

Sociology 250

November 9, 1999

Introduction to Microsociological Approaches 1. Macro and Micro

The sociological theories of Marx, Durkheim, Weber, and Parsons and the functional school are primarily large scale, macrosociological, and structural. These theories were developed in the latter half of the nineteenth and the early twentieth centuries in Europe (with Parsons adapting these theories and developing a similar model in the United States). These theories were developed by European social theorists who were attempting to understand the new social world of a modern, industrial, urban society. These classical theories established sociology as an academic discipline, their definition of the social world established the scope of sociological study, and their methodologies determined how sociology would be studied and applied.

In North America a different set of questions occupied late nineteenth and early twentieth century sociologists, and a different sociological tradition became established. North American writers were more concerned with understanding the bases of social action and interaction among individual members of society. Since it is these interactions that define the the social world, underlying

social structures, and create and maintain societies, sociologists need to understand these. These microsociological or interaction perspectives are of several main types. Symbolic interaction examines meaning, action, and interaction at the micro level, and was developed by United States sociologists George Herbert Mead and Herbert Blumer, with Erving Goffman, a Canadian, being one of its primary practitioners (Wallace and Wolf, Ch. 5). A related approach is that of ethnomethodology, originally developed by Harold Garfinkel. This is referred to as interpretive sociology or interpretive analysis, and is related to the phenomenological approach of the Austrian-American Alfred Schutz (Wallace and Wolf, Ch. 6). Contemporary sociologists have adapted and developed these ideas and have created a great variety of mixed interaction approaches.

Some of the ideas that led to these microsociological theories are examined in the following sections. A short summary of some of the different aspects of these two levels of sociology is provided first.

a. Structures or Action. The subject of macrosociological theory is the large-scale structures and features of society – social class, division of labour, power, forms of authority, rationalization, and broad historical developments. In contrast, the subject of microsociology is the individual interacting with other individuals, often in small group settings. Where the

settings are within larger scale structures, the micro focus is still on how individuals interpret the situation and interact with other individuals in these settings.

The classical approaches often began with what could be considered micro concepts, but use these to develop macrosociological theories. For example, Marx begins with a micro concept, the commodity, but derives this from a study of capitalism as a whole, and uses it to explain the structure of this system. The theories of Weber and Parsons are concerned with social action, and Parsons calls his theory an action theory, so it might be thought that these two writers span macro and micro. While the action approaches of Weber and Parsons could have led to an interaction perspective, neither author really develops such an approach. Weber is concerned with meaning, but does not devote much attention to defining this, and the bulk of his writings is devoted to groups, organizations, history, and structures of power. Parsons begins with the individual act in a situation, and considers interaction, but moves quickly from this to systems and needs. Interaction underlies these, but is not the primary focus of Parsons. The macro theories concentrate on average action and the regularities that are common to large numbers of social actors. The micro focus is on individual action, its meaning, how interaction occurs, and the uniqueness of individuals and the self

b. Determined or Creative. The macro approach to social action tends to be determined by large scale structural features of society or the cultural and value systems. In these theories, there is often little room for, or analysis of, the manner in which people are creative in their interaction with others. That is, action is determined by social norms, cultural values, laws, religion, social class, consciousness, and ideology. The effect of these macro forces on social action is determinate and can be examined empirically in sociological studies (e.g., Marx's study of capitalism, Durkheim's analysis of the division of labour and suicide, Weber's study of rationalization). In contrast, the interaction approach considers humans to be creative, with unique selves and individual forms of interaction. For these sociologists, social action and interaction can only be studied by carefully examining the ways in which symbols, structures, and organizations are understood by individuals and how different individuals come to interpret interaction differently.

The development of the child, socialization, and the formation of self and individual identity form a large part of the interaction perspective. In this perspective, personality and identity are not determined biologically, but are developed actively within the social environment. While the symbolic interaction approach identifies symbols as important, these are not so determined as the values, norms, or consciousness of the classical theorists

and Parsons. Rather, sociologists working within the interaction perspective argue that the basis for social interaction is "a common set of symbols and understandings possessed by people in a group" (Wallace and Wolf, p. 191). These are developed through socialization and continual interaction with other individuals. The sociologist must also understand how the development of the self occurs as children and adults "interpret, evaluate, define, and map out their own action" (Wallace and Wolf, p. 191), rather than merely being "passive beings who are impinged upon by outside forces." (Wallace and Wolf, p, 191).

c. Decision or Practice. A third difference is the underlying approach to social action. The theories of Marx, Weber, Durkheim, and Parsons tend to look on action as resulting from a conscious and considered decision on the part of the individual actor. This approach to action may derive from the enlightenment view of individuals as rational decision makers, weighing alternatives, and deciding on the best course of action. Parsons is explicit concerning the rational, decision making process in his theory of social action, and the three classical theorists also appear to adopt such an approach, even if only implicitly. In contrast, interaction theories tend to focus more on actual activities and actions, and what people do in social situations. Whether or not such action is consciously considered and aimed at achieving a specific goal is not

the focus of the interaction approach. Rather, interaction theorists examine the experiences, practices, actions, and situations of people to see what they do, and then attempt to understand how these features occur. Everyday life, ordinary experiences, and asking how the generally accepted features of society emerge, are the focus of interaction studies.

While interaction theorists may not have an overall theory of society, classical and Parsonian theory has little analysis of social interaction. Both perspectives have been accused of ignoring women and gender issues, and having inadequate analyses of power in society. In order to understand contemporary sociology, it is necessary to study all of these sociological approaches.

2. Weber

Since Weber defined the social world as that of social action, with the aim of the sociologist being to develop an understanding of how individuals act, his theory is one forerunner of some interaction approaches. For Max Weber, each social action has meaning associated with it, in the sense that the individual does not act as a automaton or robot, or on the basis of instinct or stimulae. Some of the acts carried out by individuals are conditioned or automatic, but a large part of what individuals do is to consider a situation, think about how to approach the situation, contemplate the possible actions of others, and act in a way that the individual

thinks will best meet his or her goals. This may be entirely a consciously worked out process, but to be considered as social action, there must be some meaning associated with the action. The task of the sociologist is to attempt to see how people interpret and attribute meaning to the situation.

Consider the situation of workers in a job as an example. Marx assumed that the situation of workers was structurally determined to be in opposition to that of the employer. Yet workers may accept the authority and power structure of the employment situation, perhaps because they need to support themselves and their families, and wish to create a reasonably comfortable life for themselves. They may consider their situation and view acceptance of the organizational structure in which they work as their best option. That is, domination within an organizational structure may be viewed as legitimate, so that there is rational- legal authority. If the employment situation becomes intolerable, this may create more active struggle, perhaps with workers combining to form a trade union. But Weber would say though that the latter does not necessarily result. Rather, the sociologist must evaluate each situation through the eyes of the actors, in order to determine what meanings they take out of the situation, how they assess alternatives, and how they decide to act (or not act).

The interactionist approach can be connected to Weber, although contemporary symbolic interactionism has dramatically refined and developed Weberian and other concepts. Historically, more of the roots of the interaction perspective are in philosophical pragmatism, psychological behaviourism, the German writer Simmel, and some early United States social science approaches.

3. Simmel

While Simmel is generally not regarded as being as influential in sociology as were Marx, Weber, Durkheim, or even Parsons, several of the early United States sociologists studied with or were influenced by Simmel. This was especially true of those who developed the symbolic interaction approach including writers in the Chicago school, a tradition that dominated United States sociology in the early part of this century, before Parsons.

Georg Simmel (1858-1918, Germany) was born in Berlin and received his doctorate in 1881. He was of Jewish ancestry and was marginalized within the German academic system. Only in 1914 did Simmel obtain a regular academic appointment, and this appointment was in Strasbourg, far from Berlin. In spite of these problems, he wrote extensively on the nature of association, culture, social structure, the city, and the economy. His writings were read by Durkheim and Weber, and Simmel contributed greatly to sociology and

European intellectual life in the early part of this century. One of his most famous writings is "The Metropolis and Mental Life" (1903) and his best known book is *The Philosophy of Money* (1907). Simmel's ideas were very influential on the Marxist scholar Georg Lukacs (1885-1971) and Simmel's writings on the city and on money are now being used by contemporary sociologists.

Simmel was influenced by Hegel and Kant and developed a sociological analysis with ideas similar to the three major classical writers. When Simmel discusses social structures, the city, money, and modern society, his approach is similar to that of Durkheim (individual and society), Weber (rationalization), and Marx (alienation). Simmel considered society to be an association of free individuals, and argued that it could not be studied in the same way as the physical world, i.e. sociology is more than the discovery of natural laws that govern human interaction. "For Simmel, society is made up of the interactions between and among individuals, and the sociologist should study the patterns and forms of these associations, rather than quest after social laws" (Farganis, p. 133). He analyzed individual behaviour "because some crucial decisions are made on the individual level, among the 'atoms of society,' which can cause reverberations throughout an entire nation" (Wallace and Wolf, p. 193). Simmel's emphasis on social interaction at the individual and small group

level, with the study of these interactions being the primary task of sociology, makes Simmel's approach different from that of the classical writers, especially Marx and Durkheim.

It is Simmel's attempt to integrate analysis of individual action with the structural approach that make his writings of contemporary interest.

Simmel began his inquiries from the bottom up, observing the smallest of social interactions and attempting to see how larger-scale institutions emerged from them. In doing so, he often noticed phenomena that other theorists missed. For example, Simmel observed that the number of parties to an interaction can effect its nature. The interaction between two people, a *dyad*, will be very different from that which is possible in a three-party relationship, or *triad*. (Farganis, p. 133)

a. Size of Group. Simmel noted that the number of individuals in a group in which social action takes place affects the form of group interaction. Relationships in a two person group, what Simmel called a dyad, are relatively straightforward, in that each individual can present themselves to the other in a way that maintains their identity, and either party can end the relationship by withdrawing from it. When a dyad changes to a triad, a three person group, the form of interaction may alter. In the triad, there may be strategies that lead to competition, alliances, or mediation. The triad is likely

to develop a group structure that is independent of the individuals in it, whereas this is less likely in the dyad (Ritzer, p. 166).

