Joseph Heath

Rational choice as critical theory

Abstract Habermas has argued that many of the endemic socio-economic problems of Western society are either symptoms or products of a 'lopsided' process of cultural rationalization, one that has emphasized instrumental forms of rationality over communicative. But other than presenting a rather general typology of lifeworld pathologies, Habermas has not done much to specify what these problems might be, nor has he provided any 'middle-range' analysis of the mechanisms through which they might be generated. This paper discusses some of the ways in which, consistent with Habermas's general framework, rational choice theory can be used for precisely this task. In this analysis, rational choice theory is not presented as a comprehensive theory of action, but is employed as a critical-diagnostic tool that allows the theorist to identify undesirable social interaction patterns that arise from a broader instrumentalization of the lifeworld.

Key words: critical theory · Habermas · instrumental rationality · market failure · rational choice theory

Jürgen Habermas has argued that many of the endemic socio-economic problems of western society are either symptoms or products of a 'lopsided' process of cultural rationalization, one that has emphasized cognitive-instrumental forms of rationality over communicative. With the breakdown of the mythico-religious worldviews that once kept the institutions and norms of premodern society relatively free from critical scrutiny, the burdens of social integration were shifted over to the 'risky' mechanism of communicatively achieved
Rational choice as critical theory

Joseph Heath

thought of Adorno and Horkheimer led them to ignore the distinction between system rationality and action rationality, and therefore to identify these undesirable consequences, not as a property of the interaction system, but as a property of instrumental action itself. (This is simply a reverse application of the compositional fallacy that for a long time plagued Marxist thought.)

By maintaining a sharp distinction between types of social action and forms of social order, Habermas is able to reformulate this critique in a much more precise manner. At the level of action, Habermas distinguishes between two types of practical rationality. Depending upon the orientation that agents adopt, they will assign priority to either instrumental or communicative considerations in their practical deliberations. By defining instrumental rationality at this action-theoretic level, Habermas is thereby able to defer to standard rational choice theory for a detailed specification of the instrumental conception of rational action. This allows him to move directly to the task of articulating an adequate characterization of communicative rationality, which he does in The Theory of Communicative Action.

At the level of social order, Habermas distinguishes between two different mechanisms for the integration of social action. While agents who are reasoning communicatively are able to directly coordinate their plans through shared norms and expectations, agents who are reasoning instrumentally are only concerned with the realization of their own preferred outcomes. This means that instrumental action cannot be integrated through the harmonization of action orientations, and so can only be integrated through non-normative steering mechanisms that alter the environmental parameters structuring each individual’s choice problem. Habermas therefore suggests that a distinction be introduced between social integration and system integration, where ‘the former attaches to action orientations, while the latter reaches right through them’ (2:150). Habermas refers to any domain in which social integration provides the primary organizational mechanism as a lifeworld, and any domain in which system integration provides the primary mechanism as a system.

Since Habermas is not interested in the bad things that people do for instrumental reasons, but rather in the unanticipated consequences that are systematically generated by interactions among instrumentally rational agents, he argues that the critique of instrumental rationality should be supplanted by a critique of what he calls (somewhat misleadingly) ‘functionalist reason’. Functionalist reason, in this context, is intended to denote the organizational logic of systematically integrated domains of social interaction.

The most important doctrinal difference between Habermas and members of the early Frankfurt School is that Habermas does not succumb to the temptation to ‘demonize’ instrumental rationality (2:333). In his view, there is nothing intrinsically wrong with instrumental rationality, it is just that when a number of agents who are reasoning instrumentally get together, their interactions often have undesirable consequences (some of which are intentional, some of which are not). The Hegelian conception of reason that informed the consensus. The inability to achieve such consensus, due to an insufficient rationalization of the public sphere, gives rise to a persistent tendency to solve goal-attainment problems by transforming communicatively integrated domains of social interaction into appropriately structured subsystems of instrumental action. This has the effect of unburdening actors of the demanding requirements of communicative interaction, allowing them to take up the less cognitively and motivationally demanding instrumental orientation. But because actions in these subsystems are not directly norm-governed, it is not obvious that the outcomes generated by these interactions will necessarily be desirable. Worse, once the society becomes dependent upon the performance of these subsystems for coordination tasks that could not otherwise be accomplished (given the achieved level of communicative rationalization), they might not even be controllable. One of the tasks of a critical theory of society is to diagnose the problems produced by this over-reliance on instrumental rationality.

