

neue hefte für philosophie

Band 11

Herausgegeben von:
Karl Popper und
Conrad Aiken
Herausgegeben von:
Karl Popper und
Conrad Aiken

1

NEUE HEFTE FÜR PHILOSOPHIE

Heft 21

Politikbegriffe

Julien Freund Die doppelte Ethik	1—21
Klaus Hartmann Kategoriale Topik im Politischen	22—34
Ernst Vollrath Ein philosophischer Begriff des Politischen?	35—46
Seyla Benhabib The methodological illusions of modern political theory: The case of Rawls and Habermas	47—74
Henning Ottmann Politische Philosophie als Gespräch	75—93

Die nächsten Hefte behandeln folgende Themen:

Kants Ethik heute. Heidegger. Argumentation in der Philosophie.

Es liegen vor: Heft 1: Phänomenologie und Sprachanalyse. 107 Seiten — Heft 2/3: Dialog als Methode. 171 Seiten — Heft 4: Theorie literarischer Texte. 139 Seiten — Heft 5: Ist eine philosophische Ästhetik möglich? 138 Seiten — Heft 6/7: Tendenzen der Wissenschaftstheorie. 168 Seiten — Heft 8: Semantik und Ontologie. 95 Seiten — Heft 9: Handlungstheorie. 160 Seiten — Heft 10: Moderne Sophistik. 111 Seiten — Heft 11: Philosophische Psychologie? 114 Seiten — Heft 12: Spinoza 1677—1977. 91 Seiten — Heft 13: Marx' Methodologie. 91 Seiten — Heft 14: Zur Zukunft der Transzendentalphilosophie. 142 Seiten — Heft 15/16: Aktualität der Antike. 182 Seiten — Heft 17: Recht und Moral. 125 Seiten — Heft 18/19: Anschauung als ästhetische Kategorie. 177 Seiten — Heft 20: Teleologie. 152 Seiten.

Die „Neuen Hefte für Philosophie“ erscheinen in unregelmäßigem Abstand, mindestens aber zweimal im Jahr. Anschrift der Redaktion: 6900 Heidelberg, Marsiliusplatz 1. In den Heften sind keine Buchrezensionen vorgesehen.

Preis dieses Heftes: 24,— DM; Subskriptionspreis (bei Abnahmeverpflichtung für mindestens 3 aufeinanderfolgende Hefte): 19,80 DM zuzüglich Porto. Kündigung der Subskription nach Ablauf der Abnahmeverpflichtung ist innerhalb von sechs Wochen nach Auslieferung eines Heftes möglich. Für unsere Preise ist ausschließlich unsere jeweils gültige Preisliste verbindlich.

© Vandenhoeck & Ruprecht, Göttingen 1982. Alle Rechte, auch das der Übersetzung, für sämtliche Beiträge vorbehalten. Satz: Fotosatz Tutte, Salzweg-Passau, Druck- und Buchbinderarbeit: Verlagsdruckerei E. Rieder, Schrobenehausen.

ISSN 0085-3917

Die doppelte Ethik

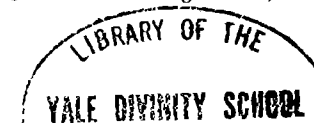
JULIEN FREUND/Straßburg

Die Beziehung zwischen Politik und Ethik bildet einen klassischen Streit der politischen Philosophie, in welchem Kulturraum es auch immer sein mag. Auf eine Weise, die uns heutzutage noch zu berühren vermag, ist das Problem in der alten indischen Literatur, z. B. im „Arthashastra“ von Kautilya, dargestellt worden. Diesem Autor zufolge ist jede Politik gut, die mit Erfolg ihre Macht behauptet, gleich welche Mittel dabei verwendet werden. Andere Autoren hingegen meinen, daß der Idealstaat derjenige sei, in dem das politische Handeln mit den Normen und den Vorschriften der Ethik übereinstimmen. Es genügt, auf die zahlreichen Verfasser von Utopien in den letzten Jahrhunderten hinzuweisen. Zwischen diesen extremen Gesichtspunkten gibt es solche, die nuancierter sind, in mehr oder weniger kritischer Weise jedenfalls darauf bedacht, die Belehrungen der Geschichte sowie die Beobachtungen des konkreten Verhaltens der Menschen zu berücksichtigen.

Die einen drücken ihre Meinung in der Form des Wunsches aus, wie A. Camus, der „die Sprache der Ethik in die Ausübung der Politik einführen“ möchte¹, oder auch in Gestalt der Forderung, daß die Politik der Ethik unterstellt sei. Andere sind mehr dazu geneigt, die Moral und das Recht dem Staat zu subordinieren, im Sinne Spinozas: „Wir sehen demnach, daß jeder Bürger nicht unter seinem, sondern dem Recht des Staates steht, dessen sämtliche Befehle er zu befolgen verbunden ist, und daß er kein Recht habe, zu entscheiden, was billig und unbillig, fromm und gottlos sei, daß er vielmehr, weil der Staatskörper wie von einem Geist geleitet wird, und folglich der Wille des Staates als der Wille aller gelten muß, das, was der Staat als gerecht und gut beschließt, so ansehen muß, als ob es von jedem Einzelnen beschlossen wäre, und daß also ein Untertan, wenn er auch die Beschlüsse des Staates für unrecht hält, sie doch zu befolgen gehalten ist“². Im allgemeinen jedoch wird angenommen, daß die Politik das Recht habe, die Ethik zu pervertieren, insbesondere in Notlagen oder in Ausnahmeverhältnissen. Tocqueville ist da anderer Ansicht, er bemerkt in seinen *Souvenirs*, daß, selbst wenn die Sitten einer Gesellschaft von denjenigen einer anderen Gesellschaft abweichen, „die Sitten der Politiker, die

1 A. Camus, *Actuelles I*, Paris 1950, S. 51.

2 Spinoza, *Sämtliche Werke*, Hrsg.: Auerbach, Stuttgart 1871, Bd. I, S. 436.



hältnis gesetzt werden. Die Maxime der Urteilskraft spricht diese bindende Bedingung aus. Auf sie ihr politisches Verhältnis gründend würden Menschen auf dem Boden kluger Vernünftigkeit in der Welt gemeinsam wohnen.

The methodological illusions of modern political theory: The case of Rawls and Habermas*

SEYLA BENHABIB/Boston

I. Introduction

Modern political philosophy began with an abstraction. "Let us . . . consider men", wrote Hobbes in *De Cive*, "as if but even now sprung out of the earth, and suddenly, like mushrooms, come to full maturity, without all kind of engagement to each other"¹. This invocation to view man in all his nakedness outside the bonds of society, history, and tradition was enough to earn Hobbes the title of being the father of modern political philosophy. Descartes had initiated modern epistemology with his call to leave behind the world of human affairs, conventions, and the senses. Removed from the fetters of prejudice, the noise of human affairs, the dogma of the Church, and the deception of the senses, for the Cartesian ego this withdrawal from the world was the beginning of a process of emancipation through doubt². By negating the appearances, the autonomous Cartesian ego gained an ontological permission to reconstruct them according to a masterplan like an architect operating with a single vision. The emancipatory thrust of Hobbes' methodological abstraction paralleled that of Descartes'. By proposing to consider "men as mushrooms", Hobbes was arguing that the current state of affairs in which they lived could no longer be taken as a beginning point for political theory. The given had lost its claim to be. A truthful reflection on politics could only begin by negating the given. With the refusal to take existing conditions as they were, Hobbes also rejected the legitimacy of time, tradition, and his-

* This essay was written under the auspices of an Alexander von Humboldt Fellowship during my stay at the Max-Planck-Institut in Starnberg 1979-1981. I would like to thank Professor Jürgen Habermas for his hospitality, the Alexander von Humboldt Stiftung for their generosity, my colleagues Wolf Schäfer, Klaus Eder, and Adina Schwartz for critical discussions of this text. Needless to say, I have been able to accommodate neither all their insights nor all their criticisms (particularly those voiced by Eder).

¹ Thomas Hobbes, "Philosophical Rudiments Concerning Government and Society", in: Sir William Molesworth (ed.), *The English Works of Thomas Hobbes*, v. II, Darmstadt 1966, 109.

² Renée Descartes, *Discours de la Methode*, Introd. by Karl Jaspers, Mainz 1945.

tory to bind the individual. The political world would be well-recovered only after the individual had journeyed through the thought experiment of the state of nature and of the social contract. Fictional though it was, like Descartes' methodical doubt, this thought experiment was devised to teach men to become the makers and architects of a new political order, and to provide them with a master plan for founding the new commonwealth. Modern political philosophy posited the methodological withdrawal from the actual world of politics to be a precondition for future political emancipation.