As group size increases even more, "the increase in the size of the group or society increases individual freedom" (Ritzer, p. 167). The small circle of early or premodern times,

firmly closed against the neighbouring strange, or in some way antagonistic circles ... allows its individual members only a narrow field for the development of unique qualities and free, self-responsible movements. ... The self-preservation of very young associations requires the establishment of strict boundaries and a centripetal unity. (Farganis, p. 140).

As the group grows in numbers and extends itself spatially, "the group's direct, inner unity loosens, and the rigidity of the original demarcation against others is softened through mutual relations and connections" (Farganis, p. 140). This implies much greater possibility of individual freedom and flexibility, with the common culture and form of association greatly weakened. Note the similarity to Durkheim's analysis.

The metropolis or city becomes the location where the division of labour is the greatest and where this individuality and individual freedom is most expanded. At the same time Simmel notes that for the individual this creates the "difficulty of asserting his own

personality within the dimensions of metropolitan life" (Farganis, p. 142). The growth of the city, the increasing number of people in the city, and the "brevity and scarcity of the inter-human contacts granted to the metropolitan man, as compared to the social intercourse of the small town" (Farganis, p. 143) makes the "objective spirit" dominate over the "subjective spirit." Modern culture in terms of language, production, art, science, etc. is "at an ever increasing distance." This results from the growth of the division of labour and the specialization in individual pursuits that is a necessary part of this. Subjective culture is "the capacity of the actor to produce, absorb, and control the elements of objective culture. In an ideal sense, individual culture shapes, and is shaped by, objective culture. The problem is that objective culture comes to have a life of its own" (Ritzer, p.162). "The individual has become a mere cog in an enormous organization of things and powers which tear from his hands all progress, spirituality, and value in order to transform them from their subjective form into the form of objective life" (Farganis, p. 143). This sounds much like Marx's alienation, Durkheim's anomie, or Weber's rationalization, although Simmel associates this with the city, rather than with the society as a whole, as do the other classical writers.

Where Simmel differs from these other classic writers, is that Simmel returns to the individual, analyzes how the individual deals with the developments of modern

society, and considers how the individual personality is developed in these circumstances. Simmel notes that one way individuals assert a personality is to "be different," to adopt manners, fashions, styles, "to appear concentrated and strikingly characteristic." The brevity and fleetingness of contact in the city mean that lasting impressions based on regular and habitual interaction with others cannot be developed. In these circumstances, obtaining self-esteem and having "the sense of filling a position" may be developed by seeking "the awareness of others" (Farganis, p. 143). This means that individuals may adopt some characteristic fashions and in their

personal mannerisms may try to appear "to the point." Note that the personality is not an isolated entity but also is a social entity, one that depends on interaction. Social interaction, looking to the reaction of others, and seeking the recognition and awareness of others is an essential aspect of individual personality. In this way Simmel ties together the individual and the social, and each requires the existence of the other.

Further, the intellect and personal psyche develop in a different way in traditional and in modern society. In rural and small town settings, impressions of others are built up gradually, over time, on the basis of habit. Many of these impressions are less conscious and are built on more deeply felt and emotional relationships. (Farganis, p. 136). In contrast, in the city, there is sharp

discontinuity, single glances, a multitude of quick impressions.

Thus the metropolitan type of man – which, of course, exists in a thousand individual variants – develops an organ protecting him against the threatening currents and discrepancies of his external environment which would uproot him. He reacts with his head instead of his heart. Intellectuality is thus seen to preserve subjective life against the overwhelming power of metropolitan life, and intellectuality branches out in many directions and is integrated with numerous discrete phenomena. (Farganis, p. 137)

Thus Simmel views objective culture as having an effect on the individual, but at the same time considers how this alters the development of the individual, how the individual understands this and develops in this context, how the individual interacts with other individuals, and how these interactions form the social life of the city. Simmel concludes his essay by noting how the city influences individuals and provides the "opportunities and the stimuli for the development of ... ways of allocating roles to men. Therewith these conditions gain a unique place, pregnant with inestimable meanings for the development of psychic existence" (Farganis, p. 144). Note "allocating roles to men" rather than "men to roles" as the structural functionalist might describe this process. While Simmel is concerned with the possible negative effects of objective culture, he considers it

possible for personalities to develop within these conditions.

b. Individual and Society. For Simmel, there is a dynamic or dialectical tension between the individual and society – individuals are free and creative spirits, yet are part of the socialization process. Simmel was troubled by this relationship. He viewed modern society as freeing the individual from historical and traditional bonds and creating much greater individual freedom, but with individuals also experiencing a great sense of alienation within the culture of urban life. Simmel notes:

The deepest problems of modern life derive from the claim of the individual to preserve the autonomy and individuality of his existence in the face of overwhelming social forces, of external culture, and of the technique of life. (Farganis, p. 136).

Simmel makes three assumptions about the individual and society. (Ashley and Orenstein, p. 312). These are:

- Individuals are both within and outside society.
- Individuals are both objects and subjects within networks of communicative interaction.
- Individuals have the impulse to be self-fulfilling and self-completing, that is, they seek an integrated self-concept. Society also tries to integrate itself (like Durkheim noted), although the effect of this may be in opposition to individual integrity.

In the social world, the various forms and styles of interaction are brought into existence by people and the above assumptions are realized as individuals interact with one another. Humans possess creative consciousness and the basis of social life is "conscious individuals or groups of individuals who interact with one another for a variety of motives, purposes, and interests" (Ritzer, p. 163) People are conscious and creative individuals and the mind plays a crucial role in this mutual orientation and social interaction. This creativity allows for flexibility and freedom on the part of the individual, and at the same time helps

to create the structures of objective culture that may constrain and stifle this freedom. That is, social interaction becomes regularized and has patterns to it, and these become forms of association. These patterns and forms, regardless of their content, is what sociologists should study.

This means that society is not a separate reality of its own, but "society merely is the name for a number of individuals, connected by interaction ... society certainly is not a 'substance,' nothing concrete, but an *event*: it is the function of receiving and affecting the fate and development of one individual by the other." For Simmel, society is nothing but *lived experience*, and social forces are not external to, nor necessarily constraining for the individual, rather it is individuals who reproduce society every living moment through

their actions and interactions. Simmel disagreed with Durkheim that "society is a real, material entity" and did not view society as merely a collection of individuals. Rather, he adopted the position of "society as a set of interactions" (Ritzer, p. 170).

The individual in a social unit must be an entity or constituent part of the unit, and Simmel distinguishes between a personal self and a social self. If there is no self-consciousness, symbolic interaction would disappear and human experience would just be the responses to stimuli. Instead, we live and die in terms of what is intersubjectively meaningful – i. e. we view ourselves in terms of responses of others – even others who we have never met.

c. Fashion. An example of how Simmel examines some of these connections in a concrete connection is his discussion of fashion. (See Ritzer p. 161 and Ashley and Orenstein, pp. 314-5). Simmel notes that fashion develops in the city, "because it intensifies a multiplicity of social relations, increases the rate of social mobility and permits individuals from lower strata to become conscious of the styles and fashions of upper classes" (Ashley and Orenstein, p. 314). In the traditional and small circle setting, fashion would have no meaning or be unnecessary. Since modern individuals tend to be detached from traditional anchors of social support, fashion allows the individual to signal or express his or

her own personality or personal values. Simmel noted that fashion provides

the best arena for people who lack autonomy and who need support, yet whose self-awareness nevertheless requires that they be recognized as distinct and as particular kinds of beings. (Ashley and Orenstein, p. 314).

Ritzer notes that fashion can be considered to be a part of objective culture in that it allows the individual to come into conformity with norms of a group. At the same time, it expresses individuality, because the individual may differ from the norm. Fashion is dynamic and has an historical dimension to it, with acceptance of a fashion being followed by some deviation from this fashion, change in the fashion, and perhaps ultimate abandonment of the original norm, so that a new norm is established. This is a dialectical process, with initial success, widespread acceptance, followed by eventual abandonment and failure. Leadership in a fashion means that the leader actually follows the fashion better than others. Mavericks are those who reject the fashion, and this may become an inverse form of imitation.

In summary, fashion allows personal values to be expressed at the same time as norms are followed. The two exist together, and the one without the other would be meaningless. In all of this, social interaction is of the

essence – what others think, what one thinks that others think, and how one conceives of fashion.

d. Philosophy of Money. Simmel's major work concerns money and the social meaning of money. In *The Philosophy of Money*, Simmel is concerned with large social issues, and this book can be thought of as on a par with *The Division of Labour* of Durkheim, although not as extensive and thorough as Marx's *Capital* or Weber's *Economy and Society*. Simmel considers money as a symbol, and examines some of its effects on people and society. In modern society, money becomes an impersonal or objectified measure of value. This implies impersonal, rational ties among people that are institutionalized in the money form. For example, relations of domination and subordination become quantitative relationships of more or less money – impersonal and measurable in a rational and calculated manner. The use of money distances individuals from objects and also provides a means of overcoming this distance. The use of money allows much greater flexibility for individuals in society – to travel greater distances and to overcome person-to-person limitations.

