
CHAPTER 1

An Outline of Habermas's Critical Theory

Jürgen Habermas's critical theory of society is a diagnostic theory of societal rationalization. A theory of societal rationalization is a theory that explains the evolution of modern society out of premodern traditional societies as a process of rationalization. Societal rationalization is generally understood either as a process in which institutions increasingly tend to secure the acceptance of society by appeal to rationally justifiable principles rather than to tradition, or as a process in which society becomes increasingly capable of appropriating nature to meet its expanding needs and interests. The former characterization is often associated with Max Weber, the prominent nineteenth century German sociologist who explained the rise of capitalism as a product of the Protestant ethic. The latter characterization is associated with Karl Marx. Both Weber and Marx, along with the Frankfurt School theorists Theodor Adorno and Max Horkheimer, and Habermas himself owe much to the precedent established by Georg Wilhelm Friedrich Hegel's attempt at reconstructing the history and structure of consciousness in modern society.

As a *critical* theory, Habermas's theory of societal rationalization is intended not only to explain the process by which modern capitalist society evolved, but also to reveal the nature and causes of its systematic failures. Unlike (ideal) psychoanalytic theories and Habermas's own earlier work (up through *Knowledge and Human Interests*), his recent work is not aimed at finding a way out of the problems confronting modern society by providing either an explanation or a critique of ideology. Its

explanation of societal evolution, then, could be characterized as a *diagnostic* explanation, as opposed to a *therapeutic* explanation.¹ A therapeutic explanation is a therapy—a method of treatment—that dissolves false consciousness and *thereby* emancipates the subject. Two clear instances of this kind of explanation are the Marxian analysis of capital as reified labor, which was intended to bring about the communist revolution by removing the ideological blindfold from the worker, and classic Freudian psychotherapeutic explanations, which are supposed to ‘dissolve’ neurosis by bringing repressed memories into the conscious mind. Therapeutic explanations rest on, but are not confined to, diagnostic explanations. A diagnostic explanation explains the causes of a crisis, and thus opens the way to confronting it, but it is not itself the means of emancipation. Nor does a diagnostic explanation or critique take a *substantive* evaluative stance with regard to the object of critique, although the questions raised within the analysis are determined by ideals that are taken to be fundamental (such as health in the case of medical diagnosis, and mutual understanding in the case of Habermas's theory). In other words, a diagnostic critique is not the same as a *moral* or *cultural* critique. In *The Theory of Communicative Action*, Habermas wished to construct a critique of societal rationalization that would reveal how the crises facing modern society may be explained as the result of a ‘one sidedness’ in the rationalization of Western societies. Although this critique ultimately is premised upon a commitment to the ideal of mutual understanding, it does not yet contain a commitment to or defense of particular substantive moral or cultural values against others.² Habermas's sociological hypothesis, which will be explained in Chapter 5, is that the processes in which culture is formed are threatened by the administrative and economic systems characteristic of modern societies. This hypothesis takes no substantive evaluative stand with regard to the particular cultural values that are threatened.

It is a feature of a critical theory of society in general, whether therapeutic or diagnostic, that it attempts a “self-clarification of the struggles and wishes of the age,” in such a way as to guide those struggles.³ Toward this end, it should also explain how a person, group, or society has come to be engaged regular-

ly in practices that, in fact, are not in his, her, or its interest, as a result of some feature of that society. Another feature of critical theories of society is that they identify deep conflicts, or potentials for society-wide crises, inherent within the social, political, and economic institutions of modern capitalist societies. For example, Marx's theory of history attempts to explain the evolution of the institution of capitalism, an institution that is seen as conflicting with the interests of the working majority of early capitalist society. Because this conflict is ineradicable from capitalist societies in Marx's view, it would eventually lead to a major crisis—the revolution of the proletariat.

Often, Karl Marx is thought to be the first critical theorist, although not everyone agrees about what makes Marx's theory of history a critical theory. It is widely believed that the central critical moment of Marx's theory is the labor theory of value, which allows for the statement of the "contradictions" or manipulative illusions whose unmasking was to bring about the socialization of the means of production and a new kind of society.