It will be the purpose of this essay to discuss the Hobbesian methodological legacy in the work of two contemporary political thinkers: John Rawls and Jürgen Habermas*. The similarities between the Rawlsian theory of justice and Habermas' theory of Diskurs have been the object of increasing attention in recent years³. There are two premises shared by Rawls and Habermas. I will call the first the "consensus principle of legitimacy" and define it as follows: the principle of rational consensus provides the only

* The following texts will be cited with abbreviations:

- TJ* = John Rawls, *A Theory of Justice*, Cambridge, Mass. 1972². The first number in parentheses gives the page references to this edition; the second, refers to the German translation: J. Rawls, *Eine Theorie der Gerechtigkeit*, Hermann Vetter, Übersetzer, Frankfurt 1975.
- KC* = John Rawls, „Kantian Constructivism in Moral Theory“, *The Journal of Philosophy*, Bd. LXXVII, 9 (September 1980), 515–572.
- BWU* = Jürgen Habermas, „Einige Bemerkungen zum Problem der Begründung von Werturteilen“, in: L. Landgrebe (Hrsg.), *Philosophie und Wissenschaft*, Meisenheim am Glan 1972, 89–99.
- HL* = J. Habermas/N. Luhmann, *Theorie der Gesellschaft oder Sozialtechnologie*, Frankfurt 1971.
- Wth* = J. Habermas, „Wahrheitstheorien“, in: H. Fahrenbach (Hrsg.), *Wirklichkeit und Reflexion* (Festschrift W. Schulz), Pfullingen 1973, 211–265.
- Lpr* = Ders., *Legitimationsprobleme im Spätkapitalismus*, Frankfurt, 1973².
- R HM* = Ders., *Zur Rekonstruktion des Historischen Materialismus*, Frankfurt 1976².

- 3 T. A. McCarthy, *The Critical Theory of Jürgen Habermas*, Boston 1979, 310ff.; Steven Lukes, „Of Gods and Demons: Habermas and Practical Reason“, in: D. Held and Thompson (eds.), *Essays on Habermas*, London (Forthcoming); Albrecht Wellmer, „Equality, Freedom and Rationality“, Ms. 1979; Herbert Kitschelt, „Moralisches Argumentieren und Sozialtheorie“, *ARSP*, Bd. LXVI/3 (1980), 391–426. The relationship between Rawls' *A Theory of Justice* and Habermas' model of discourse has been treated in the context of the evolution of normative justification procedures by a research group in Starnberg, cf. Eder, Frankenberg, Rödel and Tugendhat, „Projektvorschlag: Die Entwicklung von Gerechtigkeitsvorstellungen und Begründungsverfahren im modernen Recht als soziologisches Problem“, Ms. Starnberg 1978 and E. Tugendhat, „Zur Entwicklung von moralischen Begründungsstrukturen im modernen Recht“, *ARSP*, Bd. LXV/1 (1980), 1–20.

criterion in light of which the legitimacy of norms and of normative institutional arrangements can be justified. More significantly, Rawls and Habermas share a meta-theoretical premise: the idea of such a rational consensus is to be defined procedurally by specifying strategies and modes of argumentation through which alone such a consensus of norms and on normative institutional arrangements can be attained. Let me call this "methodological proceduralism". Rawls maintains that his theorem of justice provides us with the only procedure of justification through which valid and binding norms of collective coexistence can be established. Habermas argues that the counterfactual of the "ideal speech situation" defines the formal properties of that mode of reflexive argumentation called "diskurs", by engaging in which alone we can attain a rational consensus on norms. The fictive collective choice situation devised by Rawls and the equally fictive "ideal speech situation" devised by Habermas are *normative justification procedures* serving to illustrate the consensus principle of legitimacy.

In this essay I will argue that both counterfactual justification procedures are flawed by a dialectic of form and content, or of formal-procedural constraints and material-substantive assumptions. The first aspect of this dialectic concerns the criteria of selection presupposed by counterfactual abstractions; the second, the relationship of sociological presuppositions to normative principles in such counterfactual procedures. When examined under the first aspect, the procedural theories of Rawls and Habermas display a circularity: the initial counterfactual abstraction does not serve as justification, but rather, as an illustration, of normative principles presupposed by the theory (section 2 below). With regard to the second aspect of the form-content dialectic: while Rawls' theorem of rational choice is based upon a very questionable view of the social sciences and is, therefore, epistemologically inconsistent (section 3.1 below), in Habermas case, the problem is that the counterfactual theory of discourse can hardly fulfill the critical intentions of his sociological theory (section 3.2 below).

Despite the flaw in their internal logic, procedural argumentations continue to be very influential in normative social and political theory. In recent years, this has been undoubtedly due to the justified influence of Rawls' work. I should therefore say at the outset that since Rawls' work has been criticized on so many levels and in so many respects, to assume that at this point yet another Rawls critique would contribute much to our thinking about these issues may be almost foolhardy. This is not my purpose. My goal in this essay is to criticize a tradition of methodological proceduralism in political philosophy that runs from Hobbes to Habermas in the context of a more general question concerning the relationship of normative philosophy to critical social theory. I want to question whether

proceduralism is the correct strategy by which to integrate the achievements of critical social theory with the insights of normative philosophy⁴.

2. The Dialectic of Form and Content I: Criteria of Selection and Counterfactual Abstractions

Procedures of normative justification, which begin by counterfactually abstracting from known conditions of social and historical existence, must justify this abstraction in light of some principles of inclusion and exclusion. There must be criteria of selection in light of which certain aspects of human identity, capacities and qualities are included in the description of the counterfactual, while others are excluded from it. Since Hobbes the tendency of modern political theory has been to eliminate empirically and normatively restrictive criteria of identity like class, sex and status distinctions, and to minimize psychological and anthropological assumptions concerning what human nature is or is not⁵. As theories of rational con-

⁴ There has always reigned unclarity in the tradition of Marxism about the status of normative ideals like justice, freedom and autonomy in the name of which thinkers from Marx to Marcuse have criticized capitalist society. On the one hand, Marxism has replaced the tradition of normative philosophy by a "science" of society. On the other hand, the method of "immanent critique" juxtaposed the universalistic ideals of bourgeois society to the sociological reality of the latter, seemingly committing itself to such ideals but criticizing their particularistic realization. Cf. Habermas, "Historischer Materialismus und die Entwicklung normativer Strukturen", RHM, 9-49. Within the scope of this essay, I can do no more than mention this problem, although the question of methodological proceduralism, which concerns me here, is part of a research project on the normative grounding of a critical theory of society; cf. my "Procedural vs. Discursive Norms of Rationality: The Search for the Normative Grounding of Critique", Ms. Starnberg January 1980.

⁵ The metaphor of the "state of nature" with which contractarian theories in the modern period began, implied both an anthropology and a psychology of human nature, and a theory of pre-political institutions and relations, like domestic servitude, marriage and the right of the father over his children. By claiming that relations between husband and wife, master and servant, child and father were part of the state of nature, contract theories removed them from the jurisdiction of civil government and reduced them to relations of private personal domination. They thereby restricted the right of rational consent to the economically independent male head of household. Cf. Thomas Hobbes, *Leviathan*, C. B. McPherson (ed.), Middlesex 1968, ch. XX; John Locke, "Second Treatise of Civil Government" in *Two Treatises of Civil Government*, Peter Laslett (ed.), Cambridge 1970, part 1. Even Kant, who defined the state of nature as the sphere of *Privatrecht*, defended the privilege of the male head of household over his wife, children and servants. I. Kant, *Die Metaphysik der Sitten*, *Werke in Zehn Bänden*, Bd. VII, Darmstadt 1971, 385-397.

The first systematic distinction between the concept of the state of nature and the sphere of private-personal rights, and between an empirical theory of human nature and the concept

sent become more *universalistic* in the substantive sense, as they consider all and not only some group of Humans to be entitled to consent, they must also become increasingly formal in their methodological assumptions. For, if it is assumed that all humans are entitled to consent, or have the right to be considered as beings capable of consent, then the theory must proceed from a sufficiently general and minimal conception of the empirical and normative identity to be ascribed to human beings. If this conception were not sufficiently minimal and general, then the theory would contradict the principle of rational consent of all, for it would have restricted this principle only to certain group of humans, who are such and such, and have the capacity to act in such and such ways. This would be carried out on the basis of strong normative assumptions the validity of which would have been presupposed, but not justified. But the more minimal and general the conception of empirical and normative identity of persons from which the theory proceeds, the more unlikely that this will prove sufficient to yield normative principles of social and institutional arrangements. In this section, I will examine the various arguments by which Rawls and Habermas justify the "original position" and "the ideal speech situation", respectively.

2.1 The Justification of the Original Position in the Rawlsian Theory of Justice: Reflective Equilibrium and the Thin Theory of the Good

Rawls' initial justification for the counterfactual called the "original position" is that it "serves to make vivid to ourselves the restrictions that it seems reasonable to impose on arguments for principles of justice, and therefore on these principles themselves" (TJ, 18; 36). Principles of justice must be those that "free and rational persons concerned to further their own interests would accept in an initial position of equality as defining the fundamental terms of their associations" (TJ, 11; 28). This conception of justice is called "justice as fairness". Individuals are equal in this initial situation in the sense that they are deprived of specific information concerning their place in society, of the distribution among them of natural assets and abilities, like intelligence and strength, and of their own psychological propensities. The parties to the original position are rational, moral persons. They have their own moral ends, which are not ranked in value; they are capable of pursuing them and they are capable of a sense

of the person, was made by Hegel in the *Rechtsphilosophie*, G. W. F. Hegel, *Vorlesungen über Rechtsphilosophie*, 1818-1831, K.-H. Ilting (Hrsg.), Bd. IV, Stuttgart 1974; "Das Abstrakte Recht", 170-182; cf. my "Obligation, Contract and Exchange: The Opening Arguments of Hegel's *Philosophy of Right*", *Hegel's Political Philosophy*, II, Z. A. Pelczynski (ed.), Cambridge University Press London (Forthcoming).

of justice. By the "freedom" of the parties in this situation, is to be understood that as equal, rational, moral beings, they can conceive of the principles of their cooperation to be the result of their own actions, and view them as self-imposed (TJ, 13; 30).

