id., 'The False Roots of Western Democracy', in Fred R. Dallmayr (ed.), *From Contract to Community: Political Theory at the Crossroads* (New York and Basel: Marcel Dekker, 1978), p. 26.
3 Id., *Democratic Theory: Essays in Retrieval* (Oxford: Clarendon Press, 1973), 20, 63, 24; id., *Life and Times*, p. 43.
4 Id., 'Individualist Socialism? A Reply to Levine and MacIntyre', *Canadian Journal of Philosophy*, 6/2 (June 1976), p. 198.