Simmel thus suggests that the spread of the money form gives individuals a freedom of sorts by permitting them to exercise the kind of *individualized* control over "impression management" that was not possible in

traditional societies. ... ascribed identities have been discarded. Even strangers become familiar and knowable identities insofar as they are willing to use a common but impersonal means of exchange. (Ashley and Orenstein, p. 326)

At the same time, personal identity becomes problematic, so that development of monetary exchange has both positive and negative implications. That is, individual freedom is potentially increased, but alienation and fragmentation may occur.

e. Conclusion. In some senses, Simmel's sociology is similar to that of the other classic writers, although he had less to say about social structure or its dynamics than did Marx, Weber, or Durkheim. He discussed objective culture and his writings on money have some affinity with Weber's rationalization. Simmel's emphasis on social interaction and the resulting social world provide a unique contribution to the interaction perspective from the European classical period. His analysis of fashion, money, and the city also make his writings worthwhile reading.

4. Philosophical and Psychological Approaches

The symbolic interaction approach was first developed in the United States by social scientists who were familiar with pragmatism and behaviouralism.

Pragmatism is probably the most distinctive American school of philosophy. Dominant in the United States at the beginning of the twentieth century, pragmatism stressed an open-ended and practical conception of truth. (Knapp, p. 180).

According to the pragmatist view, reality is not "out there" in the world, but exists only as it is actively created by individuals in the world. Several principles of pragmatist thought are as follows:

- Interaction between individuals and the world around them is key to developing human knowledge.
- This interaction is not of a static or fixed type, but is continually changing.
- The individual has considerable flexibility, and is social and creative in his or her ability to interpret the social world.

All of these features of social life are demonstrated through the way that we use language and the manner in which we fill our various social roles. The philosophical pragmatism of William James and John Dewey was an important influence on Mead, Blumer, and other symbolic interaction writers.

a. William James (1842-1910) taught at Harvard and wrote about psychology and philosophy and attempted to develop moral and ethical principles for meaning and truth that depend on the definite difference these make to

people (Knapp, p. 180). For James, consciousness is active, selective and interested, and direct experience is an especially important aspect of this. Ideas are not absolutes, but are a way of preparing for and anticipating experiences.

b. John Dewey (1859-1952) was an important American writer who spent most of his academic life at Columbia University. Dewey argued that the various types of human activity are instruments that are developed to solve the various problems faced by humans. There is no eternal truth, but rather truth is based on experience, testable by all who investigate it. For Dewey, the human mind was not just a thing or a structure, but an active process by which the individual imagines, interprets, decides, defines, and acts in the world. Dewey attempted to work out principles for a democratic and industrial society, and was an

opponent of authoritarian methods in education. As founding president of the American Civil Liberties Union, Dewey was an important intellectual influence in American life. (Knapp, p. 180 and *Columbia Encyclopedia*, p. 756).

c. Behaviourism is the psychological approach that explains animal and human behaviour in terms of observable and measurable responses to stimuli (*Columbia Encyclopedia*, p. 261). In this approach, mental processes are not as important as the stimulus and

the response so that the observed stimuli and the observed set of behaviours should be the subject of psychological study. The behaviourist approach was influential on interaction theory in a negative way, with the latter developing to counter the former. While the unit of study for Mead was the act, Mead argued that there were mental processes involved in actions, processes that the behaviourists ignored. Human mental processes differed from those of non-human animals, and this meant that human behaviour had to be studied differently than non-human behaviour. Mead's approach can be considered to be a social behaviourist approach, emphasizing social rather than biological or mental processes (Ritzer, pp. 327-30).

5. Chicago School

In 1890, the department of sociology at the University of Chicago was founded by Albion Small. This department was dominant in American sociology for 40-50 years, until Parsons and Harvard University became more influential. The Chicago school also had an impact on Canadian sociology and one of the major figures in the Chicago school (Ernest Burgess) was from Canada. Small also founded the American Journal of Sociology in 1895 and translated some of Simmel's essays for this journal.

One of the major figures in the Chicago school was Robert Park (1864-1944). After working as a journalist,

he graduated from the University of Michigan, studied in Germany with Simmel, and obtained a doctorate at Heidelberg in 1904. He was associated with Booker T. Washington until 1914, when he joined the University of Chicago. Park published *Introduction to the Science of Society*, which introduced some of the major European theorists to the United States. Park looked on "the city as a giant social experiment, consisting of different worlds, neighbourhoods, and groups which are connected to each other and in conflict with each other" (Knapp, p. 175). Some of Park's ideas were derived directly from Simmel. Park told his students:

Go sit in the lounges of the luxury hotels and on the doorsteps of the flophouses; sit on the Gold Coast settees and on the slum shakedowns; sit in the Orchestra Hall and in the Star and Garter Burlesk. (from Knapp, p. 174).

In order to understand society, Park considered it necessary to study the city, the structure of the social world in which people live, and the relationships of people to each other – using "the 'moving camera' of the naturalistic approach to catch life as it was happening" (Wallace and Wolf, p. 195). He emphasized communities and changes in these, how individuals were shaped by and integrated into these communities, and how people and groups formed communities. Park stressed the importance of both social research and social reform. For Park and others in the Chicago school, the city was "a social laboratory in which human nature and

social processes could be examined" (Shore, p. 127). These researchers were also concerned with the "expanding metropolis and the influence it exerts over contiguous regions" (Shore, p. 121).

Another Chicago school sociologist was W. I. Thomas (1863-1947), co-author of the two volumes of *The Polish Peasant in Europe and America* (1918, 1920), an important empirical study of immigration and immigrant adjustment in the United States. W. I. Thomas, along with Dorothy Swaine Thomas, noted how individuals involved in interaction define the situation as important in understanding how interaction occurs. If the individual defines situations as real, they are real in their consequences (Wallace and Wolf, p. 194). Sociological researchers must "pay attention to subject meanings or definitions" in order to understand human activity (Wallace and Wolf, p. 194). The Thomases noted that individuals can ignore a particular stimulus or examine and deliberate concerning a situation, before taking action (or not acting).

6. Mead

Mead is generally regarded as the founder of the symbolic interaction approach. George Herbert Mead (1863-1931) was trained in social psychology and philosophy and spent most of his academic career at the University of Chicago. Mead's major work is *Mind, Self and Society*, a series of his essays put together after

Mead's death and originally published in 1934, a work in which he emphasizes how the social world develops various mental states in an individual.

Mead looked on the "self as an acting organism, not a passive receptacle that simply receives and responds to stimuli" (Wallace and Wolf, p, 197), as Durkheim and Parsons may have thought. People are not merely media that can be put into action by appropriate stimuli, but that "we are thoughtful and reflective creatures whose identities and actions arise as a result of our interactions with others" (Farganis, p. 145).

For Mead, what distinguishes humans from non-human animals is that humans have the ability to delay their reactions to a stimulus. Intelligence is the ability to mutually adjust actions. Non-human animals also have intelligence because they often can act together or adjust what they do to the actions of other animals. Humans differ from non-human animals in that they have a much greater ability to do this. While humans may do this through involuntary gestures, Mead thought it more important that it is only humans that can adjust actions by using significant or meaningful symbols. As a result of this greater intelligence, humans can communicate, plan, and work out responses, rather than merely reacting in an instinctive or stimulus-response manner.

Mead notes that human actions have three characteristics: (i) humans are able to organize their

minds concerning the array of possible responses open to them; (ii) humans can consider the likely implications of different actions, and test possible outcomes mentally in their own minds; and (iii) since there are a range of stimuli that impinge upon an individual, a human need not react to the immediate stimulus, but may react to one of the lesser stimuli. This means that humans are able to make choices that are better adapted to the situation and "intelligence is largely a matter of selectivity" (Ritzer, p. 339).

For Mead, rather than action being defined by:

Stimulus Response

action is more appropriately identified with the following sequence of events:

Meaning

Stimulus Interpret and Define Response

That is, the stimulus-response pattern is not what characterizes social interaction, but rather what happens between stimulus and response. Here humans go through a process of interpreting and defining the stimulus before providing a response. Associated with this is meaning – "the wedding of different attitudes and the use of significant symbols that have the same import for all concerned" (Wallace and Wolf, p. 202). "When individuals share symbolic interpretations, the act is meaningful to them" (Wallace and Wolf, p. 202). For

Mead, symbols are important in allowing human interaction to occur, and it is the shared understanding of the significance of symbols that and what they denote that makes for social interaction.

a. Mind. Mead conceived of the mind as the processes involved in responding to stimuli and contemplating action, with these being almost more important than the physiological processes of the brain, the structure of knowledge, or the contents of individual knowledge. The mind is also social, rather than being purely a characteristic of the brain or the individual. That is, the mind develops as a result of social interaction, the mind is part of social processes, and since the latter precede the mind, society is prior to the mind and self for Mead. While Mead's concept of the mind is less clear than that of the self, Ritzer notes that the mind "has the ability to respond to the overall community and put forth an organized response." This is not just a particular response, but one that can have meaning for the community as a whole, with symbols playing a major role. Further, the mind "involves thought processes oriented toward problem solving."

b. Self. The self is the central social feature in the symbolic interaction approach. Instead of being passive and being influenced by values or structures, Mead considers the self as a process that is active and creative – taking on the role of others, addressing the self by

considering these roles, and then responding. This is a reflexive process, whereby an individual can take himself or herself to be both subject and object. This means that "the individual is an object to himself, and, so far as I can see, the individual is not a self in the reflexive sense unless he is an object to himself" (Mead, quoted in Farganis, p. 148).

Wallace and Wolf note that Mead distinguished "things" from "objects" and the "I" from the "me." When an individual is involved in a situation and acts, this action occurs in an environment. Physical things or stimulæ exist in the environment, prior to action, and people encounter these. By considering these things and acting in response to them, following self-reflection and interpretation, these things become objects. In doing this, individuals are active and creative. Those things of which the individual is conscious are those that the individual takes note of and indicates to the self. This has two consequences. (i) By being conscious of certain things, the individual makes these things into objects, and these are more than stimuli. The individual "constructs his objects on the basis of his on-going activity." These objects then become meaningful for the individual and "This is what is meant by interpretation or acting on the basis of symbols" (Blumer in Wells, p. 92). (ii) This also means that acts are "constructed or built up instead of being a mere release." The act is considered, in the context of the surroundings including the possible

responses of others, and the overall consequences that are anticipated by the actor. Action is thus conscious and is not just a reaction to a stimulus.