Throughout the initial sections of *A Theory of Justice* Rawls expresses a preference for "generally shared and preferably weak conditions" (TJ, 20; 37). The original position can also be regarded as "the result of a hypothetical course of reflection", called reflective equilibrium (TJ, 20ff.; 38ff.), a procedure of going back and forth between our existing judgments and philosophical-moral principles. Rawls takes great pains not to introduce as justificatory steps in his argument the strong conception of the good presupposed by the theory of justice as fairness – goodness as rationality (TJ, 407ff.; 445ff.), the Aristotelian moral psychology (TJ, 424ff.; 463ff.), the theories of moral development and of the sense of justice (TJ, 490–496, 504–505; and unanimously acceptable definition of the "original position" on the basis of weak and generally shared premises about moral and rational human agency. If this can be granted, Rawls assumes, then it can also be granted that principles of justice, reached through a correct deliberation procedure in this position, are also justified.

Although Rawls does not use this term, I will describe "justice as fairness" as the Rawlsian equivalent of the principle of rational consent. This principle is a second-order principle defining an acceptable level of moral justification, which first-order principles must be such as free and rational persons would accept in an initial position of equality as defining the fundamental terms of their associations. But it is no accident that Rawls does not distinguish between justice as fairness as a second and a first-order principle, or between the level of normative justification and a specific theory that would satisfy this level of justification. This conflation changes the meaning of the original situation: whereas the original position was supposed to make "vivid to ourselves" the restrictions that it seems *reasonable* to impose on arguments for principles of justice, it comes to define all those *rational* restrictions that a *specific conception* of justice presupposes⁶. Rawls writes: "we have to ascertain which principles it would be rational to adopt given the contractarian situation. This connects the theory of justice with the theory of rational choice . . . The concept of the original position, as I refer to it, is that of the *most philosophically favored interpretation of this initial choice situation* for the purposes of a theory of justice" (TJ, 17–18; 35. *Emphasis mine*). The most "philosophically favored" interpretation of the original position yields a strong justification strategy, called that of the "privileged description".

⁶ The distinction between the "rational" and the "reasonable", which is central to my argument, is not always made explicit in the German translation. I discuss this distinction below, cf. 2.1.2.

2.1.1 The Privileged Description of the Original Position

The philosophically favored interpretation of the original position as a rational choice situation, involves the introduction of the following presuppositions: "thick veil of ignorance" (TJ, § 24); the theory of primary goods (TJ, § 15); the definition of the circumstances of justice; the formal constraints of the concept of right (TJ, § 23); and the minimax rule as the principle of rational choice (TJ, 144ff.). The move from the weak to the "philosophically preferred description" of the original position, leads to the following problem: the weak and generally accepted assumptions that ought to have described the reasonable constraints of arguments for justice become strong and constraints of arguments for justice become strong and controversial, and a theoretically defined conception of a "rational choice situation" is offered as the *only possible interpretation* of these constraints. The original position no longer serves as a *justification* of the two principles of justice; rather, it functions as an *illustration* of those principles whose justification has been presupposed.

That the original position is not a justificatory, but an illustrative device, is by now a shared criticism in the literature. These criticisms can be divided into two groups: according to the first, even the minimal interpretation of the original position, under the constraints of a "thin theory of the good" – the assumption that individuals are to be considered as moral persons whose ends are not ranked in value – is not minimal. Thomas Nagel points out that there are many groups of individuals, Marxists, conservatives, aristocrats, and hedonists, among others, whose conception of the good would contradict the conditions stipulated by Rawls for admission to the original position⁷. According to the second kind of criticisms, not the minimal, but the privileged description of the original position, constitutes the problem. These critics have questioned the so-called "neutrality" of Rawls' theory of primary goods (Schwartz); Rawls' criterion of intersubjective comparison in terms of average income (Wolf); the plausibility of Rawls' moral psychology, which presupposes risk aversion (Barber); the game theoretic and economic definition of rationality (Höffe); the use of the maximin as a decision procedure (Harsanyi).⁸ Both the minimal interpretation of the original position in light of reflective equilibrium and

⁷ Thomas Nagel, "Rawls on Justice", in: Norman Daniel (ed.), *Reading Rawls*, Oxford 1975, pp. 97ff.

⁸ Adina Schwartz, "Moral Neutrality and Primary Goods", *Ethics*, 83 (July, 1973), 294–307; Robert Paul Wolf, *Understanding Rawls*, Princeton 1977; Benjamin Barber, "Justifying Justice: Problems of Psychology, Politics and Measurement in Rawls", in: Daniles (ed.), *Reading Rawls*, 292–319; Otfried Höffe, *Ethik und Politik*, Frankfurt 1979, 227–243; John C. Harsanyi, "Morality and the Theory of Rational Behavior", *Social Research*, 4 (1977), 623–656.

the thin theory of the good, and its privileged description as a rational choice situation,⁹ are highly controversial, and do not provide a *prima facie* plausible interpretation of the principle of rational consent. Rather than engaging in a similar criticism, I want to consider what light Rawls' Dewey lectures throw on this problem of the justification of the original position.

2.1.2 Kantian Constructivism and the Original Position

Rawls' strategy in these lectures modifies the earlier one in *A Theory of Justice* at least in two essential respects. First, *A Theory of Justice*, the Kantian conception of the person was not used as a justificatory device in the construction of the original position. Rawls subsequently explained that he considered the original position to be an interpretation of the Kantian conception of autonomy (TJ, 252 ff.; 283 ff.). In the Dewey lectures, it is stated that "a particular conception of the person is an element in a reasonable procedure of construction, the outcome of which determines the content of the first principles of justice" (KC, 516). Whereas *A Theory of Justice* claimed to proceed from weak and generally accepted premises of rational and moral agency to a conception of the person, Rawls now proceeds from a conception of the person to the description of the original position. According to this new interpretation, Rawls attributes to persons not only the capacity for a sense of justice, and the capacity to form, to revise and to pursue rationally a conception of the good, but also two "highest order interests to realize and to exercise these powers" (KC, 525)¹⁰. The attribution to individuals of such interests is a strong moral assumption; for it does not follow from Rawls' definition of the first two capacities that individuals would also have these interests. It only follows

⁹ There is disagreement in the literature as to whether the "rational choice situation" can be described as a *collective* bargaining game at all, and whether it does not merely redescribe a *monological universalizability principle* of the sort proposed by Kant. Cf. Sidney Alexander, "Social Evaluation through Notional Choice", *The Quarterly Journal of Economics*, 4 (1974), 597–625; John Rawls, "Reply to Alexander and Musgrave", *Ibid.*, 633–655; J. Hampton, "Contracts and Choices: Does Rawls have a Social Contract Theory?", *Journal of Philosophy*, 36 (1980), 315–338. The criticisms of Musgrave and Hampton have convincingly shown, I believe, that the Rawlsian construction of rational choice is not sufficiently different than the utilitarian principle of adopting "for society as a whole the principle of rational choice for one man" (TJ, 27; 45). In section 3.1 below I will argue that if the Rawlsian theory of justice purports to be a rational choice model, it does so at the cost of epistemological inconsistency.

¹⁰ Cf. Rawls' initial discussion of "The Rationality of the Parties", in TJ, § 25 with the Aristotelian theory of motivation in § 65. In "Kantian Constructivism and Moral Theory", the Aristotelian principle of motivation becomes an element in the construction of the concept of the person, not its elaboration, as it was in TJ.

that it is not excluded that individuals be motivated by such an interest, but not that they are in fact motivated by it. Rawls now seems to imply that the construction of the original position is only plausible for those who share a moral concept of the person and who have an interest in realizing this in society.

Whereas in *A Theory of Justice*, it was assumed that proceeding from a minimal interpretation of the "reasonable conditions for arguments of justice", one could reach the *rational* solution to the problem of justice, Rawls now admits that only those who find the moral conception of the person reasonable and worthy of realization as an ideal, will find the specific construction of rational autonomy, embodied in the principles of justice, also reasonable (KC, 528 ff.). In his earlier formulation, Rawls thought to be able to proceed from a "thin" definition of rational and moral agency to specify a rational choice situation. The new strategy of moral constructivism addresses individuals who are ready to share a normative conception of the person, also because this conception seems to articulate the most reasonable solution to certain dilemmas in a democratic culture. "The aim of political philosophy is to articulate and to make explicit those shared motives and principles thought to be already latent in common sense, or, as is often the case, if common sense is hesitant and uncertain and does not know what to think to propose to it certain conceptions and principles congenial to its most essential convictions and historical traditions" (TC, 518 ff.).

Second, this consideration is quite different that what Rawls had earlier meant by "reflective equilibrium": the "we" that referred to the theorist as well as to the reader was never given an explicit *hermeneutic identity* in *A Theory of Justice*. Rawls could even claim that to see our place from the perspective of the original position was "to see it sub specie aeternitatis" (TJ, 587; 637). The standpoint of moral constructivism is now described as one in which "we are not trying to find a conception of justice suitable to all societies regardless of their particular social and historical circumstances. We want to settle a fundamental disagreement over the just form of basic institutions within a democratic society under modern conditions" (KC, 518).