But other than presenting a rather general typology of lifeworld pathologies, Habermas has not done much to specify what these problems might be, nor has he provided any ‘middle-range’ analysis of the mechanisms through which they might be generated. I would like to discuss some of the ways in which, consistent with Habermas’s theoretical framework, rational choice theory can be used for precisely this task. In the following, I will (i) identify the position that Habermas assigns instrumental rationality in his overall analytic framework, (ii) examine the status of his ‘colonization of the lifeworld’ thesis, (iii) modify this analysis in order to eliminate some naive functionalism, (iv) give three examples of how rational choice theory can then be used to give a more concrete diagnosis of problems arising from instrumentalization, and (v) explain how a critical theory of this type differs from rational-choice liberalism, analytical Marxism and ‘critical hermeneutics’.
This means that, in Habermas's view, the modern social critic cannot score points just by uncovering some type of concealed or unnoticed instrumental action, she or he must actually specify in what way the consequences of this pattern of interaction are problematic. Once the problem is posed in this way, it becomes clear that there will often be nothing wrong with the use of a 'system' rather than a 'lifeworld' pattern of organization. For example, consider the following two approaches to solving a familiar allocation problem. At Northwestern University, course registration in the Faculty of Arts is organized by having students queue up alphabetically. This is a typical 'lifeworld' pattern of social organization — the allocation is directly normed, success depends simply upon following the rule. In the Faculty of Management, on the other hand, students bid for course positions in a competitive auction, using an initial stock of currency that is allocated by administrative fiat. This is a typical 'system' pattern of organization — success depends upon each student adopting an instrumental orientation, and developing a strategically sophisticated bidding strategy.

Ignoring for the moment its rather dubious pedagogical objectives, it should be clear that there is nothing inherently wrong with the 'system' solution to this problem. In fact, it undoubtedly provides a far more satisfactory allocation of resources. Queuing is a notoriously inefficient method of distribution, effectively rewarding individuals, not in accordance with merit or need, but in accordance with their willingness to waste inordinate amounts of time standing in line. Furthermore, the queuing solution is essentially premodern. The 'pure communicative action' resolution to the problem would be to gather all the students together in a room and have them decide cooperatively who should get which course positions. If widespread social criticism were to undermine the traditional authority of the queuing norm, and course registration were suddenly shifted over to the 'risky' mechanism of communicatively achieved consensus, then the 'system' solution of a competitive auction would begin to look not only attractive, but even indispensable.

So — we might then like to ask — what is the problem? Is there anything special about instrumental action that would lead us to think that these types of interactions will work out poorly, more so than human interactions generally? In order to answer this question at a theoretical level, we need a model of instrumental action. The potential payoff of such a model is double: by allowing us to specify the types of problems that will be systematically generated by instrumental interactions, such a model would not only give us principled reasons for challenging the status of instrumental rationality as a general resource for solving organizational problems, but also provide us with a specific characterization of problematic interaction patterns that we could then use to determine the etiology of concrete social pathologies.

There are a number of different models of instrumental rationality, the most sophisticated being preference-based decision and game theory (what I have been calling rational choice theory). But whatever the specific implementation, one thing that always shows up as a general problem with instrumentally rational action is that it cannot serve as the basis of a stable social order. The argument needed to establish this conclusively is somewhat complex: Habermas takes it as having been adequately demonstrated by Talcott Parsons, and does not elaborate upon it. The basic idea is that instrumentally rational actors cannot come up with a credible set of rules to constrain either their own or each other's actions, and are therefore 'stuck' with dramatically limited organizational capabilities. Without outside assistance, interaction systems among instrumentally rational actors would never leave the state of nature.

This means that in order to shift organizational objectives over to systems of instrumentally rational action, it is not sufficient for actors simply to abolish normative controls. In order to harness the potential of instrumental rationality, it is necessary to design a set of institutions that will specifically channel and regulate these interactions. As they are currently discovering in the former Soviet Union, markets do not arise spontaneously out of the individual pursuit of advantage, they are extremely sophisticated legal constructions. Eliminating constraints without replacing them with anything simply creates chaos, disorder, or anomie.