Similarly, the I is the impulsive and spontaneous, unorganized response of the individual, whereas the me is the organized self that is learned in interaction with others and which guides the behaviour of the socialized person. The I allows for spontaneity, innovation, and individuality, and the me is that part of the self which involves the influence of others. The self involves both the I and the me, with acts resulting from the dialogue between the two.

Self-interaction is the way in which the individual takes things into account and organizes himself or herself for action. As the social environment changes, or as individuals encounter new or altered experiences, they experiment and interact with themselves in order to find an appropriate response. This involves taking on the role of the other, considering how others will respond, having a conversation with oneself, and forming a means of response which takes all these considerations into account. This may sometimes be quick and not entirely conscious, as in fairly routine situations such as buying food at the cafeteria. At other times, it may involve a long period of conscious role-playing, for example in preparing for a job interview. In either case though, some self-interaction does take place, in that each action is unique and is a result of the individual using the

information from previous experiences and what the individual understands about the environment and situation, in order to act appropriately in the future.

Humans are distinctive in having the ability to be able to have a conversation with themselves, to imagine themselves in the position of other people, to consider what the other person imagines, and contemplates what the reaction of the other person is likely to be. This is evident in communication with the other person, where the individual carries on a conversation with himself or herself (although this is covert and

in the mind, and is not stated for others to hear) at the same time as the conversation with the other person is carried on.

Development of the Self. Mead spends considerable time discussing the development of the child, because this is how the self is created. The first stage of development of the self involves imitative acts on the part of the child. This is the pre-play stage, around age two, where the child does not have the ability to take on the role of others, but merely imitates the actions of others. A play stage follows, where the child can act out the parts of others but cannot yet relate to the role of others. That is, the child repeats what others say, and takes on several roles, one at a time. Later, the child is able to act with others and anticipate the actions of others. This is the game stage, where the child can take

on the role of all the others involved in the game or situation. In doing this, the child learns the organized attitudes of the whole community, and is able to act in common with others. The final stage in socialization is the internalization of the generalized other, whereby people can put themselves in the position of the other person, imagine how others will react, and from that contemplate various courses of action. Once this ability is developed, the individual has a self which is individual, yet could not have developed apart from the community. That is, "one has to be a member of a community to be a self" (Wallace and Wolf, p. 202).

Simple and then more complex situations and games are means by which the child develops a self, and these situations illustrate the nature of more general social processes – interacting with family and friends and taking part in social relationships. Other than games, the development of the individual's ability to communicate using language and other symbols also play an important role in this. In the use of different forms of language, the child learns what others think and how others might respond. Games and learning a language are both social – they could not occur in the isolated individual.

c. Society. The third major part of Mead's approach is society. The ongoing symbols and social processes that exist are logically and historically prior to the development of the mind and self. Institutions that give the common responses of society and the regular habits

of the community are the context within which the mind and self are created. Socialization and education are the means by which individuals internalize these common habits. Mead does not see these as coercive or oppressive, and feels that individual creativity can exist within this. Social institutions can be viewed as constraining on individuals but these same institutions can also be viewed as enabling people to become creative individuals (Ritzer, p. 347). Mead did not develop a macro view of society and social institutions as a whole, but his approach might be combined with some of the more structural approaches to provide a more integrated view of the macro and micro approaches. Note that the classical sociologists have a similar conception of society to that of Mead, but they do not have a theory of the self, and they do not emphasize interaction.

d. Symbolic Meaning. For Mead, significant symbols are those "which will call out in another that which it calls out in the thinker" (Mead in Farganis, p. 150). Symbols of this sort are universal (rather than particular) and are involved in the process of thinking – "an internalized or implicit conversation of the individual with himself" (Mead in Ritzer, p. 338) using gestures or symbols. Language is a set of vocal gestures which are significant symbols carrying social meaning. Thinking is implicit conversation, or covert behaviour -- that is, it "is

not a mentalistic definition of thinking; it is decidedly behavioristic" (Ritzer, p. 338).

While Weber considered meaning to be essential to defining what is social, he did not provide a very clear idea of how he defined meaning or what aspects of meaning were important. In contrast, Mead makes meaning an essential part of definition and development of self. "Meaning as such, i.e., the object of thought, arises in experience through the individual stimulating himself to take the attitude of the other in his reactions toward the object" (Wallace and Wolf, p. 201). That is, meaning develops through experiences, as different individuals develop a common understanding of social situations and symbols. When symbolic interpretation is shared, people see things in the same light, and acts are meaningful to actors. As a result of this common understanding, the gesture or symbol arouses the same attitude in the individuals, and this is sufficient to trigger a reaction.

A symbol is "the stimulus whose response is given in advance" (Wallace and Wolf, p. 204). This could be a set of words, a gesture, a look, or a more fixed, material symbol such as a flag, crest, or money. When actors in a situation have developed a common understanding of symbols, it is significant for them and has meaning, in that they understand how the symbol will be interpreted by others and what the response is likely to be. This understanding is developed from previous experiences

where the likely responses of others to the symbols has been observed or understood. As a result, the symbols have meaning for the individuals, and these allow individuals to interact with each other.

References

Ashley, David and D. M. Orenstein, *Sociological Theory: Classical Statements*, Allyn and Bacon, Boston, 1990, second edition.

Blumer, H., "Society as Symbolic Interaction," in A. Wells, ed., *Contemporary Sociological Theories*, Goodyear Publishing, Santa Monica, 1978, pp. 91-98 and also in Rose, pp. 179- 192.

Denzin, Norman K., "Sexuality and Gender: an Interactionist / Poststructural Reading," in P. England, ed., *Theory on Gender / Feminism on Theory*, Aldine de Gruyter, New York, 1993, pp. 199-221.

Denzin, Norman K., *Interpretive Interactionism*, Sage, Newbury Park, California, 1989, Applied Social Research Methods Series, Volume 16. HM 24 D438 1989

Denzin, Norman K., *Symbolic Interactionism and Cultural Studies: the Politics of Interpretation*, Blackwell, Oxford, 1992. HM 291 D454 1992

Farganis, J., *Readings in Social Theory: the Classic Tradition to Post-Modernism*, McGraw- Hill, New York, 1993.

Frisby, David, *Georg Simmel*, Ellis Horwood, Chichester and Tavistock, London, 1984, Key Sociologists Series. HM 22 G3S482 1984.

Frisby, David, *Simmel and Since: Essays on Georg Simmel's Social Theory*, London, Routledge, 1992. HM 22 G3 S4828.

Frisby, David, *Sociological Impressionism: A Reassessment of Georg Simmel's Social Theory*, London, Routledge, 1992. Second edition. HM 22 G3 S483.

Knapp, P., *One World – Many Worlds: Contemporary Sociological Theory*, Harper-Collins, New York, 1994).

Sharrock, W., "Individual and Society," in R. J. Anderson, J. A. Hughes and W. W. Sharrock, *Classic Disputes in Sociology*. HM51 C54.

Simmel, Georg, *The Philosophy of Money*, Routledge, London, 1990), second edition. HG 221 S5913 1990.

Turner, H. H., *The Structure of Sociological Theory*, Wadsworth, Belmont, California, 1991, fifth edition.

Wolff, Kurt, *The Sociology of Georg Simmel*, Free Press, Glencoe, Illinois, 1950. HM 57 S482

Zorbaugh, Harvery W., *The Gold Coast and the Slum*, Chicago, University of Chicago, 1929). F 548.5 Z89

Last edited on November 9, 1999 Return to [Sociology 250](#).

Internet Encyclopedia of Philosophy

A Peer-Reviewed Academic Resource



Internet Encyclopedia of Philosophy

SKIP TO CONTENT

ABCDEFGHIJKLMNOPQRSTUVWXYZ

Jürgen Habermas (1929—)