Some scholars prefer to confine the use of the term *critical theorist* to the Frankfurt School theorists Theodor Adorno, Max Horkheimer, and Herbert Marcuse and their intellectual affiliates. Frankfurt School critical theory differs from Marx's theory of history principally in its evaluation of societal rationalization. For Marx, the evolving use of human rationality in improving our means of controlling nature contained an emancipatory potential, leading toward an efficient as well as liberated mode of production in the future communist society. The Frankfurt School, on the other hand, shared Weber's ambivalence toward the emancipatory potential of rationalization. In Weber's view, the more society becomes functionally organized and removed from tradition, the less freedom and meaningfulness are available to the individual.⁴ Similarly, the Frankfurt School was concerned that rationalization entailed an ultimately senseless proliferation of bureaucratic offices, each issuing in further constraints on the individual. This absurd pursuit of perfect control leads, in the end, only to barbarism. "Enlightenment destroys itself."⁵ The external constraints imposed by a rationalized society, then, are not its only regrettable feature; reason itself is inherently coercive, and its hold can be broken only by a flight into the irrational.⁶

Habermas does not share Adorno's pessimistic belief that a special form of barbarism is inherent to reason and modern society. As an aside, it is worthy of note that he also does not share the experiences of the early Frankfurt School theorists, who spent several years in exile in the United States during the Third Reich. Habermas, born in 1929, was a child when the Frankfurt School left Germany and later a youthful member of the Nazi Jugend (a children's organization concerned to instill German nationalism and loyalty to Nazi aims), learning of the Holocaust as a university student, after the surrender.

Habermas's conception of rationality—what he calls *communicative rationality*—more than anything else distinguishes his theory of societal evolution from the theories of Marx, Weber, and the Frankfurt School. Habermas rejects the equation (which he believes these individuals made) of rationality with what Weber called *Zweckrationalität* or “instrumental rationality.” *Instrumental rationality* is defined as the rationality that governs the choice of means to given, usually material, ends. *Communicative rationality*, on the other hand, characterizes the activity of reflecting upon our background assumptions about the world, bringing our basic norms to the fore, to be questioned and negotiated. Instrumental rationality takes these background assumptions for granted, in the pursuit of new gains (cf. *TCA* I, 285–286). Habermas believes that the notion of instrumental rationality is insufficient to capture either the nature of *cultural evolution*, which is not governed merely by instrumental thinking, or the nature of economic and administrative systems, whose organization is too complex to be characterized as the product of instrumental planning. Rather, cultural evolution is a process fed by collective reflection on the whole range of values to which the society is committed. Hence it is Habermas's understanding of the nature of culture, and not a misguided desire for a society without conflict or dissent, that motivates his positioning of the ideal of mutual understanding at the center of his critical theory. On the other hand, economic and administrative systems in modern welfare states are characterized by *functional rationality*, which holds of these systems independently of the intentions of the instrumentally acting agents within them. Thus, the functional rationality of a system is decided on the basis of its ability to

achieve goals such as economic and political stability, whereas the instrumental rationality of agents' actions is determined by their ability to maximize (those agents') utility. In summary, instrumental rationality explains neither the culture nor the behavior of the capitalist welfare state.

With Weber, Habermas rejects the Marxist determinist assumption that the instrumentally rational development of capitalist production determines not only the objective or institutional conditions, but also the subjective or ideological conditions, of societal change. Habermas believes that the social construction of social reality is part of the process of social change. Social reality is constructed through what Habermas calls *communicatively rational* action, or communication between participants attempting to reach a rational consensus. An understanding of societal evolution can be gained, in his view, by understanding how social reality is constructed through consensus building.

Communication, according to Habermas, is inherently oriented toward mutual understanding, and the standards that govern communication are therefore conditioned upon reaching mutual understanding and, ideally, rational consensus (*CES*, 3). Communicatively rational action, then, is action that conforms to these standards. There are three fundamental types of valid rational consensus, corresponding to three basic ways in which a communicative act can be claimed to be valid. In asserting a statement, one implicitly claims that the sentence stated is *true*. In stating a prescriptive norm, one claims that the norm is *normatively valid*. Finally, in expressing a subjective state, one is implicitly claiming to be *sincere* or truthful. Habermas calls truth, normative validity and sincerity the *validity claims*.⁷ The validity claims are made in everyday acts of speaking. By asserting something, I make the claim that an assertion is true; in issuing a demand on moral grounds, I make the claim that a norm is justifiable, and in expressing my subjective point of view, that my expression is authentic. These validity claims are made in everyday acts of speaking. To be capable of communicatively rational action is to be well versed in the use and defense of validity claims. Communicative rationality, then, or the rationality of understanding oriented action, is an ability to produce rationally motivating justifications for validity claims.