Habermas argues that the problem of introducing institutional constraints on systems of instrumental action is addressed at its most general level through the introduction of steering media, in particular, money and power. The idea behind these media is that they are easily quantifiable, so agents can use them as a substitute for utility in making their decisions. Institutional constraints can then be expressed in terms of these media, providing a very simple way of articulating what is to count as the 'rules of the game'. By constantly making slight changes in these rules, governments can then change what counts as a winning strategy, thereby indirectly controlling the actions of players whose only interest is in winning, in order to bring about a better overall outcome. Thus system integration is achieved through institutional constraints expressed in the form of steering media.
The basic idea underlying Habermas's critical diagnosis of the reliance upon instrumental rationality is quite simple. Once a society begins to rely upon system integration in order to meet certain core organizational challenges, e.g. production and distribution, goal-setting, etc., these systems develop an internal dynamic that extends the system ever further into socially integrated domains of interaction. This analysis involves three key concepts: instrumentalization, mediatization and colonization. **Instrumentalization** is the process through which spheres of action in which agents traditionally acted communicatively are switched over to instrumental action. **Mediatization** occurs where interactions that were once coordinated through language are switched over to steering media (2:183). The difference here is subtle: instrumentalization is an action-level concept, dealing with how agents decide what to do; mediatization is an order-level concept, dealing with how the actions of a number of agents are coordinated. Habermas understands instrumentalization and mediatization to be complementary phenomena.

The uncoupling of system integration and social integration . . . presupposes a differentiation on the plane of interaction not only between action oriented to success and to mutual understanding, but between the corresponding mechanisms of action coordination [i.e. language or media]. (2:180)

Naturally, it is not possible for every sphere of social action to be instrumentalized/mediatized. Since social order cannot be achieved in a purely instrumental context, certain types of interactions must remain norm-governed, i.e. socially integrated, in order to maintain the institutions that guide and constrain the system. Habermas adopts a classic Parsons schema in which the prerequisites of social integration are analyzed under the categories of culture, society, and personality. In this view, social order is maintained through a combination of socialization and social control. The **colonization** of the lifeworld occurs therefore when spheres of action that are essential to the operation of either of these two mechanisms are switched over to instrumental action (2:196).

Given the choice, people would not mediatize spheres of interaction that are essential to the reproduction of social order. The problem, according to Habermas, is that in order to keep the system running smoothly, it is often necessary to extend the scope of media-steered interactions. For instance, corporations are constantly trying to externalize costs, often in ways that are not easily accounted for in monetary terms, e.g. pollution, layoffs, exploitation of public resources, etc. In order to avoid socially undesirable consequences, the government must introduce new regulations, 'costing' these outcomes in such a way as to restructure corporate incentives. But given that instrumentally rational actors have a constant incentive to find a way around any set of rules, there will be a constant increase in both the density and scope of regulation, and hence a constant expansion of media-steered interactions.

So even though instrumentalization and mediatization are, as such, harmless, they set into motion a dynamic that gradually increases the sphere of system integration.10 When this sphere expands to the point where it interferes with the maintenance and reproduction of the lifeworld, mediatization becomes colonization (2:318), and social pathologies develop. It is these pathologies that Habermas considers it the task of a critical theory to diagnose.

Before proceeding, one reservation is in order. As it stands, this analysis presupposes an objectionable form of functionalism. Habermas treats integration as if it were an automatic consequence of all social interaction, and therefore simply equates instrumentalization with mediatization. Thus he presents social integration and system integration as an exhaustive dichotomy, mapping directly on to communicative and instrumental action. The problem is that while actions must be either communicative or instrumental, there is no reason that an interaction has to be integrated. It could simply be anomie. For instance, in the Northwestern course registration system, there is no reason that the procedure has to be carried out through **either** an orderly line-up, or a competitive auction. If the queuing norm were to break down, students might just mob the registration booth in an instrumental free-for-all. This means that instrumentalization, as the process through which communicative action is switched over to instrumental, should be accorded conceptual primacy. The switching process, when left to its own devices, leads directly to anomie.11 Mediatization is one possible solution to this problem, that may or may not be implemented.

It is worth emphasizing that this criticism does not presuppose or imply any sort of general objection to functionalist explanations. I am therefore not challenging Habermas's claim that intentionalist and functionalist explanations (corresponding to lifeworld and system perspectives) are complementary modes of analysis. My concern is
with Habermas’s implied claim that instrumental interaction always becomes integrated through steering media. My objection is not that he leaves unspecified the mechanism through which this integration is achieved, it is that there can be no such mechanism, because the claim itself is false. The question of functionalism arises because the only semiplausible reason for thinking that it might be true would stem from some broadly organicist view that regards integration as an inherent telos of all social interaction (such as can be found in the work of Niklas Luhmann or the later Parsons). 12

This correction induces a few minor changes in the overall analysis. First of all, it becomes incorrect to say that the mediatization of the lifeworld becomes colonization when it encroaches upon the core regions of lifeworld reproduction. Mediatization may provide the dynamic that extends the system, but it is actually instrumentalization that becomes colonization, since it is instrumental action that blocks the reproduction of social order. Habermas’s formulations are often misleading on this point. Consider the following:

If it is true that the symbolic reproduction of the lifeworld cannot be transposed onto the base of systemic integration without pathological consequences, and if precisely this trend is the unavoidable side effect of a successful welfare-state program, then in the areas of cultural reproduction, social integration, and socialization an assimilation to formally organized domains of action would have to take place under the conditions . . . [etc.]. (2:356–7; my emphasis).