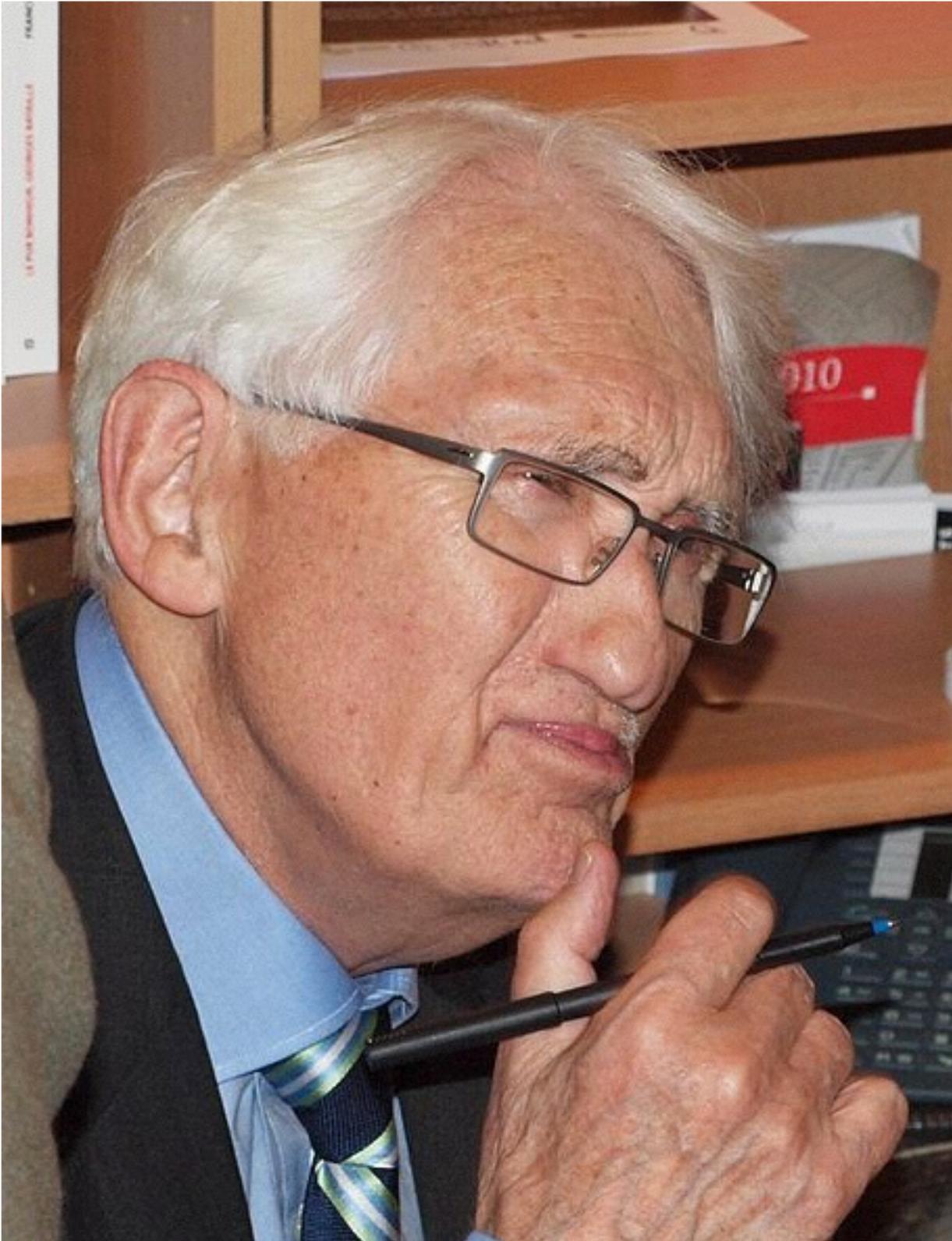


photo by Ziel

Jürgen Habermas produced a large body of work over more than five decades. His early work was devoted to the public sphere, to modernization, and to critiques of trends in philosophy and politics. He then slowly began to articulate theories of rationality, meaning, and truth. His two-volume *Theory of Communicative Action* in 1981 revised and systematized many of these ideas, and inaugurated

his mature thought. Afterward, he turned his attention to ethics and democratic theory. He linked theory and practice by engaging work in other disciplines and speaking as a public intellectual. Given the wide scope of his work, it is useful to identify a few enduring themes.

Habermas represents the second generation of Frankfurt School Critical Theory. His mature work started a “communicative turn” in Critical Theory. This turn contrasted with the approaches of his mentors, Max Horkheimer and Theodor W. Adorno, who were among the founders of Critical Theory. Habermas sees this turn as a paradigm shift away from many assumptions within traditional ontological approaches of ancient philosophy as well as what he calls the “philosophy of the subject” that characterized the early modern period. He has instead tried to build a “post-metaphysical” and linguistically oriented approach to philosophical research.

Another contrast with early Critical Theory is that Habermas defends the “unfinished” emancipatory project of the Enlightenment against various critiques. One such critique arose when the moral catastrophe of WWII shattered hopes that modernity’s increasing rationalization and technological innovation would yield human emancipation. Habermas argued that a picture of Enlightenment rationality wedded to domination only arises if we conflate instrumental rationality with rationality as such—if technical control is mistaken for the entirety of communication. He subsequently developed an account of “communicative rationality” oriented around achieving mutual understandings rather than simply success or authenticity.

Another enduring theme in Habermas’ work is his defense of “post-national” structures of political self-determination and transnational governance against more traditional models of the nation-state. He sees traditional notions of national identity as declining in importance; and the world, as faced with problems stemming from interdependency that can no longer be addressed at the national level. Instead of national identity centered on shared historical traditions, ethnic belonging, or national culture, he advocates a “constitutional patriotism” where political commitment, collective identity, and allegiance coalesce around the shared principles and procedures of a liberal democratic constitutionalism

facilitating public discourse and self-determination. Habermas also claims that emerging structures of international law and transnational governance represent generally positive achievements moving the global political order in a cosmopolitan direction that better protects human rights and fosters the spread of democratic norms. He sees the emergence of the European Union as paradigmatic in this regard. However, his cosmopolitanism should not be overstated. He does not advocate global democracy in any strong sense, and he is committed to the idea that democratic self-determination requires a measure of localized mutual identification in the form of civic solidarity—a legally mediated solidarity around shared history, institutions, and rooted in some shared “ethical” pattern of life (see *Sittlichkeit* discussion below) fostering mutual understandings.

Table of Contents

- 1 [Biography: Early Life to Structural Transformation](#)
- 2 [Enduring Themes in Formative and Transitional Work](#)
 - 1 [Public Deliberation Over Positivist Decisionism and Technocracy](#)
 - 2 [From Philosophical Anthropology to a Theory of Social Evolution](#)
- 3 [The Linguistic Turn into the Theory of Communicative Action](#)
- 4 [Discourse Ethics](#)
- 5 [Political and Legal Theory](#)
- 6 [References and Further Reading](#)
 - 1 [General Introductions to Habermas](#)
 - 2 [Introductory Books and Articles on Specific Themes](#)
 - 1 [Biography](#)
 - 2 [Linguistic Turn](#)
 - 3 [Discourse Ethics](#)
 - 4 [Political Theory](#)
 - 3 [Works Cited](#)
 - 4 [Secondary Scholarship Beyond the Subject-Specific Recommendations Cited Above](#)

1. Biography: Early Life to Structural Transformation

Habermas was born in 1929 in Düsseldorf, Germany. He has noted that early corrective surgeries for a cleft palate sensitized him to human vulnerability and interdependence, and that subsequent childhood struggles with fluid verbal communication may partly explain his theoretical interest in communication and mutual recognition. He has also cited the end of WWII and frustrations over postwar Germany's uneven willingness to fully break with its past as key personal experiences that inform his political theory.

Habermas belongs to what historians call the "*Flakhelfer* generation" or the "forty-fivers." *Flakhelfer* means antiaircraft-assistant. At the end of the war, people born between 1926 and 1929 were drafted and sent to help man antiaircraft artillery defenses. Over a million youth served as such personnel. The second "forty-fiver" label captures how this generation came of age with the 1945 Nazi defeat. These experiences fostered a political skepticism and vigilance born out of having been exploited, and an affinity for the nascent liberal democratic principles of postwar Germany. Both labels capture formative features of Habermas' biography (Specter 2010, Matustik 2001).

Reflecting on his upbringing during the war, Habermas describes his family as having passively adapted to the Nazi regime—neither identifying with nor opposing it. He was recruited into the Hitler Youth in 1944 and sent to man defenses on the western front shortly before the war ended. Soon thereafter he learned of the Nazi atrocities through radio broadcasts of the Nuremberg trials and concentration camp documentaries at local theaters. Such experiences left a deep impact: "all at once we saw that we had been living in a politically criminal system" (AS 77, 43, 231).

After the war, he studied philosophy at the universities of Göttingen (1949-50), Zurich (50-51) and Bonn (51-54). He wrote his thesis on Schelling under the direction of Erich Rothacker and Oskar Becker. He was increasingly frustrated with the unwillingness of German politicians and academics to own up to their role in the war. He was disappointed in the postwar government's failure to make a fresh political start and distressed by continuities with the past. In interviews, he has recalled leaving a campaign rally in 1949 after being disgusted by the far-right connotations of the flags and songs used. He was similarly disappointed by German academics. At

university he studied the work of Arnold Gehlen and Martin Heidegger extensively, but their prior Nazi ties were not discussed openly. In 1953 Heidegger reissued his 1935 *Lectures on Metaphysics* in a largely unedited form that included reference to the “inner truth and greatness of the Nazi movement.” Habermas published an op-ed challenging Heidegger, and the lack of response seemed to confirm his suspicions (NC, 140-172). He wrote a piece critiquing Gehlen a few years later (1956). Around the same time he was distressed to learn Rothacker and Becker had also been active Nazi party members.

Near the end of his studies Habermas worked as a freelance journalist and published essays in the intellectual journal *Merkur*. He took an interest in the interdisciplinary Institute for Social Research affiliated with the University of Frankfurt. The Institute had returned from wartime exile in 1950, and Adorno became director in 1955. Adorno was familiar with Habermas’ essays and took him on as a research assistant. While at the Institute Habermas studied philosophy and sociology, worked on research projects, and continued to publish op-ed pieces. One such piece, *Marx and Marxism*, struck Horkheimer as too radical. Horkheimer wrote to Adorno suggesting he dismiss Habermas from the Institute. The following year Horkheimer rejected Habermas’ Habilitationsschrift proposal on the public sphere. Habermas did not want to alter his project, so he completed his dissertation at the University of Marburg under the Marxist political scientist Wolfgang Abendroth. His Habilitationsschrift, *The Structural Transformation of the Public Sphere* (German 1962, English 1989), was well received in Germany. It chronicled the rise of the bourgeois public sphere in 18th and 19th century Europe, as well as its decline amidst the mass consumer capitalism of the 20th century. Habermas gave an account of the way in which newspapers, coffee shops, literary journals, pubs, public meetings, parliament and other public forums facilitated the emergence of powerful new social norms of discourse and debate that mediated between private interests and the public good. These forums functioned as mechanisms to disseminate information and help freely form the public political will needed for collective self-determination. These norms also partly embodied important principles like equality, solidarity, and liberty. By the late

19th century, however, capitalism was increasingly monopolistic. Large corporations easily influenced the state and society. Economic elites could use ownership of the media and other (previously public) forums to manipulate or manufacture public opinion and buy-off politicians. Citizens deliberating about the common good were transformed into atomized consumers pursuing private interests. Habermas describes this as the “re-feudalization” of the public sphere. While his narrative was pessimistic, the end of *Structural Transformation* seems to hold out hope that the truncated normative potential of the public sphere may yet be revived. The work solidified Habermas’ place in the German academy. After a short stint in Heidelberg, he returned to the University of Frankfurt in 1964 as a professor of philosophy and sociology, taking over the chair vacated by Horkheimer’s retirement.