What light does this modification throw on the problem of the form-content dialectic? Upon Rawls' own admission, the criteria of selection guiding the description of the original position are not minimal and general; they are strong moral assumptions whose plausibility depends upon their *hermeneutic capacity to interpret* for social and political agents an answer to a certain impasse in their culture. Given the admission, that the original position formulates the conditions of *reasonable* rational consent on questions of justice in light of a strong moral conception of the person and a certain hermeneutic-historical understanding of democratic societies,

then it cannot be said to interpret the conditions of rational consent *simpliciter*. This conception of the person may not speak to those who do not share or sympathize with its motivational structure; others may find it too formal, too repressive or too individualistic and private. Equally, the impasse in democratic culture may not be interpreted as a conflict between two conceptions of liberty – the liberties of civic life and those of public life (KC, 519ff.) –, but as a conflict between political democracy and the nondemocratic control of social and economic life by corporations and the state serving their interests (Lasch, Wolin); as the conflict between the increasing complexity of social life and the “unmanageability” of democratic participation demands (Samuel Huntington); as the incompatibility of cultural hedonism, generated by an excessive democratic culture, and the work ethic (Daniel Bell)¹¹. By making his standpoint more hermeneutically and historically specific, Rawls raises the question of the relationship of normative philosophy to social-theoretical assumptions, but does not resolve it.

The dialectic of form and content in the Rawlsian theory of justice can be stated in the form of a dilemma: either to proceed from weak and substantively non-controversial assumptions, and to forfeit the claim to a formal solution to the problem of justice, or to proceed from strong and substantively controversial assumptions, but to forfeit the claim that the formal solution to the problem of justice can be the *unique* procedural interpretation of the principle of rational consent. By shifting from the weak to the strong privileged description of the original position, *A Theory of Justice* offered a formal but highly controversial answer to the problem of justice. In light of Rawls' new strategy the question would be: can these new admissions and modifications allow a formal, game-theoretic answer to the problem of justice at all? (Cf. section 3.1 below).

2.2. The Justification of the Ideal Speech Situation in Habermas' Theory of Discourse

In the previous section, I have attempted to show that the justification of the original position within the Rawlsian theorem of justice undergoes a modification which can be briefly described as “from the rational to the reasonable”. One could say that Habermas' theory of practical discourse

11 For the first interpretation, cf. articles by Christopher Lasch, S. Wolin, L. Goodwyn and D. Dickson in *Democracy*, v. I, No. 1 (January 1981); for the second, Samuel Huntington (ed.), *No Easy Choice: Political Participation in Developing Countries*, Cambridge, Mass. 1976; and for the third, Daniel Bell, *The Cultural Contradictions of Capitalism*, New York 1976.

still operates with the assumption that the rational is prior to the reasonable, or that the formal conditions of rational argumentation and communication should define the reasonable normative posture of agents who are capable of speech and action. The “cognitive ethics of language” is posterior to, in the sense of being grounded by, the universal pragmatic analysis of the conditions of rational communication and argumentation¹².

Following a suggestion made by K. H. Ilting, I will distinguish between four possible interpretations of the ideal speech situation in Habermas' writings: 1. the ideal speech situation as the model of an ideal communication; 2. the anticipation of the ideal speech situation as the condition of the possibility of linguistic communication; 3. the anticipation of the ideal speech situation as the condition of the possibility of a discourse; 4. the ideal speech situation as a sufficient condition for the attainment of a rationality motivated consensus on controversial validity claims¹³. The parallelisms between Habermas' justification of the ideal speech situation and the Rawlsian justification of the original position can be established as follows: while the first and the second interpretations of the ideal speech situation correspond to Rawls' conception of a “thin theory of the good” and “reflective equilibrium” respectively, the third and fourth interpretations parallel the privileged description of the original position as a rational choice situation. I will argue that according to all three interpretations, the justification of the ideal speech situation is circular, for it presupposes, rather than grounding, the normative validity of communicative ethics.

2.2.1 The Ideal Speech Situation as a Thin Theory of the Good

According to Habermas, „Die kontrafaktischen Bedingungen der idealen Sprechsituation erwiesen sich als Bedingungen einer idealen Lebensform“ (HL, 139). These counterfactual conditions are: a symmetrical distribu-

12 In his critique of the „dezinistische Restproblematik“ in K. O. Apel's reconstruction of communicative ethics (Apel, „Das Apriori der Kommunikationsgemeinschaft“, in: *Transformation der Philosophie*, v. II, Frankfurt 1973, 405ff.), Habermas claims that „Die soziokulturelle Lebensform der kommunikativ vergesellschafteten Individuen erzeugt in jedem Interaktionszusammenhang den „transzendentalen Schein“ reinen kommunikativen Handelns, und zugleich verweise sie jeden Interaktionszusammenhang strukturell auf die Möglichkeit einer idealen Sprechsituation . . .“ (L pr 152). My argument in this section is that the „dezinistische Restproblematik“ remains in Habermas' methodological proceduralism as well, but that it is a reasonable „Restproblematik“, whose consequences we should accept.

13 Karl-Heinz Ilting, „Geltung als Konsens“, *Neue Hefte für Philosophie*, 10 (1976), 22–50.

tion of chances to choose and to lead speech acts, and to assume various dialogue roles. While these first two conditions provide for equality of opportunity to use communicative, constative and partially regulative speech acts, the third formal precondition stipulates that participants in a discourse be "truthful": „die Sprecher weder sich noch andere über ihre Intentionen täuschen dürfen“ (HL, 138). The fourth precondition stipulates that only those participants can be allowed into the ideal speech situation, „die als *Handelnde* gleiche Chancen haben, Representative zu verwenden“ (HL, 138). Neither the third nor the fourth preconditions can be satisfied by contexts of linguistic communication alone, but presuppose action contexts as well. For, truthfulness is a quality that pertains not only to linguistic expressions, but to the relationship between saying, doing and meaning. Likewise, the fourth precondition, guarantees reciprocity in behavioral expectations by excluding norms that are binding only one of the parties¹⁴.

The ideal speech situation gives a linguistic determination for what we "usually" understand by truth, justice and freedom: *truth* is to be defined as the consensus reached by participants under the conditions of an ideal speech situation. This implies that the rationality of the parties can be defined as their capacity and willingness to engage in, comply with, and accept as binding the outcomes of such communication. The ideal speech situation defines the meaning of justice in terms of the *equal right* of all participants to an *equal* distribution of the use and initiation of speech acts. The universal pragmatic meaning of *freedom* would be the *right* to accept as binding only the "force" of the better argument; freedom means the suspension of all internal and external constraints of action that lead persons to act in certain ways, and the right to follow the force of reason only.

As with Rawls' thin theory of the good, the Habermasian original position does not *justify* the moral and normative content of the theory,, rather it serves to *illustrate* it. The symmetry condition already stipulates a prior normative commitment to the *equal right of all* to be considered as rational and free beings. In fact, Habermas' normative assumptions are stronger

14 The characterization of this fourth precondition is confusing: while the condition of symmetry and reciprocity in the use of speech acts, are built into the very logic of the ideal speech situation, the requirement that those who are to be admitted to the ideal speech situation also have, as acting agents, the same chances to use representatives, seems to limit participation in discourses to those who are in a position of equality in real social life. First, how does Habermas interpret „gleiche Chancen“ as a concrete social norm? Does it mean the *right* to equality, or the *opportunity* of equality, or both? Second, while most norms of institutional life imply reciprocity of duties and privileges, they do not imply *equality* of duties and privileges, for e. g. between a doctor and his patient, there is reciprocity but no equality. Habermas interprets reciprocity as equality in an undefined and obscure way.

than those of Rawls', for they already define a certain moral conduct, that of truthfulness, and a certain form of life, one in which the complete reciprocity of the parties in power and authority relations is presupposed. The universal pragmatic interpretations of rationality, justice and freedom are indeed *interpretations*, which presuppose the normative validity of communicative rationality, of equality of opportunity, of reciprocity and truthfulness. But what grounds have been given to choose this particular normative interpretation, if it was first the outcome of the ideal speech situation that would establish norms which we would have to consider binding? The ideal speech situation is a circular construction; it presupposes those very norms whose validity it was supposed to establish.

2.2.2 The Ideal Speech Situation as Reflective Equilibrium

Habermas could admit this circularity, but, like Rawls, argue that the validity of these norms need not be first established, for they are always already presupposed. Just as the weak description of the original position claimed to make clear to us those "considered judgments duly pruned and adjusted" (TJ, 20; 38), Habermas argues that the ideal speech situation is anticipated in the structure of all empirical speech (HL, 136; Wth, 258) and of all human interaction (BWU, 95; L pr, 153).

Upon a first reading this argument is trivially implausible. Non-symmetrical, non-reciprocal, deceitful and irrational speech and interaction are the norm, rather than the exception in human life. If this were not the case, there would be no need for normative philosophy either. All would be as it ought to be.

This objection is trivial because universal pragmatics establishes the conditions for the possibility of speech and interaction in the transcendental sense. Just as the time perception of schizophrenics or the space perception of cubists do not constitute arguments against Kant's pure forms of intuition, so the presence of distorted communication patterns does not constitute an argument against the universal pragmatic interpretation of the possible conditions of human speech and interaction. Universal pragmatics analyzes the deep structure of the possibility of specifically human modes of speech and communication.