This makes it appear as if it is the form of integration that blocks the reproduction of the lifeworld, rather than the type of action. In fact, the symbolic reproduction of the lifeworld cannot be transposed onto the base of instrumental action, regardless of whether this action is organized or disorganized. But this means that mediatization is involved in colonization only indirectly, insofar as it provides the mechanism that extends instrumentalization. 13

The second induced change is somewhat more significant. Since Habermas assumes that instrumental interactions will automatically become integrated through media, he paints rather too rosy a picture of the impact of instrumentalization. He concludes that it is only when mediatization extends to the point where it induces colonization that anomie will arise (2:385). But without the dubious functionalist assumption, one does not have to look so far to find problems. Anomie can arise as a direct result of instrumentalization. While colonization may place objective limits on the capacity of agents to correct this through media, there is nothing to guarantee that the media solution will be implemented in the first place. Thus the real problem with modern societies is not the ‘uncoupling of system integration and social integration’, but rather the shift from communicative to instrumental action. This is what creates anomie, and this is what undermines the reproduction of the lifeworld. Whether instrumental interactions are reintegrated through systemic mechanisms is a separate issue, and is relevant only insofar as it determines whether there will be a tendency toward increased, decreased, or stable levels of instrumentalization.

It is therefore a mistake to want to diagnose lifeworld pathologies by analyzing the effects of steering media. The task is rather to analyze the effects of switching interactions from a communicative to an instrumental orientation. As a result, the significance of the colonization thesis is dramatically reduced. In effect, the colonization thesis provides an in principle limit on the scope of systemic integration. This is a good argument to have around as backup, but it is not really required until we are given some reason to think that many of the various problems generated by instrumentalization are even susceptible to media treatment. And barring naive functionalism, this is not at all obvious. Looking at debates among economists over even relatively straightforward problems like pollution control, it is obvious that there are often serious problems (both practical and conceptual) with the implementation of media solutions.

At this point, the role of rational choice theory as a tool for diagnostic critique becomes clear. If instrumentalization means that agents adopt an instrumental orientation, then the basic postulates of rational choice theory should be satisfied in these domains, i.e. it should be possible to model these interactions game-theoretically. One of the most interesting aspects of game theory is that its models suggest that interactions among instrumentally rational agents will often produce characteristically counterintuitive results. In particular, agents will often experience difficulties in coordinating their actions in order to bring about mutually beneficial outcomes. Now, if some of the typical coordination and organization problems that game theory models predict begin to show up in our actual patterns of social interaction, then we would have excellent heuristic grounds for attributing these problems to an over-reliance upon the instrumental action orientation.

IV

In this way, rational choice theory would not provide a general theory of social action, it would serve as a diagnostic tool that could help us
identify the problems generated by instrumentalization. Rational choice models would be used to generate specific critical theories in the form of testable social-scientific hypotheses. The lifeworld pathologies that Habermas identifies as the product of colonization would persist as a subset of these problems. But a rational choice framework would also allow the critical theorist to identify problems that arise when socially integrated domains of interaction are instrumentalized, but not reintegrated through systemic mechanisms. Although there are many such cases, the following three types strike me as particularly interesting.

1 Collective action problems

When agents adopt an instrumental orientation, it can no longer be assumed that they will act in order to bring about outcomes that are in their common interest. Games in which the common interest does not translate into a set of coincident individual interests are called prisoner’s dilemmas, or market failures. Consider the following example: during a drought in California, if everyone waters their lawn, then the reservoir will run dry. The city calls on citizens to refrain voluntarily from watering their lawns. However, each citizen knows that the amount of water his or her lawn needs is negligible, and so realizes that regardless of what others do, he or she is better off watering. If others refrain from watering, he or she can ‘free ride’ on their conservation efforts; if others do not refrain, then the reservoir will run dry anyway, and so it is best to get as much water as possible, while one can. Since everyone reasons in the same way, everyone winds up without water the next day.