In the spirit of his early call for renewed public sphere debate, Habermas has consistently engaged political movements as a public intellectual and taken part in various scholarly debates. This has not always been easy. After returning to Frankfurt he had been a mentor for the German student movement, but had a falling out with student radicals in 1967. In June of that year a variety of simmering protests—over the restructuring of German universities, proposed “emergency laws,” the Vietnam War, and other issues—boiled over. The breaking point was when a student at a protest against the Shah of Iran was shot and fatally beaten by plainclothes police, who then tried to cover up the incident. This stoked the flames of student protests. Sit-ins and protests crippled everyday life. Under the leadership of Rudi Dutschke students occupied the Free University of Berlin.

Habermas worried that protest leaders seemed to be advocating an unsophisticated and extra-legal opposition to any and all authority that could easily lead to violence. At a conference in Hannover shortly after the shooting he publically reproached Dutschke by calling his model of extra-legal direct-action “left fascism.” That charge alienated Habermas from the leftist student movement and inspired an essay collection *Die Linke antwortet Jürgen Habermas* (*The Left Answers Habermas*—German 1969). Rapprochement would only come a decade later when, in the aftermath of a series of

killings by the radical left-wing Red Army Faction, politicians on the right tried to garner political capital by suggesting that such terrorism was rooted in the ideas of Frankfurt School Critical Theory. Habermas and Dutschke published pieces repudiating the accusation. A decade later, the editor of the essay collection apologized for how the book made it seem like Habermas' falling out with the student movement marked a conservative turn that meant he was no longer part of the left.

As a public intellectual, Habermas has engaged a variety of topics: the anti-nuclear movement of the late fifties, the "Euromissile" debate of the early eighties and, in the early two-thousands, both the terrorism of 9/11 and the second Iraq War. In the second half of the eighties he was also a key voice in the *Historikerstreit* debate between historians, philosophers, and other academics about the proper way for Germany to situate and remember the Holocaust amidst the history of other atrocities. In 1989 he made important contributions to public debate about the reunification of Germany. While Habermas was not against reunification, he was critical of the speed and manner in which reunification was carried out. More recently, he has approached public debate on the European Union along the broadly similar lines of a cautious optimism that is also on guard against a forced, rushed, or duped false unity that would lack legitimacy and stability over the long term.

In a more academic vein, he has had numerous exchanges with thinkers like Jacques Derrida, Richard Rorty, Hans-Georg Gadamer, Niklas Luhmann, John Rawls, Robert Brandom, Hilary Putnam, and Cardinal Joseph Ratzinger (before he was Pope Benedict XVI). His ongoing debate with postmodernism is arguably the most enduring line of debate. Broadly speaking, thinkers like Michel Foucault, Jacques Derrida, and Richard Rorty have levied criticisms to the effect that reason is little more than a historically and culturally contingent social form, that notions of universally valid morality and truth are ethnocentric projections of power, that interests shaped by radically different ways of life are irreconcilable, and that our belief in the emancipatory moral progress of humankind is a myth. Habermas has tried to meet such challenges in much the same way as he responded to Horkheimer and Adorno's *Dialectic of Enlightenment*: by relying on his account of

communicative rationality in *Theory of Communicative Action*. However, before turning to that more mature theory, we must survey a few major phases of his formative and transitional work.

2. Enduring Themes in Formative and Transitional Work

a. Public Deliberation Over Positivist Decisionism and Technocracy

The essays in *Towards a Rational Society* (German 1968 and 1969, English 1970) and *Theory and Practice* (German 1971, English 1973b) were written on the heels of *Structural Transformation*. They were written amidst the “positivism dispute” in Germany about the relation between the natural and social sciences. The (somewhat inaccurately labeled) “positivist” side of this debate took scientific inquiry as the sole paradigm of knowledge and generally thought of the social sciences as analogous to the natural sciences. Following Adorno, Habermas argued against a positivistic understanding of the social sciences.

For Habermas, positivism is comprised of three claims: (1) knowledge consists of causal explanations cast in terms of basic laws or principles (for example, laws of nature), (2) knowledge passively reflects or mirrors independently existing natural facts, (3) knowledge is about what *is*, not what *ought to be*. He calls these claims scientism, objectivism, and value-neutrality. He said each can be pernicious, especially in the social scientific realm. Scientism fosters the view that only causal and empirically verifiable hypotheses can count as true knowledge. Objectivism seems to falsely naturalize the world by ignoring how lived experiences, human subjectivity, and interests can structure the object domain that gets identified as relevant or worthy of study. Lastly, value-neutrality misleads us into thinking that the role of knowledge is purely descriptive and technical. Values or preferences are seen as separate from knowledge and, as such, wholly subjective “givens” lying beyond rational justification. In turn, knowledge is seen as a tool for efficiently controlling the environment so as to realize whatever values an agent happens to hold. Ironically, this fails to

see the tacit value commitments already inscribed in this general paradigm of knowledge.

Habermas' critique makes sense given his place in Frankfurt Critical Theory. Despite differences with the first generation, he shares the decidedly non-neutral commitments to human emancipation, interdisciplinarity, and self-reflexive theory. Like Horkheimer and Adorno, Habermas worried the prior ascendancy of positivism had left influences on our conceptualizations of knowledge and social inquiry that were hard for even reflective positivists to leave behind. Indeed, he critiques Karl R. Popper's account of inquiry and knowledge even though it rejects what Habermas calls objectivism. In opposition to a positivist picture of knowledge merely mirroring the world, Habermas holds the Frankfurt School's Hegelian-Marxist-inspired conception of a dialectical relation between knowledge and world. Finally, like his Frankfurt School contemporaries, Habermas was concerned that positivism had left subtle yet pernicious impacts on politics.

In early writings Habermas is especially critical of two related trends, decisionism and technocracy, that stem from a positivistic understanding of political science and practice. Decisionism starts from the assumption that there is no such thing as *the* public interest, but rather a clash of inherently subjective values that do not (even in principle) admit of rational persuasion or agreement. It follows that political elites must either simply decide between competing values or base policy on their aggregation. Either way, political value preferences are taken as brute or static facts; there is no sense in which reasoned argumentation and persuasion could genuinely transform such preferences or lead people to a new understanding of their values. Technocracy builds from this point by emphasizing the "objective necessities" (*Sachzwänge*) supposedly involved in a political system—economic growth, social stability, national security—and highlighting the increasing ability of policy experts to advise political leaders about strategies for optimally realizing these goals. The worry with this approach is that questions about *what specific type* of growth, stability, and security we seek (and why) are removed from debate by definitional fiat. In decisionism, political legitimacy flows from periodic expressions of acclamation or disapproval at the way leaders have manifested

predefined values. In technocracy, legitimacy supposedly flows from the ability of politicians to find and follow expert advice so as to attain fixed outcomes pre-defined by “objective necessities.” Both models render the potentially transformative effects of public deliberation superfluous. Legitimacy is seen as flowing from either certain outcomes or periodic expressions of aggregate preference. Habermas thinks both models are extremely problematic accounts of democratic political practice and legitimacy. While *Structural Transformation* only gestured at how the normative potential of the public sphere could be reinvigorated in contemporary circumstances, this theme received increasing attention in works such as *Legitimation Crisis* (German 1973, English 1975), *Theory of Communicative Action*, and *Between Facts and Norms* (German 1992, English 1996). An account of democratic legitimacy that combats decisionism and technocracy is an enduring concern. Indeed, despite championing the European Union he has continued to critique technocracy by criticizing the way in which it has arisen and is currently structured (2008, 2009, 2012, 2014).

b. From Philosophical Anthropology to a Theory of Social Evolution

Knowledge and Human Interests (German 1968, English 1971) and *Communication and the Evolution of Society* (German 1976, English 1979) are two early attempts at a new systematic framework for Critical Theory. The approaches he uses are akin to the tradition of “philosophical anthropology” in the German social theory of the early 1900s that grew out of phenomenology—a tradition that is quite different from contemporary anthropology. *Knowledge and Human Interests* sought to overcome positivist epistemology that saw knowledge as simply discerning static facts, and to give a plausible account of the dialectical relation between knowledge (theory) and world (practice). Habermas’ main claim was that the knowledge of scientific and social progress is tacitly guided by three types of “knowledge constitutive interests”—technical, practical, and emancipatory—that are “anthropologically deep-seated” in the human species.

Knowledge and Human Interests tries to recover and develop alternative models of the relation between theory and practice. The

approach is historical and reconstructive in that it interprets the attempts of prior theorists as part of a trajectory that Habermas wants to extend. He reviews prior reformulations of Kant's "transcendental synthesis" (the form-legislating activity making objective experience possible) and his "transcendental unity of apperception" (the unity of the subject having such experience). He also tries to articulate the way in which Hegel relocated such synthesis in the historical development of human subjectivity (*absolute spirit*) and how Marx relocated it in the material use of tools and techniques (*embodied labor*). Habermas wants to add to such a trajectory by rehabilitating their shared insight that the constitution of experience is not generated by transcendental operations but by the worldly natural activities of the human species. Yet he wants to do this in a way that avoids the mistakes of Marx and Hegel as well. He tries to do this by building on his interpretation of Hegel, which was already concisely captured by his essay *Science and Technology as Ideology* (German 1968, included in English 1970).