The ability to raise validity claims, Habermas believes, is the basis of the social bond.⁸ The act of raising a validity claim affirms the mutual commitment to standards of validity that make communication, and thus the social relationship possible. This mutual commitment enables one to recognize one's own and the other's respective places in an intersubjectively shared world; it enables each of us to agree on "where we stand" with respect to each other and with respect to an objective world of "states of affairs," an understanding of which we share. Now, it is the claim that the validity standards are the basis of the social bond, more than any other, which establishes the significance of Habermas's theory for sociology and for any theory of the evolution of social institutions. If the validity standards are the shared basis on which we negotiate our evolving social institutions and arrangements, as well as our shared understanding of the objective world, then they must be part of—indeed central to—the explanation of societal change.

Although Habermas rejects the Marxian and Weberian hypothesis that instrumental or "strategic" rationality is *central* to societal evolution, he does not deny that strategic rationality regulates individuals' interactions in the economy and in political decision making. Strategically rational action is constrained on two fronts: on the one hand by the values underlying communicative rationality (theoretical, ethical, cultural, and aesthetic values), and on the other hand by the functional, systemic requirements of the economic and administrative systems. This reevaluation of the place of strategic and functional rationality in the explanation of societal evolution is the chief distinguishing feature of Habermas's theory of societal rationalization, as opposed to other theories of societal rationalization and the more recent systems theories. For Habermas, it is not the degree to which a society has harnessed nature to produce goods that determines its degree of rationalization, but the degree to which the use of the validity claims have been developed in that society's communicative practices.

It may be unclear how a theory intended to describe the evolution of human rationality and the development of social institutions that results, could also be a critical theory. Furthermore, if there could be such a theory, it would seem to violate the basic scientific

requirement of value freedom. A theory cannot both contain normative content—that is, be critical of society—and be a scientific explanation. In response to the first concern, the explanatory-descriptive and normative-critical aspects of Habermas's critical theory both are derived straightforwardly from the theory of communicative rationality. Any model of human rationality is a (descriptive) model of a normative structure. The hypothesis of a theory of societal rationalization, then, is that the norms of rationality, as described by the theory of rationality, are operative in history. That is, they figure centrally in the history of our institutions and practices, by setting the standards for the choices we make. But the same norms provide the foundation for a critique of society.

The second concern, a familiar one among American philosophers and social scientists, is trickier. Although Habermas's theory of societal rationalization is intended to explain, among other things, how values take shape within a culture, rather than to take up the cause of particular values, it nonetheless is true that his theory is deeply critical of the modern capitalist welfare state. The normative force or validity of this critical content stands squarely on the validity of the norms of communicative rationality. If human rationality is based on the norms that Habermas believes it is based upon, then the critical content of his theory of society is strong, for it is framed by those norms. However, many of Habermas's critics believe that the necessary conditions of communication cannot provide the foundation for the very strong normative content that he draws from them.⁹ We shall return to this question in Chapters 3 and 4.

Habermas himself explains the intertwinement of explanatory and critical purposes in his theory in another way, which might be helpful here. According to Habermas's introduction to his critical theory of societal rationalization, the concept of communicative rationality is employed in three ways in its construction.¹⁰ First, it answers the metatheoretical question; In what ways *can* action be rational? The answer provides the conceptual framework for Habermas's model of the evolution of *culture*. Culture is modeled within this framework as evolving in three increasingly differentiated "spheres" that correspond to each of the three validity claims. Historically, the idea of culture dividing into these three spheres derives from Kant's three faculties of

reason: the theoretical sphere of science, the practical sphere of morality and law, and the aesthetic sphere. The distinctions drawn among these spheres reflect Weber's belief that each of these three cultural arenas possesses its own internal logic of development (*TCA I*, 159–164). According to Habermas, developments in each of these spheres are governed by the standards of validity (truth, normative validity, and expressive validity).

Second, the concept of communicative rationality is applied in the interpretation of culture: to grasp the cultural learning processes that lead to modern culture, the theorist is obliged, in effect, to engage in consensus-oriented communication with past and present cultures. He or she can understand them only by acknowledging and challenging their claims to validity (*TCA I*, 115–116). At this second level of involvement of the theory of communicative rationality in the theory of society, a critical dimension is introduced. The theorist's participation in a "dialogue" with the evolving culture is a two-way exchange, both listening to its lived crises and challenging it to identify and reflect on its normative and theoretical assumptions.