Granted, however, that this deep structure is not one that we as acting and speaking subjects are immediately aware of, but one which we can anticipate and understand with the help of adequate means of construction (HL, 136), the problem still remains: even if general conditions of symmetry and reciprocity are presupposed by all human speech and interaction as

structural conditions¹⁵, these structural conditions do not imply a *unique semantic content*¹⁶. It does not follow that symmetry and reciprocity ought to be interpreted such as to mean the right of *all* to symmetrical and reciprocal participation, conduct, behavioural expectations and the like. We must presuppose a specific concept of *equality*; we must first normatively stipulate a specific content and interpretation in light of which the *identity* of those who are *entitled* to symmetry and reciprocity is defined. The claim that *all* humans, insofar as they are capable of speech and action, are entitled to the full symmetry and reciprocity of the conditions defined by the ideal speech situation, is itself a normative commitment that is not justified by the ideal speech situation, but rather one, in light of which, the structural conditions of the ideal speech situation are given a specific semantic interpretation. Whereas Rawls' procedure of "reflective equilibrium" refers to an already *shared* semantic interpretation of normative concepts, Habermas falsely assumes that this semantic content can be *univocally deduced* from universal pragmatics. But as A. Wellmer as well points out, a universal pragmatics cannot justify the step of abstraction which defines the transition to an universalistic ethics, because it presupposes it methodologically¹⁷. Universal pragmatics proceeds at a level of abstraction from the standpoint of which each and every individual is a being capable of consensus. Universal pragmatics reconstructs a system of rules which will be recognized to be binding by acting and reasoning agents of a certain sort. These have attained, or can attain, a level of moral and practical *Bildung* that corresponds to the level of abstraction of universal pragmatics.

2.2.3 The Privileged Description of the Ideal Speech Situation

There is yet another interpretation given to the ideal speech situation by Habermas. The ideal speech situation is said to anticipate the possibility of, and to define the formal procedures to be followed within; special modes of argumentation called Diskurs. In discourses, validity claims which have become problematical are examined. In such argumentation situations, facts and norms lose their bindingness, and are viewed under

15 Cf. Alvin Gouldner, "The Norm of Reciprocity: A Preliminary Statement", *American Sociological Review*, v. 25 (April 1960), 161–178.

16 E. Tugendhat also distinguishes between „semantischen d. h. bedeutungskonstituierenden und pragmatischen Regeln“ in discourses, „Zur Entwicklung von moralischen Begründungsstrukturen“, 6 ff. Our arguments differ, however, in that I claim that pragmatic rules of discourse – symmetry, reciprocity – themselves already presuppose a semantic content.

17 Albrecht Wellmer, *Praktische Philosophie und Theorie der Gesellschaft: Zum Problem der normativen Grundlagen einer kritischen Sozialwissenschaft*, Konstanz 1979, 50.

the aspect of possible existence (Wth, 214). The acting and speaking subject "abstracts"¹⁸ from the given constraints of the world shared by others, to assume an attitude of inquiry and reflection both with respect to what is said to exist, and with respect to what is claimed to be right (Wth, 213 ff.; Lpr, 174 ff.).

The ideal speech situation anticipates the possibility of discourses, in the sense that, once the universal pragmatic presuppositions of everyday communication and interaction – like truth, truthfulness, and rightness – become problematical, individuals are led to assume a counterfactual attitude towards these presuppositions: one is led to suspend belief in the truth of assertions, the truthfulness of one's partners, and the rightness of norms, and is led to question them argumentatively. As a counterfactual in which the reestablishment of consensus is anticipated, the ideal speech situation is a quasi-empirical fact: „Die ideale Sprechsituation ist weder ein empirisches Phänomen noch bloßes Konstrukt, sondern eine in Diskursen unvermeidliche reziprok vorgenommene Unterstellung“ (Wth, 258).

The ideal speech situation also *defines* the formal properties of argumentation strategies which are to lead to a rational and grounded consensus. In addition to the symmetry and reciprocity conditions discussed above (2.2.1), Habermas introduces a further formal requirement: the equal right of participants to challenge, to shift, and to revise the level of accepted justification procedures (Wth, 250). Discourses are procedures of argumentation characterized by a radicalization of the self-reflection of acting and speaking subject. Concrete forms of action, the norms legitimizing them, the warrants grounding these norms, and the level of justification presupposed by these warrants can all be called into question (Wth, 253).

I have called this third interpretation of the ideal speech situation a "privileged description", because Habermas, like Rawls, argues that if a rational procedure of deliberation on normative questions can be defined, then the norms arrived at through such procedures would have to be *rationally consented to by all*. But as rational deliberation procedures, discourses are not normatively neutral; the rationality of the procedures is dependent upon the reasonableness of the norms presupposed by it, and not vice versa.

Habermas now admits that discourses are the result of learning processes that are dependent on experience (Wth, 246), and attempts to give an evolutionary justification of the ideal speech situation. Learning processes

18 It is unclear whether Habermas means more by "abstraction" than the simple suspension, or bracketing, of the constraints of action, and assuming an attitude of "reflection". For a detailed analysis, H. Schnädelbach, *Reflexion und Diskurs: Fragen zu einer Logik der Philosophie*, Frankfurt 1977, 135–171.

are both ontogenetic and phylogenetic. While in the history of the development of the individual, discourses correspond to the speech and interaction competencies of post-conventional personality types, in the history of the development of the human species, discursive justification procedures are first initiated by modern bourgeois natural right theories, and institutionalized in part by the legal system, by science, and by democratic-parliamentary procedures¹⁹.

Even this new strategy, however, cannot *justify* "the step of abstraction" that leads to communicative ethics: either Habermas is claiming that because individuals can attain the level of post-conventional ego-identity, communicative ethics is justified – a procedure, which L. Kohlberg calls from "is to ought"²⁰ –, or he is merely saying that communicative ethics is historically contingent, because it is dependent on certain empirical developments. If the first argument is chosen, and indeed it would have to be, for it is the only one that raises a normative claim, then this amounts to assuming that the conditions for the possibility of the development of communicative ethics can serve as its justification. This would be analogous to saying that because individuals *can* act in these ways, they *ought* to do so. But while ought implies can, can does not imply ought. Evolutionary argumentation strategies may define the *range of possibilities* from which one can select, and also establish the *irreversibility* of certain courses of linguistic and interactive development, but from within the stage of a *shared level* of cognitive and moral competence, evolutionary strategies cannot establish one alternative to the exclusion of others. From the standpoint we have reached, we cannot know what will turn out to be the level of moral competence or personality type to be favored by evolutionary development. Since we are not at the end of history, but within it, we can only know what our range of possibilities are and at what costs a certain course of development may be reversed.

Discourses are not the unique normative procedures that the evolutionary development of individual and collective competencies allow. They are ones which are preferred to others on account of certain prior normative considerations. This becomes amply clear in Habermas' critique of Rawls: Rawls, according to Habermas, considers needs to be the "natural properties" of individuals, inaccessible to public discussion (R HM, 84). A *Theory of Justice* is based on the assumption that „Jeder Einzelne soll

19 For a general statement, Habermas „Historischer Materialismus und die Entwicklung normativer Strukturen“, in: *RHM*, 15ff.; on the relationship of ontogenetic development and discourse theory, „Moralentwicklung und Ichidentität“, *Ibid.*, 84ff.; on the relationship of phylogenetic development and discourse theory, „Legitimationsprobleme in modernen Staat“, *Ibid.*, 278ff.

20 Larry Kohlberg, "From is to Ought", in: Mischel (ed.), *Cognitive Development and Epistemology*, New York 1971, 165ff.

monologisch die Verallgemeinerungsfähigkeit der jeweiligen Normen prüfen" (*ibid.*). Rawls is also criticized for presupposing a conception of happiness that mirrors the way of life of private consumers, but the "pursuit of happiness" might one day mean something different: not accumulating material objects which one disposes of privately, but bringing about social relations in which mutuality predominates and satisfaction does not mean the triumph of one over the repressed needs of other (R HM, 292).

Habermas proposes a dialogic instead of a monologic model of need constitution, and prefers a system of justice in which individuals are not allowed to pursue their "thin" conceptions of the good at the sake of "repressing the needs of others". But evolutionary arguments can provide evidence neither *for* or *against* such a criticism. Rawls, like Habermas, presupposes a post-conventional personality type, a human being who has acquired the capacity to accept norms of action only by reflecting upon the correct procedures of their generation. Furthermore, Rawls also presupposes the institutional and collective achievements of western democracies, but maintains that "the triumph of one over the repressed needs of the other" can only be avoided by the establishment of the two principles of justice. Evolutionary argumentations cannot tell us which interpretation of the post-conventional personality type is the one that must be chosen: this depends upon the plausibility, the enlightening and emancipatory potential of conceptions of personality and the theory of the good implied by them.

My conclusion in this section is a largely negative one: the assumptions leading to the justification of counterfactual abstractions are, in the case of both Rawls and Habermas, equally "strong" and "privileged"; each can elicit rational dissent and disagreement. Furthermore, because of the circularity of their justification procedures, in each case we confront assumptions among which we cannot arbitrate by recourse to proceduralism, for whether we accept the procedure itself depends upon whether we accept the norms that are presupposed by the procedure.