In that essay he responded to Herbert Marcuse's claim that the technical reason of science *inherently* embodies domination. According to Marcuse, under late capitalism the technical reason of science functions ideologically to collapse intersubjective practical questions about how we want to live together into technical questions about how to control the world to get what we want. Habermas shares Marcuse's concerns, as his criticism of technocracy makes clear. Yet he thinks this dynamic is *contingent* because, taken as an emergent collective project, humankind constitutes how the world shows up in experience through its worldly activity. More specifically, Habermas identifies two irreducibly distinct and dialectically related modes of human self-formation, "labor" and "interaction." Whereas labor is an action type that aims at technical control to achieve success, interaction is an action type that aims at mutual understandings embodied in consensual norms. Marcuse's claim (and his remedy of a "new science") would only stand if the "interaction" of intersubjective collective political choice—including the question of how we use technology—was somehow subsumed or rendered superfluous by the "labor" of technological progress in controlling the external

world. But, given Habermas' views in this period, this is impossible. Interaction and labor seem to be pitched as irreducible and invariant categories of human experience. Neither can be dropped nor can one be subsumed in the other—even if their relation becomes unbalanced.

In *Knowledge and Human Interests*, this division between labor and interaction is recast as the technical and practical interests of humankind. The technical interest is in the *material reproduction* of the species through labor on nature. Humans use tools and technologies to manage nature for material accommodation. The practical interest is in the *social reproduction* of human communities through intersubjective norms of culture and communication. Human social life requires members who can understand each other, share expectations, and achieve cooperation. In a sense, these interests are the “most fundamental.” Moreover, the knowledge that flows from them is supposed to slowly accrue over time in the enduring institutions of society: theoretical knowledge driven by the technical interest in controlling nature accrues in the “empirical-analytic” sciences, and normative knowledge driven by the practical interest in mutual understandings accrues in the interpretive “historical-hermeneutic” sciences.

But, going beyond *Science and Technology as Ideology*, in *Knowledge and Human Interests* Habermas adds a third “emancipatory” human interest in freedom and autonomy. The labor of material reproduction and the interaction norms of social reproduction require, in a weak sense, psychosocial mechanisms to repress or deny basic drives and impulses that would destroy material and social reproduction. For instance, labor requires delayed gratification and social interaction requires internalized notions of obligation, reciprocity, shame, guilt, and so forth. Unfortunately, psychosocial mechanisms of control are often used far more than they need to be to secure material and social reproduction. Indeed, perverse incentives to rely on such mechanisms may even arise: if the burdens and benefits of material and social reproduction processes become unfairly distributed across groups and solidified over time, then those in power may find psychosocial mechanisms useful. If women are falsely taught there

are natural laws of gender relations such that the dominant patterns of marriage and domestic work that consistently disadvantage them are the best they can hope for, this is an *ideological* mechanism of social control. It is the limitation of freedom and autonomy for no purpose other than domination, and it “functions” through systematically distorted communication.

Habermas posits a human interest in using self-reflection and insight to combat ideologically veiled, superfluous social domination so as to realize freedom and autonomy. While there is no clearly institutionalized set of sciences where the knowledge spurred on by such an interest would accrue, Habermas points to Marx’s critique of ideology and Freud’s psychoanalytic dissolution of repression as demonstrating a cognitive viewpoint that focuses on neither (efficient) work nor (legitimate) interaction but (free) identity formation liberated from internalized systematically distorted communication. Here Habermas takes his lead from Kant’s idea that reason aims to emancipate itself from “self-incurred tutelage,” and tries to forge a link between theory (reason) and practice (in the sense of self-realization) through using critical reflection on self and society to unveil and dissolve internalized oppressive power structures that betray one’s own true interests.

Knowledge and Human Interests was envisioned as a preface for two other books that would jointly challenge the separation of theory and practice. However, the project was never finished. On the one hand, Habermas felt that vibrant critiques of positivism in the philosophy of science made the rest of the project superfluous. On the other, the work encountered heavy criticism. For starters, Habermas seems to pitch work and interaction as *real* action types. But, if we account for how work is communicatively structured, interaction is teleologically ordered, and how historical notions of work and interaction structure one’s sense of freedom, then it is clear these can be at best idealizations. Moreover, as even sympathetic interpreters noted, his account of an emancipatory interest seemed to blur together reflection on “general presuppositions and conditions of valid knowledge and action” with “reflection on the specific formative history of a particular individual or group” (Giddens, McCarthy, 95). Lastly, his stipulation

of knowledge-constitutive interests seemed to reproduce the sort of foundationalism he wished to avoid.

Given such criticism, it may seem surprising that *Communication and the Evolution of Society* reconstructs Marx's historical materialism as a theory of social evolution. This sounds foundationalist and deterministically teleological. These impressions are misleading. Around this time Habermas began presenting his work as a "research program" with tentative and fallible claims evaluable by theoretical discourses. Moreover, while he speaks of evolution, he uses the term differently than 19th century philosophies of history (Hegel, Marx, Spencer) or later Darwinian accounts. His "social evolution" is neither a merely path-dependent accumulative directionality nor a progressive, strongly teleological realization of an ideal goal. Instead, he envisions a society's latent potentials as tending to unfold according to an immanent developmental logic similar to the developmental logic cognitive-developmental psychologists claim maturing people normally follow. Lastly, Habermas' theory of social evolution avoids worries about determinism by distinguishing between the *logic* and the *mechanisms* of development such that evolution is neither inevitable, linear, irreversible, nor continuous. A brief sketch of his theory follows.

Habermas characterizes human society as a system that integrates material production (work) and normative socialization (interaction) processes through linguistically coordinated action. This is qualitatively different from the static and transitive status hierarchy systems of even other "social" animals. In various human epochs the linguistic coordination of these processes crystalizes around different "organizational principles" that are the "institutional nucleus" of social integration. In the most basic societies kinship structures play this role by (to take just one possible configuration) dividing labor and specifying socialization responsibilities through sex-based roles and norms. Habermas claims this organizational principle was replaced by political order in traditional societies and the economy in liberal capitalist societies. Social evolution in general and the particular movements from one "nucleus" to the next stem from learning in material and social reproduction.

Understood as ideal types, work and interaction mark out different ways of relating to the world. On the one hand, in material production one mainly adopts an instrumental perspective that tries to control an object in conformity to one's will. In this orientation, learning is gauged by success in controlling the world and the resultant knowledge is cognitive-technical. On the other, in social reproduction one mainly adopts a communicative perspective that tries to coordinate actions and expectations through consensually agreed upon normative standards. In this orientation learning is gauged by mutual understanding and the resultant knowledge is moral-practical. Each learning process follows its own logic. But, since the processes are integrated in the same social system, advances in either type of knowledge can yield internal tensions or incongruities. These cannot be suppressed by force or ideology for long, and eventually need to be solved by more learning or innovation. If these internal tensions are too great, they induce a crisis requiring an entirely new "institutional nucleus."

For Habermas, the slow social learning in history is the sedimentation of iterated processes of individual learning that accumulates in social institutions. While there is no unified macro-subject that learns, social evolution is also not mere happenstance plus inertia. It is the indirect outcome of individual learning processes, and such processes unfold with a developmental logic or deep structure of learning: "the fundamental mechanism for social evolution in general is to be found in an automatic inability not to learn. Not learning but *not-learning* is the phenomenon that calls for explanation" (LC, 15; also see Rasic 2014, 68). Habermas posits a universal developmental logic that tends to guide individual learning and maturation in technical-instrumental and moral-practical knowledge. He discerns this logic in the complementary research of Jean Piaget in cognitive development and Lawrence Kohlberg in the development of moral judgment. As social and individual learning are linked, such underlying logic has slowly created homologies—similarities in sequence and form—between: (i.) individual ego-development and group identity, (ii.) individual ego-development and world-perspectives, and (iii.) the individual ego-development of moral judgment and the structures of law and morality (Owen 2002, 132). Habermas pays more attention to the

last homology and later writings focus on Kohlberg, so it is instructive to focus there (1990b).

Kohlberg's research on how children typically develop moral judgment yielded a schema of three levels (pre-conventional, conventional, and post-conventional) and six stages (punishment-obedience, instrumental-hedonism/relativism, "good-boy-nice-girl", legalistic social-contract/law-and-order, universal ethical principles). Two stages correspond to each level. Habermas follows Kohlberg's three levels in claiming we can retrospectively discern pre-conventional, conventional, and post-conventional phases through which societies have historically developed. Just as normal individuals who progress from child to adult pass through levels where different types of reasons are taken to be acceptable for action and judgment, so too we can retrospectively look at the development of social integration mechanisms in societies as having been achieved in progressive phases where legal and moral institutions were structured by underlying organizational principles. Habermas slightly diverges from the six stages of Kohlberg's schema by proposing a schema of neolithic societies, archaic civilizations, developed civilizations, and early modern societies. Neolithic societies organized interaction via kinship and mythical worldviews. They also resolved conflicts via feuds appealing to an authority to mediate disputes in a pre-conventional way to restore the status quo. Archaic civilizations organized interaction via hierarchies beyond kinship and tailored mythical worldviews backing such hierarchies. Conflicts started to be resolved via mediation appealing to an authority relying on more abstract ideas of justice—punishment instead of retaliation, assessment of intentions, and so forth. Developed civilizations still organized interaction conventionally, but adopted a rationalized worldview with post-conventional moral elements. This allowed conflicts to be mediated by a type of law that, while rooted in a community's (conventional) moral framework, was separable from the authority administering it. Finally, with early modern societies, we find certain domains of interaction are post-conventionally structured. Moreover, a sharper divide between morality and legality emerges such that conflicts can be legally regulated without presupposing shared morality or

needing to rely on the cohering force of mythical worldviews backing hierarchies (McCarthy 1978, 252).