A third way in which the concept of communicative rationality enters into Habermas's theory is as the *object* of observation. The evolution of communicative rationality in society is observed, and hypotheses about the patterns in its emergence formulated. This level of involvement of the concept of communicative rationality is also central to the critical purpose of Habermas's theory. It is the pathological tendencies or distortions of the ideal of communicative rationality, in the actual historical unfolding of communicative rationality, that his theory is intended to diagnose.

Habermas's theory of societal rationalization is not a complete history of society: it is an analysis of history and exists only in the form of a rough outline of the processes that have governed the emergence of modern society. In a completed theory of societal rationalization, hypotheses about the course of development of communicative rationality would take a more specific form, similar to that of Jean Piaget's model of cognitive development, Lawrence Kohlberg's model of moral development, and George Herbert Mead's reconstruction of the evolution of communication. These models share the feature of being empirical recon-

structive hypotheses—they attempt to follow, empirically, the development of abilities with highly complex internal structures. Habermas's own work thus far is more of a rational or analytic reconstruction of the necessary components of communicative ability. However, he has made forays in the direction of integrating this reconstruction with the theories of Piaget, Kohlberg, and Mead. Because the evolution of communicative ability is part of cultural development as well as a development experience by each individual, however, a full reconstruction of the history of communicative ability would recount the stages of learning of societies as a whole, as they gradually came to distinguish between claims of truth and right, right and beauty, and beauty and truth.

Thus far I have been concerned with introducing the idea of a critical theory of societal rationalization, the theory of communicative rationality on which this theory rests, and the way in which the theory of rationality is marshaled to construct the theory of societal rationalization. However, I have not yet introduced the actual critical hypothesis about modern society defended in this theory. Habermas believes that the modern capitalist welfare state is undergoing a process of "one-sided rationalization." To explain this hypothesis, it is necessary to return to the distinction among strategic, functional, and communicative rationality. As was remarked earlier, Habermas does not completely reject the view that strategic and functional rationality are part of the explanation of societal evolution, but he does believe that they are derivative of communicative rationality, both conceptually and historically. Only at a relatively recent stage in our history do we find that functional and communicative rationality begin to "uncouple," so that some activities are regulated by functional rationality and others by communicative rationality. According to Habermas, modern society is characterized by a rift between two "worlds," each with its own basis for thought, action, and organization. The one world, called the *lifeworld*, is still based upon the norms of communicative rationality, which have consensus as their aim. The other, called the *system*, is based on the demands of material production (as in Marx's theory of history). Habermas hypothesizes that the system increasingly tends to interfere with and distort the communicative activity in which the pursuit of knowledge and ethical understanding takes place.

The consequences of this bifurcation of rational action can be described from two perspectives, corresponding, not surprisingly, to the division between the system and the lifeworld. From the global, unengaged "noncommunicative" perspective of the system, the system's tendency to undermine the lifeworld results in system dysfunction, as the individuals entering into it are less and less prepared to accept the cultural presuppositions on which the system itself rests. For example, the system depends on individuals willing to sacrifice short-term gains for the sake of anticipated long-term gains. However, such willingness does not occur "naturally"; it is the product of an ethos. Unless this ethos prevails, fewer individuals will be prepared to make the kinds of sacrifices that our system demands.

From the engaged perspective of the lifeworld, the result of one-sided rationalization is a loss of autonomy and meaningful activity. This loss is real and not merely perceived; it cannot be answered with a "free" exercise of a will to give meaning to activity. As the system expands to meet the demands placed on it, it reorganizes or otherwise disrupts just those domains of activity in which freedom and meaning were to be founded. The critical claims of Habermas's theory, then, arise not only from the engaged perspective of the communicatively rational theorist in dialogue with the lifeworld, but also from the perspective of the observer of the functionally rational system in interaction with the lifeworld.

If Habermas's claims about the history of modern society are to hold up, then it is important that his critique of rationality be viable. For this reason, the following three chapters will be devoted to the arguments that he makes in support of it. Chapters 2 and 3 will describe Habermas's defense of his consensus theories of truth and normative correctness. The norms that govern argumentation concerning truth claims and argumentation concerning normative validity claims constitute the principal part of the norms that govern communication in general, or the norms of "universal pragmatics." Before turning to the theory of societal rationalization, the fourth chapter will explain the sincerity claim and Habermas's defense of his universal pragmatics as a whole.