This is a classic example not only of collectively self-defeating behavior, but of a specific type of collectively self-defeating behavior that is characteristic of interactions among instrumentally rational agents (in the same way that a ‘feud’ is characteristic of certain systems of conventional norm-governed action). In this example, the water supply is what is known as a ‘collective good’. A collective good can be defined as an outcome that is better for everyone, yet unobtainable because it is not a strategic equilibrium. Other examples of collective goods include: a clean environment, high wages, nuclear disarmament, an education system, major economic infrastructure, long-term savings, resource conservation and full employment.

If, as Habermas suggests, there is a tendency toward instrumentalization of the lifeworld, it means that collective action problems will begin to show up where none existed before. Handgun ownership in the United States is a good example. The problem is a classic prisoner’s dilemma, i.e. it can never hurt to be better armed than your neighbor, but a society in which everyone has a gun is much worse than one in which no one does. Since this basic incentive structure has always been in place, the recent escalation in gun ownership in the United States can be attributed to the development of an increasingly instrumental attitude among citizens, which is to say, an erosion of the moral commitments that traditionally restrained them.

Once the normative contexts that originally allowed agents to avoid these collective action problems break down, it becomes necessary to establish a set of formal sanctions in order to motivate compliance for group projects. In this way, the rational choice analysis could be used to supply micro-foundations for the account of juridification. On a more general level, the identification of hidden collective action problems can often provide the clue to understanding otherwise mysterious processes of social change. For instance, the erosion of professional standards in print journalism, or of quality programming in broadcast television, is arguably not the consequence to any particular cultural changes, but is a self-reinforcing dynamic generated through the aggregation of instrumental choices.

2 Problems of legislative deliberation

Social choice models of democratic decision-making have shown that when agents vote their preferences in an instrumentally rational fashion, it results in very specific types of legislative instability and irrationality. The most common problem arises from the effects of ‘cyclical majorities’. Consider the following example: individuals A, B and C must vote on a proposal to divide $30 among themselves, with amendments voted on as they are introduced. Suppose that the initial proposal is to divide the money equally, so that each gets $10. If A introduces an amendment that modifies this distribution to give himself $19, B $11 and C nothing, it can be expected to pass with A and B in the majority. But then B can suggest, e.g. giving herself $29 and C $1, and expect this to pass with B and C in the majority, and so on. It can easily be shown that there is no proposal that cannot be defeated in some similar way. Not only can the procedure go on forever, but there is no outcome that is not obtainable from some point. The consequence is that whatever distribution is eventually settled upon will be either arbitrary or an artifact of the procedure through which it was obtained.

The point is that there is often no effective method for aggregating individual preferences; all of the ‘I wants’ simply do not add up to a coherent ‘we want’. Strictly speaking, this is true only when the
Rational choice as critical theory

collective preference ordering is intransitive. However, as the number of options increases, the chances of intransitivity increase as well, so that, for example, in the case of budgeting, the existence of intransitivity (and hence the potential for cyclical majorities) is virtually guaranteed. One of the consequences of this problem is that it places an enormous amount of power in the hands of those who set the agenda. In the case of simple intransitivities, it can be shown that whoever determines the sequence in which options are considered is effectively able to decide on the outcome.

What this shows is that the entire apparatus of legislative democracy is pointless unless there is a very widespread commitment to a deliberative self-understanding. If legislators maintain a generally communicative orientation, and are thus motivated by the force of the better argument, then in principle there will be no problem settling upon a legislative outcome. But once legislators begin to interpret their role as that of pursuing certain individual or community objectives in an instrumental fashion, the procedure becomes arbitrary, and the system increasingly ineffectual.

This analysis provides a powerful critical tool, particularly in the United States, where the lack of an effective party system, along with the idea of institutional 'checks and balances', encourages an instrumental orientation among legislators. It suggests that the range of phenomena referred to under the rubric of legislative 'gridlock' may not be the result of partisanship, but may be an endogenously generated consequence of a broader instrumentalization of the public political sphere.

3 Problems with the market economy

With the realization that the rate of profit does not tend to fall, and that labor is not the source of all value, the traditional Marxist critique of capitalism (which Habermas still subscribed to in Legitimation Crisis) is in need of reconstruction. Specifically, what needs to be shown is that transferring production and exchange over to a market mechanism will result in income and wealth distributions that generate problems of justification. Naturally, the specification of what will count as problematic is not obvious, since it is generally acknowledged that a market enables much higher aggregate production levels than any known alternative. Some formula along the lines of Rawls's difference principle is needed to specify what will represent acceptable tradeoffs between aggregate and distributive gains.