3. The Dialectic of Form and Content II: Social-theoretical Assumptions and Normative Philosophy

The second aspect of the form-content dialectic to be examined concerns the relationship of sociological presuppositions to normative principles in counterfactual procedures. I will first state this problem as it arises in the Rawlsian theory of justice: the initial counterfactual description with which the theory begins must be supplemented by some additional as-

assumptions concerning the content of the fictive choice procedure. In introducing such assumptions, the theory must remain consistent. It must be shown that these assumptions either follow from, or are not inconsistent with, a. the principle of rational consent and b. the privileged description provided at the outset. Now, if these assumptions deductively follow from b., then the procedure cannot be described as one of choice, for how one chooses, or what one chooses would be already contained in the original description, and the theory would merely deductively elaborate its consequences. But if these assumptions do not deductively follow from the original privileged description, then on what grounds does the theory resort to them, and are they consistent with other assumptions? In the one instance, the theory would become deductive and dogmatic; in the other, it would run the risk of inconsistency.

3.1 The Status of "General Facts" in the Rawlsian Theory of Justice

There is a major inconsistency in the Rawlsian theory of justice: social and political agents under the "veil of ignorance" would have to know either less or more than what Rawls allows them to know, but they cannot know *all* and *only* what Rawls allows them to know. Rawls commits himself to the following two assumptions: individuals deprived of any knowledge of their social positions, natural and psychic endowment, the economic or political condition of their society, its level of culture and civilization, and its point in history, also know *general facts* about human society. They understand the principles of political science and of economics, the basis of social organization and the laws of human psychology. "Indeed, the parties are assumed to know whatever *general facts* affect the choice of the principles of justice" (TJ, 137; 160).

It has only been noticed by Robert P. Wolff that these two premises are epistemologically incoherent²¹. Rawls assumes that there can be a sort of *general knowledge* concerning society, politics, economics and psychology, the possession of which would not inform individuals of the *particulars* of their social-historical conditions, the level of civilization of their society, its political and economic condition. But the concept of "particular facts" refers to two quite different kinds of propositions, which do not have the same epistemic status. While Rawls can justify the exclusion of knowledge of one kind of proposition under the conditions of the original position, he is inconsistent in excluding the second kind.

Under "particular facts", Rawls understands *existentially indexical* as well as *socially and historically indexed* knowledge. An example of the first

21 R.P. Wolff, *Understanding Rawls*, 119ff.

kind of proposition would be: "Individual A belongs to the third generation of a well-to-do Protestant family", where A refers to Jones, or James. An example of the second kind of proposition is given by a statement like, "The crash of the world-exchange market in 1929 was the beginning of a period of great economic instability and unemployment for the capitalist countries of that period". Rawls is consistent in excluding propositions of the first kind under the conditions of the "veil of ignorance", for such knowledge is inconsistent with the requirement of justice as fairness. We should be willing to reason on the first principles of justice by excluding any specific knowledge we might have about our own socio-economic condition, psychic abilities, natural endowments, social and historical lineages, etc. But by confusing such existentially indexical knowledge with socially and historically indexed knowledge, Rawls is led to assume a. that the social sciences can provide us with socially and historically non-indexical laws and generalizations, and b. that the "general facts" relied upon by Rawls himself, and which are provided by marginalist economic theory, game-theory, Pareto sociology, etc. are themselves non socially and historically specific. Both assumptions are, if not erroneous, at least highly debatable.

When we consider the first point: those committed to a hermeneutic methodology in the social sciences, like Peter Winch and Charles Taylor for example, would question that it is ever possible to formulate general laws in history and sociology that would be more than trivial²². Ethnomethodologists (Garfinkel, Cicourel) would debate that a theoretical understanding of the social world, such as abstracts from the perspective of the participant, is possible²³. They would maintain that social knowledge is essentially knowledge of the participant and not of the observer. Marxists would reject that one could formulate laws and generalization in the social sciences which would not specify the social and historical range of their own validity – for example, neither the profit maximization nor the risk aversion motives, they would claim, can be attributed to social agents under precapitalist economic conditions²⁴. Consequently, from the standpoint of a hermeneutic, ethnomethodological or Marxian social theory, the propositions of game theory, marginalist economics and

22 Peter Winch, *The Idea of a Social Science and its Relation to Philosophy*, London 1958; P. Winch, "Understanding a Primitive Society", in: *Rationality*, Oxford 1970, 78–112; Charles Taylor, "Interpretation and the Sciences of Man", *The Review of Metaphysics*, (1971); Taylor, *Erklärung und Interpretation in den Wissenschaften von Menschen*, Frankfurt 1975.

23 H. Garfinkel, *Studies in Ethnomethodology*, New York 1967; A. Cicourel, *Method and Measurement in Sociology*, New York 1964.

24 Polanyi, C.M. Arensberg and H. W. Pearson (eds.), *Trade and Market in Early Empires*, Glencoe, Ill. 1975.

Pareto sociology would themselves be viewed as highly socially and historically indexed.

By ignoring such debates concerning the foundations, methodology and epistemology of the social sciences²⁵, Rawls falls into a major inconsistency: if it can be argued that the kind of general facts that Rawls allows into the theorem of justice are no more general than other facts, provided by different theoretical traditions in the social sciences and which Rawls does exclude from the theory of justice, than social agents could not know all and only what Rawls assumes they could not know. Either Rawls has to provide a principled argument as to why marginalist economic theory is more general than Marxian economics, or Pareto sociology more general than Weberian sociology, or he would have to admit that social agents could also dispose of information provided by these other theoretical traditions.

It might appear that Rawls' new strategy in "Kantian Constructivism and Moral Theory" has solved this difficulty. It is no longer assumed that individuals do not possess socially and historically specific knowledge about their societies, its development, culture and civilization. Despite this admission, however, the "veil of ignorance" remains as thick as before. According to Rawls, "whether certain facts are to count as reasons for justice and what their relative force is to be can be ascertained only on the basis of principles that result from the construction" (KC, 565). It is up to the parties in the original position to decide how simple or complex moral facts are to be, and to agree upon the number and complexity of principles which identify "which facts are to be accepted as reasons for justice by citizens in society" (KC, *ibid.*). The original position involves an *epistemic* as well as a *moral* choice.

Rawls does consider the following objection: "what is the reason for limiting the parties to those beliefs and not allowing them to take into account all true beliefs? (KC, 541) Rawls' answer is that, if one is to have a "public" conception of justice, then such a conception must be limited to those generally shared beliefs and convictions in society, which can accommodate different points of view. Procedures and methods of scientific evidence and verification are included in such generally shared beliefs and convictions (KC, 540-541). If, however, there are *no* such generally shared beliefs and convictions in the community of social scientists about methods of scientific evidence and verification, then is not Rawls simply recapitulating the

25 Cf. R. Bernstein, *The Restructuring of Social and Political Theory*, Philadelphia 1978; A. MacIntyre and D. Emmet (eds.), *Sociological Theory and Philosophical Analysis*, London 1970; Alfred Schutz, *The Phenomenology of the Social World*, G. Walsh and F. Lehnert (trans.), Evanston 1967; J. Habermas, *Zur Logik der Sozialwissenschaften*, Frankfurt 1970; Anthony Giddens, *Studies in Social and Political Theory*, London 1977.

old positivist belief that social science is what the philosopher of science, committed to the unity of science, thinks it ought to be?

Rawls' final answer to this objection would be the following: If a theory of justice is to be a part of political culture, then it must rely on information on moral matters that goes back a long time, and "is available to the common sense of any thoughtful and reflective person" (KC, 566), for "advances in our knowledge of human psychology and social theory might be relevant . . . in the *application* of the principles of justice, as opposed to the *adoption* of principles in the original position" (KC, 566)²⁶. When we consider Rawls' theory of "primary goods", however, it becomes clear that social scientific information is relevant not only in the *application*, but also in the *choice* of the second principle of justice.

In the Dewey lectures, Rawls admits that "the question whether the account of primary goods is a matter for social theory or depends essentially on a conception of the person" was not satisfactorily discussed in *A Theory of Justice* (KC, 527). In the new version of Rawls' theory, the account of primary goods ought to depend upon the moral conception of the person. It should be assumed that individuals would choose primary goods like basic liberties, freedom of movement and occupation, the powers and prerogatives of office and of positions of responsibility, income, wealth and the social basis of selfrespect, for these "are generally necessary as social conditions and all purpose means" (KC, 526) to enable humans to realize and to exercise their moral powers and to pursue their own ends. But that individuals would choose such primary goods in *general* is not equivalent to the claim that individuals would choose a *specific* pattern of the *distribution* of these goods as formulated in the second principle of justice: social and economic inequalities are to be arranged so that they are to the a. greatest benefit of the *least advantaged*, consistent with the just savings principle and b. attached to offices and positions open to all under conditions of fair equality of opportunity (TJ, 302; 336). The second principle of justice presupposes that a specific interpretation of primary goods in light of marginalist economic theory can be given. For if individuals cannot be said to have a preference curve for ordering their desires and wishes, and if these desires and wishes cannot be correlated with quantitatively commensurable criteria of intersubjective comparison, like "average income and wealth", then comparisons like "better off" and "worse

26 In view of this claim it must be asked whether Rawls can explain or even admit to changes in our moral concepts and principles in light of "advances in our knowledge of human psychology and social theory". Morality, like any other system of belief, can be studied as a social reality, which develops and changes in relation to other systems of belief, socio-economic factors, etc. Cf. E. Durkheim, *On Morality and Society*, R. Bellah (ed.), Chicago 1973, 63 ff.

off" would be meaningless²⁷, and there would be no generally agreed upon criterion for evaluating the justice of major social and economic arrangements (TJ, 143 ff.; 167 ff.). Thus, if Rawls abandons the specific interpretation of primary goods in terms of economic theory, then he cannot argue that a specific pattern of the distribution of these goods, one which benefits the least disadvantaged, is the correct formal resolution to the problem of rational choice. But if Rawls wants to claim that the moral interpretation of primary goods as "generally necessary social conditions and all purpose means" also entails the adoption of a specific pattern of the distribution of these goods, then the second principle of justice could not be presented as the unique outcome of a rational choice situation. This specific distribution pattern would be the result of no choice at all; it would merely deductively stipulate that pattern of distributive justice considered to be most just by the theorist.