What rational choice theory can provide is a general model to specify the distributive effects of the market mechanism. John Roemer, for instance, has used standard microeconomic models to reconstruct the basic Marxist complaint, viz. that, granted a few basic assumptions, using markets to organize production will generate exploitation (some people work harder), inequality (some people earn more) and classes (some people work for others). With this sort of analysis, the distributive consequences of relying on instrumental action in the economic sphere can be precisely specified. As an added bonus, by focusing on instrumentalization rather than mediatization, the analysis separates out a number of distinct issues that tend to get run together in classical Marxism under the heading of 'commodification'.

A rational choice model of this type, because of its articulation at a microeconomic level, would be of considerably greater value in the everyday critical analysis of economic policy. Because of certain peculiarities of the classical economic framework, pre-analytic Marxism was necessarily restricted in its analysis to a macroeconomic level. This meant that other than providing a sweeping indictment of capitalism as a whole, it could not provide much in the way of specific policy recommendations. A rational choice model, on the other hand, could be used to specify the distributive effects of, for example, transferring a single industry or market sector from private to public ownership. This sort of fine-grained analysis is becoming increasingly indispensable, as critical theorists begin to face up to the necessity of balancing distributive against aggregate gains.

Note that the types of problems canvassed here all involve unpleasant consequences that arise as a direct result of instrumentalizing socially integrated domains of interaction. Since the segments of the lifeworld subjected to this process may or may not be involved in producing resources required for the reproduction of the lifeworld, none of these problems is necessarily connected to the colonization of the lifeworld. This means that a rational choice analysis of this type complements, rather than replaces, Habermas's analysis of lifeworld pathologies.

To a certain extent, this simply involves a change in emphasis. Habermas's colonization thesis, which is designed to guarantee a limit on the scope of system integration, is clearly motivated by his desire to respond to the pessimistic vision of a 'totally-refined society' that haunted members of the early Frankfurt School. But at the present historical juncture, and perhaps especially in a North American context, it is rather the lack of effective social organization that seems to provide a more tangible threat. Otherwise put, the Hobbesian state of nature appears to be a more proximate dystopia than the fascist state.
Rational choice as critical theory

V

When Habermas's analysis is formulated in this way, some question may arise as to the sense, if any, in which it remains a form of critical theory. In order to address this issue, I would like to show that the analysis continues to integrate 'explanatory-diagnostic' with 'anticipatory-utopian' moments in a way that is characteristic of critical social theory. This is the crucial feature, since, as Seyla Benhabib puts it, 'without an explanatory dimension, critical social theory dissolves into mere normative philosophy; if it excludes the dimension of anticipatory-utopian critique, however, it cannot be distinguished from other mainstream social theories that attempt to gain value free knowledge of the social world'.

In the formulation of Habermas's analysis presented here, these two components are preserved, because rational choice theory serves only an explanatory-diagnostic role, while communicative rationality continues to provide the normative standards in terms of which forms of social organization are to be evaluated. Because Habermas does not attempt to reduce communicative to instrumental action, or vice versa, he is able to preserve these as distinct components of the theory. This distinguishes his theory from that of rational-choice liberals like William Riker, or analytical Marxists like Roemer, who treat all action as instrumental, and so are constrained in their ability to introduce normative standards. But it also sets his analysis apart from that of, say, Charles Taylor, who treats instrumental action as just one instance of normatively governed action.

Consider first the case of those who treat all action as instrumental. Riker, for instance, in Liberalism Against Populism, correctly interprets the limited results of social choice theory to indicate that decision-making among instrumentally rational agents will be arbitrary and unstable. However, because he regards instrumental action as the only rational action type, he concludes that legislative decision-making must be strictly devoid of normative content. He then draws the anti-democratic conclusion that the normative content of liberal regimes lies, not in their collective decision-making practices, but in their enforcement of individual rights. Not only is this conclusion an embarrassment for the moral self-understanding of democratic societies, it is also question-begging, insofar as Riker does not give a satisfactory account of how respect for individual rights is secured among instrumentally rational agents.

Roemer puts himself in a similarly peculiar situation. By replacing the praxiological conception of labor that was central to the Marxist tradition with the strictly instrumental version that dominates contemporary microeconomics, he is able to establish a genuinely neutral framework for the analysis of labor and capital markets. But because he views all action as instrumental, when he moves to introduce the key normative category of Marxian analysis—exploitation—he has only limited resources at hand. He opts to define exploitation using a cooperative bargaining model with social classes as coalitions. However, he is unable to get the right results by taking the social status quo as the initial bargaining point, and so is forced to introduce an idealized redistribution of resources as the basis of the bargaining problem. This is question-begging on a number of levels, not the least of which is that cooperative game theory itself lacks instrumental microfoundations.