To summarize: There is a major inconsistency in the Rawlsian theory of justice, for individuals cannot know all and only what Rawls allows them to know. On the one hand if they were to know more, the rational choice situation would be hopelessly complicated, involving disagreements not only about the facts themselves, but about principles of verification and evidence as well, and a formal solution could not be attained. On the other hand if these agents were to know less, if they were ignorant, say of game-theoretic reasoning, then there could be no choice at all. Finally, if all general facts which Rawls wants to rely upon are said to be entailed by the normative definition of the person and the counterfactual abstraction with the theory begins, then the device of counterfactual choice procedure is superfluous, for the theory then merely elaborates a specific normative

27 In: TJ, Rawls repeatedly returns to the question of how we actually should define "worse off" and "better off" in society, and resolves this with the remark that "These comparisons are made in terms of expectations of primary social goods. In fact, I define these expectations as the index of those goods which a representative individual can look forward to" (TJ, 92; 112). But how is the "representative individual" to be defined? According to Rawls, individuals in the original position would either choose a relevant social position, and count all "those with the average income and wealth of this group, or less" as representatives of the least disadvantaged individual (TJ, 98; 118). Or, they would define the representative individual "solely in terms of income and wealth with no reference to social position" (Ibid.). Now, individuals in the original position could not do the first, for under the constraints of the veil of ignorance they lack knowledge of social stratification patterns in their society. They cannot know whether stratification in their societies runs along class, sex, race or status distinctions, and which relevant social positions would be entailed by the patterns of division of labor in their society. They could actually only choose income and wealth as the index for defining the "least advantaged" representative individual. But this means that the primary goods are not all purpose moral means for individuals to realize their ends; they presuppose a specific interpretation in light of economic theory.

conception of justice, whose plausibility in no way depends upon its being the outcome of a choice procedure.

3.2. Legitimation Crises and the Theory of Practical Discourse

The radically divergent social-theoretical assumptions presupposed by Rawls and Habermas are made nowhere more evident than in the following counterfactual. „Wie hätten die Mitglieder eines Gesellschaftssystems bei einem gegebenen Entwicklungsstand der Produktivkräfte ihre Bedürfnisse kollektiv verbindlich interpretiert und welche Normen hätten sie als gerechtfertigt akzeptiert, wenn sie mit hinreichender Kenntnis der Randbedingungen und der funktionalen Imperative ihrer Gesellschaft in diskursiver Willensbildung über die Organisation des gesellschaftlichen Verkehrs hätten befinden können und wollen“ (Lpr, 156). Discourses of legitimacy are procedures of argumentation through which the worthiness of a political order to be just is rationally evaluated. As opposed to the Rawlsian construct of a collective choice situation under the restrictions of the "veil of ignorance", the Habermasian counterfactual procedure of discursive argumentation does not exclude any relevant social knowledge and places no restrictions on the kind of social information that individuals can use in order to judge the legitimacy of a social and political order. A dialectic of form and content between social-theoretical assumptions and formal-procedural restrictions does not exist in Habermas' theory of Discourse, although many aspects of this theory would remain unintelligible if Habermas' diagnosis of the motivation and legitimation crises of late capitalist societies was ignored²⁸. In this case, one must ask whether the critical intentions of Habermas' social theory can be fulfilled by the proceduralism of his normative justification strategy.

In developing the counterfactual of discursive will formation on generalizable interests, Habermas intends to criticize decisionist, empiricist, and authoritarian answers to the authority and legitimation problems of late capitalist societies. That power relationships presuppose the cognitive validity of a set of legitimating beliefs, is contested by the decisionist

28 The role played by the concept of "needs" in Habermas' discourse theory has been ignored even in the most extensive treatments of the subject, cf. Reiner Wimmer, *Universalisierung in der Ethik*, Frankfurt 1980, 22 ff. From Habermas' description of the formal properties of discourses alone, it does not follow that the subject matter of discourse ought to be the truthful interpretation of common and consensually articulable needs (W. th, 252). The concept of "needs" serves a double function: first, it bridges the gap between socialization theory and the theory of practical discourse (L pr, 211–265); second, the ideal of consensual need articulations through discursive procedures has an unmistakable utopian dimension through which Habermas seeks to reformulate the Frankfurt School ideal of socialization without repression (RHM, 87; 110).

and by the empiricist (L pr, 133 ff.). While the decisionist denies the cognitive necessity of basing power upon truth, the empiricist denies that the exercise of power requires the truthfulness of the beliefs which justify it. The theory of practical discourse is designed to show that practical questions permit of rational confirmation and warranted assertability, even if they cannot be measured by an emphatic conception of truth. Vis à vis the empiricist who uncouples the functioning of power from the truth content of its cognitive component, Habermas maintains that truth and power cannot be disassociated without endangering the intersubjective processes of socialization and identity formation of the species that presuppose the cognitive justifiability of social norms.

Such a disassociation seems to be at work in late capitalist societies. In such societies religions, cosmologies, and worldviews that have previously had an identity securing function have lost their claim to cognitive significance, and mechanisms of meaning-generating and identity stabilizing social integration are increasingly uncoupled from the functioning of mechanisms of systems integration (the world economy, institutions of formal-rational law, administration). Discursive will formation on generalizable interests is suggested as a rational procedure through which norms and values that secure individual and collective identities can be generated without regressing back to mythological, religious, or cosmological forms of thought, or into apathetic socialization patterns for which the relationship of truth and power would become fully irrelevant. The norm of discursive will formation excludes these alternatives, and proposes collective participation and further democratization as the only desirable means through which the crises of individual and collective identity in late capitalist societies ought to be resolved.

If the theory of practical discourse is offered as an answer to the crises of late capitalist societies, is the counterfactual procedure of discursive will formation on generalizable interests adequate to serve the purpose of Habermas' critical theory of late capitalism? Habermas himself admits that conditions of empirical speech are rarely, if at all, identical with those of an ideal speech situation, and that furthermore *we lack criteria to judge* if under given circumstances an actual discourse has taken place, or if a consensus has not been attained under conditions of external constraints (Wth, 157 ff.). The gulf between the ideal and the actual, the normative and the empirical is unbridgeable. Second, is the discourse theory adequate to define the necessary *and* sufficient conditions of rational consensus? The equal right of all participants to use speech acts, to shift levels of discourse, to challenge levels of argument does not guarantee an insightful, unbiased, or intelligent exercise of these rights. This depends on the nature of the arguments advanced, on their quality, and rationality. What we have to know beforehand is whether these claims are true and rational.

That they are advanced in accordance with a formally correct procedure does not guarantee their rationality²⁹. Finally, given that we lack criteria to mediate between the ideal and the actual, and that the discourse model is inadequate to define the sufficient conditions of rational consensus, can we actually compare „die jeweils bestehenden normativen Strukturen mit dem hypothetischen Zustand eines *ceteris paribus* diskursiv gebildeten Normensystems“ (L pr, 156). Do we know in the name of which or what we are to criticize? Once again we confront an epistemological vacuum. The ideal is so counterfactual that it cannot serve as a criterion to evaluate the actual.

The radical counterfactualness of the discourse theory, which threatens to reduce it to the status of a “transcendent”, and not merely a “transcendental” illusion, is not due to the formalprocedural constraints of the ideal speech situation alone. The discourse model cannot serve as a critical criterion of democratic legitimacy, for it presupposes the possibility of a *self-transparent* and *self-identical collectivity*. This difficulty has been well-stated by G. Markus: “Since participatory democracy is here directly identified with the notion of an unlimited, all-embracing community of critical discourse . . . the theory in which real consensus appears not only as the basic *strategic* principle of democratic action, but also as the ultimate end of democratic life, seems to resurrect the idea of the subject, if not as the “bearer” of historical process than as the immanent “goal” of historical development”³⁰.

In this respect the counterfactual ideal of discursive argumentation is a paradoxical project: the necessity of discursive argumentation arises when through conflict and crises social and political agents challenge an established background consensus. Yet the very step of “abstraction” that leads such agents to engage in discourses, namely the virtualization of the constraints of action, can only take place when such agents are willing to suspend the motivating force and content of these real conflict situations. Discourses arise when the intersubjectivity of ethical life is *endangered*; but the very project of discursive argumentation presupposes the ongoing validity of a *reconciled* intersubjectivity.