Both Riker and Roemer get into difficulties because they attempt to use rational choice theory, not just to diagnose problems that emerge from actual patterns of social interaction, but also to construct the normative standards against which these patterns can be evaluated. Within Habermas's framework, the normative standards are supplied by the conception of communicative rationality. This means that he can agree with Riker's critique of instrumental populism, without drawing any anti-democratic conclusions. Rational choice theory does not prove the futility of legislative decision-making, it gives us tools to identify the consequences of an erosion of deliberative politics. In the same way, the analysis of collective action problems is not intended to supply normative foundations for the state, but rather to identify undesirable social patterns that can be corrected through state action.

It is equally important, however, that one does not go too far in the opposite direction, and allow the normative component of the theory to obscure the explanatory-diagnostic dimension. Taylor, for instance, while denying that all action is instrumental, presents a model of rational action that effectively treats all action as organized around a set of moral commitments. In his view, a variety of strongly evaluated goods underlie all social practices. What we call instrumental action, in this view, is actually action oriented toward a good called 'efficacy'. This good then competes on all fours with a variety of other goods that are important for the modern identity, such as 'equality', 'creativity', etc. The instrumentalization of social interaction is interpreted as simply a reordering of our moral commitments in such a way as to rank efficacy' highest.

The problem with this view is that it encourages a simplistic understanding of the problems of modern societies. For instance, Taylor treats a variety of perverse and undesirable effects generated by the market mechanism as a straightforward consequence of our
excessive commitment to the good of efficacy. The implication is that it is within our conscious control to prevent these outcomes by simply reordering our moral priorities. Rational choice analysis suggests that it is not the way that individuals rank their values that creates the problem, but rather the way in which their actions interact when they attempt to realize these values. The nature of this interaction is often obscured from the actors themselves, so that it is impossible for agents to resolve the problem simply by reorganizing their priorities. Most obviously, there is good reason to believe that a number of unpleasant side-effects of markets, such as recessions and classes, will continue to be generated regardless of the value system we adopt.  

In short, it is important to maintain a clear distinction between perverse effects that are generated by having the wrong norms institutionalized, and perverse effects that arise as a consequence of having nothing institutionalized. By treating all social action as norm-governed, Taylor creates an unhelpful confusion of the two. Habermas’s typology of rational action allows us to maintain this distinction. Within this framework, rational choice models allow us to determine when social problems are of the latter type.

VI

In summary, the argument advanced here suggests that Habermas develops his analysis using the order-level concept of system integration and ‘functional reason’ when he should use the action-level concept of instrumental rationality. The attempt to base a critical social theory on order-level concepts is misguided, simply because in order to get social order you already have to have solved most of your problems. Switching to the action-level concept means that we should be using rational choice theory to develop specific diagnostic critical theories. Naturally, this does not replace the analysis of lifeworld pathologies. Insofar as instrumentalization does erode the cultural, social and personal underpinnings of the system of steering institutions, Habermas’s basic analysis will continue to hold. What the rational choice analysis does is bring to light an entire region of suboptimality and social dysfunction that is obscured in Habermas’s treatment. Thus the action-level analysis does not undermine the order-level analysis, it merely broadens it in such a way as to expose difficulties that arise prior to the achievement of order.

This reformulation of the diagnostic component of Habermas’s critical theory has a number of advantages, which I will conclude by summarizing briefly.

- The rational choice framework allows us to drop the objectionable form of functionalism that underlies Habermas’s analysis, and gives us access to a number of very powerful theoretical choice models that, unlike Luhmann’s systems functionality, are both rigorously developed and empirically verifiable.
- The rational choice analysis has much greater generality than the systems analysis, precisely because it does not assume that the actions of instrumentally oriented agents will be integrated. Habermas ties his analysis of lifeworld pathology directly to mediatisation, and mediatisation directly to instrumentalization. But once we grant that this second connection is not a direct one, given that mediatisation is only one possible response to instrumentalization, it is obvious that Habermas has implicitly restricted his analysis to a very narrow range of phenomena. The rational choice approach eliminates this circuitous route by taking up the analysis of instrumentalization directly.
- The use of rational choice theory allows us to develop small-scale, concrete analyses of interaction patterns, using tools that do not require extensive commitment to one or another theoretical system. This relative theory-neutrality creates the possibility for fruitful dialogue between critical theory, analytical Marxism and mainstream economics. And in the context of a society in which the critical theorist is not a cultural hero, establishing these sorts of connections is almost the only way of exercising any influence over actual public policy discussions.

University of Toronto, Canada

Notes

I would like to thank James Johnson for persuading me of the correctness of many of these views, Thomas McCarthy for very detailed and helpful comments, participants in the 1994 Midwest Critical Theory Roundtable for discussion of an earlier version of this paper, along with the Social Sciences and Humanities Research Council of Canada for financial support during the period in which it was written.

Rational choice as critical theory

2 See Russell Hardin, Collective Action (Baltimore, MD: Johns Hopkins University Press, 1982).
3 For an extended discussion of the action-order distinction, see Jeffrey Alexander, Theoretical Logic in Sociology, 4 vols (Berkeley: University of California Press, 1983 [1982]).
4 It should be noted that Habermas very confusingly designates communicative action as a 'mode of action coordination', the idea being that agents act communicatively when they harmonize their action orientations through linguistically mediated exchanges. This definition blocks the action-order distinction that I am drawing, unhelpfully building functionalist commitments into the very concept of communicative action. But since no one would claim that instrumental action is a 'mode of coordination', an action-order distinction must be available between instrumental action and system integration. This means that a pure action-level concept of communicative action can simply be defined in contrast to instrumental action, which can then be distinguished from the order-level concept of social integration. Naturally, it is always then open to Habermas to argue, as Parsons did, that social integration will be the equilibrium state of communicative interactions. The point is that this cannot be established through mere stipulation.

5 For some reason, the fact that Habermas adopts a 'rational choice' (i.e. preference-based decision and game theory) account of instrumental rationality has escaped the attention of numerous commentators. See, for example, Maeve Cooke, Language and Reason (Cambridge, MA: MIT Press, 1994), pp. 20–2. This is surprising, not only because Habermas explicitly refers to John von Neumann and Oskar Morgenstern in the Theory of Communicative Action, but also because he adopts the term 'strategic action' in the same technical sense in which it is used by game theorists.


8 This suggests an interesting line of inquiry (which Habermas does not pursue). Instrumental reasoning can be very demanding, both in terms of the amount of information that agents require and in the computational skills they must exercise. For instance, strategic interactions require that each player's preference ordering be common knowledge among all players. This is a condition that is not met under most circumstances. However, when preference is translated into, e.g. demand using the money medium, or votes using a power mechanism, it becomes both public and measurable. Media also allow certain strategic variables in the environment to be fixed as parameters, e.g. by guaranteeing binding decisions. This dramatically simplifies a variety of choice problems.

10 The basic idea for such a dynamic comes from Karl Polanyi, The Great Transformation (Boston, MA: Beacon Press, 1944).
11 Anomie understood in Parsons's sense as 'an absence of structured complementarity of the interaction process'; see The Social System, p. 39.
12 Note that although Habermas has himself criticized this form of functionalism, he has done so on the grounds that it ignores the distinctive characteristics of lifeworld patterns of integration, not on the grounds that it assumes that integration is an automatic consequence of interaction. However, one might choose to interpret Habermas charitably, and suggest that his remarks are simply misleading, and that he does not believe instrumentally rational action is automatically integrated through media. In this case, the position I am outlining here can be regarded as simply a more detailed articulation of Habermas's actual view.

13 This having been said, it is not even obvious that steering media alone supply the dynamic through which instrumentalization is extended. Alexis de Tocqueville's observation that self-interest is the 'rust of societies' reflects the fact that instrumental action is, in many respects, self-propagating. This is because the trust that underlies cooperative relationships is usually built up over time, but can easily be undermined by a single defection. Agents who have been exploited through un reciprocated adherence to cooperative schemes may be more likely to adopt instrumental orientations in the future.

16 Jack Knight and James Johnson suggest that proponents of deliberative conceptions of democracy (like Habermas) would find considerable support for their views in the rational choice literature. But
unfortunately 'they seem rarely to read it, except superficially as an example of what they take to be anemic apology for liberalism', in 'Aggregation and Deliberation: On the Possibility of Democratic Legitimacy', Political Theory 22 (1994): 277–96.


21 David Gauthier has run into similar problems in his attempt to use bargaining theory to establish a basis for fair cooperation in prisoners' dilemmas. For criticisms along this line, see Ken Binmore, 'Bargaining and Morality', in David Gauthier and Robert Sugden (eds) Rationality, Justice and the Social Contract (Hemel Hempstead, Herts: Harvester Wheatsheaf, 1993).