By identifying participatory democracy with the notion of an ideal community of reconciled intersubjectivity, Habermas interprets *all diversity* of interests leading to conflict as illegitimate particularisms. Interests are “subjectivized needs” that have freed themselves from mutually shared traditions and values (L pr, 157). This means that a community of recon-

²⁹ This critique is most convincingly developed by A. Wellmer in an unpublished Manuscript, „Thesen über Vernunft, Emanzipation und Utopie“, Konstanz 1980; cf. Ilting, „Geltung als Konsens“, 35 ff.

³⁰ György Markus, “Practical-Social Rationality in Marx: A Dialectical Critique – Part 2“, *Dialectical Anthropology*, v. 5, No. 1 (1980), 10.

ciled intersubjectivity eliminates conflict and diversity of interests by allowing only universalizable need interpretations. This is a transcendent illusion, for, although the very project of an emancipated society means eliminating *certain structural* sources of interest conflict in present societies, a human society freed from all sources of conflict and diversity of interests cannot be conceived of without eliminating the radical plurality of ways of life, of cultural traditions and individual biographies which differentiate humans from one another. Such a society of reconciled intersubjectivity would be a self-identical collectivity, but not a human community. When the critique of the false universalism of particularistic interests is conceived as the desubjectivization of all interests, then those interests that are rooted in the pluralisms of autonomous life forms, living communities and self-conceptions are excluded from political discourse. Not only is the project of such a self-transparent and self-identical collectivity impossible, it is not even desirable.

Admittedly, Habermas distinguishes between the discourse principle of legitimacy and concrete principles of democratic organization. He accuses Rousseau precisely of having mixed up „Die Einführung eines neuen Legitimationsprinzips mit Vorschlägen zur Institutionalisierung einer gerechten Herrschaft“ (RHM, 279). Consequently, „Das hat die Demokratiediskussion bis auf den heutigen Tag verwirrt . . . Wenn man genau die politischen Ordnungen Demokratien nennt, die dem prozeduralen Legitimitätstypus genügen, dann können Fragen der Demokratisierung als das behandelt werden, was sie sind: als Organisationsfragen“ (RHM, 279). But how meaningful is this appeal to a procedural type of legitimacy? For, as I have tried to show in this essay, procedures of normative justification and legitimation are not normatively neutral constructs, but presuppose a moral theory of the person and of the good, as well as a certain sociological understanding of the demands and conflicts of the present. What distinguishes the Rawlsian legitimation procedure from the Habermasian theory of discourse are precisely assumptions concerning the meaning of human autonomy and of the good life, as well as a diagnosis of the dynamics, needs and tendencies of present democratic societies. The theory of discourse advocates the procedure of discursive will formation on generalizable interests, precisely because this would encourage democratization and repoliticization. But by distinguishing between procedures of legitimacy and problems of democratic organization, Habermas retreats into a meta-theoretical neutrality which is illusory. This leads me to the following negative conclusion: either the discourse principle of collective will formation on generalizable interests is a principle of participatory democracy, or it is an empty meta-theoretical procedure which cannot arbitrate among various first order theories of legitimacy, because for such an arbitration, more than a procedure, certain material assumptions and

criteria in light of which to judge procedures, are needed. On the other hand, if the discourse principle is not a meta principle of procedural legitimation, but a principle of democratic participation, then it is equally inadequate, for it identifies participatory democracy with the ideal of a unified and self-reconciled subjectivity.

4. Conclusion: Methodological Abstractions and Political Vision

The methodological withdrawal from politics with which modern political theory began can assume different forms. While Hobbesian proceduralism entails an *absolutization* of politics, Rawlsian proceduralism entails its *privatization*, and Habermasian proceduralism the *self-abolition* of politics through the ideal of a self-identical collectivity.

The Hobbesian methodological model of considering “men as mushrooms” would lead to the absolutism of the Hobbesian state, because by abstracting individuals from given bonds of community and social life, Hobbes would allow the sovereign to reconstruct social life only in accordance with the dictates of the political. Not only could the sovereign alone decide which civil and public associations were to be permitted among his subjects, but political reconstruction, argued Hobbes, would have to entail a semiotic and moral reconstruction as well³¹. Only the sovereign would be allowed to establish the meaning of those public signs by which men judged honor, dignity and infamy; only the sovereign would decide the public hermeneutic in light of which the sources of morality would be interpreted. Within the confines of civic, religious and moral life allowed by the sovereign, each subject was free to pursue his rational self-interest as long as it did not conflict with the like interest of others.

Rawlsian agents under the conditions of a “veil of ignorance” are likewise mushrooms insofar as they have no bonds to each other besides limited egoism: they take no interest in each others’ interests. But unlike Hobbesian agents, they jealously guard their private semiotic and their morality, and choose a political system whose virtue consists in allowing them to promote the conception of autonomous self-agency. Autonomy is irreconcilable with altruism; rationality must be limited to the possibility of forming and pursuing a consistent life-plan; and the demands of collective justice must not overstep the priority of liberty. Civic liberty and public justice can be best reconciled by a vision of politics in which the distributive role of the state does not overstep the boundaries of the private conception of the good cherished by autonomous agents. Admittedly, there is an unresolved tension in the Rawlsian conception of politics. The restric-

31 Thomas Hobbes, “Of the Rights of Sovereigns by Institutions”, ch. XVIII, *Leviathan*.

tion of the role of the state to the function of a general agency of distribution might create all sorts of political problems of control that would conflict with the privatism of the public conception of justice. Rawls remains in the uneasy limbo between a liberal privatist and a distributive participatory conception of politics. But by opting for the second principle of justice, which allows the existence of great social and economic inequalities even under conditions of Pareto optimality, he prevents the distributive state from becoming the interventionist state.

Habermas, by contrast, aims at a critique of the state which is interventionist, but which seeks to conceal the driving social interests behind its interventionism by resorting to a technocratic ideology that uncouples power from truth, or by encouraging apathetic socialization, which eliminates the necessity to question to legitimacy of power. However, since the thematization of needs and interests that are systematically blocked off from public discourse, are said to only occur under the conditions of an ideal speech situation, the model of participatory democracy, required by Habermas' crises theory, dissipates into a methodological chimera. The procedure for attaining rational consensus, interprets the meaning of consensus in such counterfactual categories, that the ideal of participatory democracy is reduced to the unattainable ideal of the real consensus of a self-identical collectivity. If all conflict of interests derives from the subjectivization of need interpretations, and if only those needs which are universalizable are legitimate, then the goal of rational consensus is a moral community freed from conflict, but also freed from politics and the experience of human plurality.

The dialectic of form and content analyzed in this essay was urged by the following question: has the time not come to discard the methodological illusions of modern political theory that the correct procedure of reasoning about political life is by first withdrawing from it through counterfactual abstractions?

Politische Philosophie als Gespräch. Überlegungen zum Stellenwert einer normativen politischen Philosophie

HENNING OTTMANN/München

I

(1) Im Zeitalter der verwissenschaftlichten Politik erfährt die politische Philosophie die Schwankungen der wissenschaftlichen Konjunktur. Sie wird totgesagt¹ und wieder gesundgeschrieben – manchmal von denselben Ärzten². An ihrem Krankenbett wird die besorgte Frage aufgeworfen, ob es sie denn noch gebe³. Und, wenn ihre – wie auch immer schwachen – Herztöne der Diagnose nicht entgingen, dann wird dem Patienten doch geraten, sich, angesichts seines angegriffenen Gesundheitszustandes, nicht länger zu übernehmen. Die alten Fragen, welche die ehemalige Königswissenschaft der praktischen Philosophie erörtert hatte, die Frage nach dem „guten“ Leben, der „guten“ oder gar „besten“ politischen Ordnung, nach der „guten“ Regierung, Verwaltung und Erziehung verfallen dem Verdikt der Unwissenschaftlichkeit. Allenfalls die Ideengeschichte bleibt da als Reservat historisch gesinnter Köpfe. Aber man wird dort, so sagt man, über das „Wirken moderner politischer Parteien, den amerikanischen Kongreß, das politische System der UdSSR oder über eine Vielzahl anderer Themen . . . nicht soviel . . . finden“⁴ wie in neueren empirischen Untersuchungen.

In den angelsächsischen Ländern hat der Patient politische Philosophie überlebt, und er erfreut sich, allen Todesanzeigen zum Trotz, einer erstaunlichen Gesundheit⁵. In der deutschen Politikwissenschaft ist sein

1 P. Laslett, Introduction, in: ders. (Hrsg.), *Philosophy, Politics and Society*. First Series (1956), Oxford 1975, VII.

L. Strauß, *What is Political Philosophy? And Other Studies*, New York 1959, 17.

2 P. Laslett/W. G. Runciman/Q. Skinner, Introduction, in: diess. (Hrsg.), *Philosophy, Politics and Society*. Fourth Series, Oxford 1972, 1ff.

3 I. Berlin, Does Political Theory still exist?, in: P. Laslett/W. G. Runciman (Hrsg.), *Philosophy, Politics and Society*. Second Series, Oxford 1962, 1ff.

4 R. Dahl, *Die politische Analyse*, München 1973, 7.

5 Dazu M. Richter, Editor's Introduction, in: ders. (Hrsg.), *Political Theory and Political Education*, Princeton 1980, 3–57. K. G. Ballestrem, *Politische Wissenschaft und Politische Philosophie: Überlegungen am Beispiel neuerer Entwicklungen in der angelsächsischen Philosophie der Politik*, in: P. C. Mayer-Tasch (Hrsg.), *Münchener Beiträge zur Politikwissenschaft*, Freiburg 1980, 86–100.