Obviously, this sketch is rather vague and needs further elaboration. This is especially true in light of the ways a superficial reading (that takes social evolution as strictly parallel rather than homologous to individual development) lends itself to unsavory developmentalist narratives. Yet, apart from a few later writings, Habermas has not returned to his theory of social evolution in a systematic way. Several secondary authors have tried to fill in the details (Rockmore 1989, Owen 2002, Brunkhorst 2014, Rapic 2014). Nevertheless, Habermas still endorses the contours of his theory of social evolution: these ideas show up in *Theory of Communicative Action* and his later writings on the nature and development of legality and democratic legitimacy bear a loose connection to this early work (especially the final homology above) insofar as they are tailored for specifically post-conventional societies. Yet, before turning to his democratic theory, we must tackle the hugely important intervening body of work concerning his communicative turn and its articulation in his *Theory of Communicative Action*.

3. The Linguistic Turn into the Theory of Communicative Action

Habermas' engagement with speech act theory and hermeneutics in the late 1960s and 70s started a linguistic turn that came to full fruition in *Theory of Communicative Action*. This turn makes sense after both *Knowledge and Human Interests* and *Communication and the Evolution of Society*. He came to see the knowledge-constitutive interests of the former as illicitly relying on assumptions in the philosophy of consciousness and Kantian transcendentalism, while the reconstructed phases of social learning and evolution in the latter can seem far too naturalistic or foundationalist. In contrast, a focus on communicative structures let him form his own pragmatic theory of meaning, rationality, and social integration based in reconstructions of the competencies and normative presuppositions underlying communication. This approach is transcendental and naturalistic but only weakly so. Far from an account of ultimate foundations, his approach takes itself to be a post-metaphysical methodology for philosophical and social

scientific research into practical reason. From the start of his linguistic turn until well after *Theory of Communicative Action* this approach underwent revisions. In what follows, only a broad outline of this trajectory is given.

Habermas has cited his 1971 Gauss lectures at Princeton (German publication 1984b, English publication 2001) as the first clear expression of the linguistic turn, but it was also evident in *On the Logic of the Social Sciences* (German 1967, English 1988a). His first truly systematic foray in Anglo-American philosophy of language came with *What is Universal Pragmatics?* (German 1976b, included in English 1979). His ideas were then revised further in *Theory of Communicative Action*. While the development of his ideas throughout this period is an important exegetical task, for present purposes the broad way he takes up speech act theory is what is important: he accepts the division in linguistics between syntax, semantics, and pragmatics. He considers each division to be reconstructing the tacit system of rules used by competent speakers to recognize the well-formed-ness (syntax), meaningfulness (semantics), and success (pragmatics) of speech. His main interpretive twist is that the theories of truth-conditional propositional meaning often associated with philosophical projects regarding language only locate *part* of the meaning of speech. Thus, he moves away from meaning based on the correspondence theory of truth and gives an account of the unique pragmatic validity behind the meaning of speech.

While his linguistic turn is sometimes cast as a break with prior theory, his interpretive approach actually coheres quite well with his early critique of positivism. He has always rejected the idea that language simply states things about the world. Instead of *merely* analyzing propositions that either do (true) or do not (false) obtain in the world, he is interested in the full range of ways people *use* language. He claims that, instead of focusing on sentences, a complete theory of language would focus on contextual *utterances* as the most basic unit of meaning. Thus, he developed a formal pragmatics (called “universal pragmatics” in early work). Building on the work of Karl Bühler, he conceives of the pragmatic use of language in context as embedding sentences in relations between speaker, hearer, and the world. This embedding helps to

intersubjectively stabilize such relations. Habermas claims that, in uttering a speech act speakers *mean* something (express subjective intentions), *do* something (interact with or appeal to a hearer) and *say* something (cognitively represent the world). While truth-conditional theories of meaning focus on cognitive representations of the world, Habermas prioritizes the *pragmatics* of speech acts over the semantic or syntactical analysis of sentences. What is *done* through speech is taken to be what is most basic for meaning.

During his linguistic turn Habermas appropriated several ideas from John Searle. Even though Searle has not always fully agreed with such appropriations, two of them are useful points of orientation (Searle 2010, 62). Habermas adopts Searle's idea of the constitutive rules underlying language: just like the rules of a game define what counts as a legitimate move or status, so too there is an implicit rule-governed structure to the use of language by competent speakers. He also adopts Searle's view, built on JL Austin's work, that speech has a double structure of both propositional content and illocutionary force. For instance, the propositional content of "it is snowy in Chicago" is a representation of the world. But the same content can be used in different illocutionary modes: as a warning to drive carefully, as a plea to delay travel, as a question or answer in a larger conversation, and so on. Moreover, beyond such illocutionary force, all speech acts also have derivative perlocutionary effects that, unlike illocution, are *not internally* connected to the meaning of what is said. A warning about snow may elicit annoyance or gratitude, but such responses are contextually inferred and not necessarily connected to either the propositional content or the warning itself.

These ideas about the structure of speech highlight a few key points. First, Habermas takes perlocutionary success (for example, eliciting gratitude) to be parasitic on illocutionary force (for example, the speech is perceived as a warning, not a plea). Attaining success with others by realizing one's intention in the world is *secondary* to achieving an understanding with them. For example, even when lying, the lie only works by first coming to a false understanding that what is being said is true. Second, he identifies three modes of communication—cognitive, interactive, and expressive—that depend on whether a speaker's main illocutionary intention is to

raise a truth claim of propositional content, a claim of rightness for an act, or a claim of sincerity about psychological states. Third, he identifies corresponding speech act types—constatives, regulatives, and expressives—that, seen from the perspective of a competent language user, contain immanent obligations to redeem the aforementioned claims by respectively providing grounds, articulating justifications, or proving sincerity and trustworthiness. In short, Habermas thinks there are general presuppositions of communicative competence and possible understanding that underlie speech and which require speakers to take responsibility for the “fit” between an utterance and inner, outer, and social worlds. For any speech act oriented towards mutual understanding, there is a presumed fit of *sincerity* to the speaker’s inner world, *truth* to the outer world, and *rightness* to what is inter-subjectively done in the social world. Naturally, these presumptions are defeasible. Yet, the point is that speakers who want to reach an agreement have to *presuppose* sincerity, truth and rightness so as to be able to mutually accept something as a fact, valid norm, or subjectively held experience.

For Habermas these elements form the “validity basis of speech.” He claims that, by uttering a speech act, a speaker is seen as also potentially raising three “validity claims”: sincerity for what is expressed, rightness for what is done, and truth for what is said or presupposed. Depending on the speech act type, one claim often predominates (for example, constatives raise a validity claim of truth) and, more often than not, speech rests on undisturbed background agreements about facts, norms, and experiences. Moreover, minor disagreement can be quickly resolved through clarifying meaning, reminding others of facts, asking about preexisting commitments, highlighting situational features, and so on. Habermas sometimes refers to such minor communicative repairs as “everyday speech.” But when disagreement persists we may need to transition to what Habermas calls “discourse”: a particular mode of communication in which a hearer asks for reasons that would back up a speaker’s validity claim. In discourse the validity claims that are always immanent within speech become explicit.

Clearly, Habermas uses “validity” in an odd way. The notion of validity is most often used in formal logic where it refers to the preservation of truth when inferentially moving from one proposition to another in an argument. This is not how Habermas uses the term. What then does he mean by validity? It is instructive to look at the assumptions behind his theory of meaning. When his model of meaning emphasizes what language *does* over what it merely *says* or *means* the operative assumption is that the *primary function* of speech is to arrive at mutual understandings enabling conflict-free interaction. Moreover, at least with respect to claims of truth and rightness, he assumes genuine and stable understandings arise out of the give and take of *reasons*. Claims of truth and rightness are paradigmatically *cognitive* in that they admit of justification through reasons offered in discourse. What Habermas means by validity then is a close structural relationship between the give and take of reasons and either achieving an understanding or (more strongly) a consensus that allows for conflict-free interaction. This yields an “acceptability theory” of meaning where the acceptance of norms is always open to further debate and refinement through better reasons.

As we cannot know in advance what reasons will bear on a given issue, only robust and open discourses license us to take the (provisional) consensuses we do achieve as valid. Habermas therefore formulates formal and counterfactual conditions—the “pragmatic presuppositions” of speech and the “ideal speech situation”—that describe and set standards for the type of reasoning that mutual understandings must pass through before we can regard them as valid (on these formal conditions and how understanding and consensus may differ see below and section 4). At the same time, we never start this give-and-take of reasons from scratch. People are born into cultures operating on background understandings that are embodied in inherited norms of action. Borrowing from Husserl and others, Habermas calls this stock of understandings the “lifeworld.”

The lifeworld is an important if somewhat slippery idea in Habermas’ work. One way to understand his particular interpretation of it is through the lens of his debate with Gadamer. Broadly speaking, Habermas agrees with the view of language held

by Gadamer and hermeneutics generally: language is not simply a tool to convey information, its most basic form is dialogic *use* in context, and it has an inbuilt aim of understanding. On such a view, objectivity is not just correspondence to an independent world but instead something that is ascribed to mutual understandings (about the world, relations to others, and oneself) intersubjectively achieved in communication. Moreover, communication has an underlying structure that makes understandings possible in the first place. Meaning is therefore in some sense parasitic on this background structure.

On this much Gadamer and Habermas agree. But Gadamer takes all this to mean that explicit understanding and misunderstanding are only possible due to a taken-for-granted understanding of cultural belonging and socialization into a natural language. Habermas agrees that culture and socialization are important, but is worried that Gadamer's take on the background structures that form the "conditions of possibility" for meaning yields a relativistic "absolutization of tradition." On Habermas' interpretation the lifeworld encompasses the sort of belonging and socialization referred to by Gadamer, but it works with and is underpinned by certain deep structures of communication itself. For Habermas, the *complementarity* between the lifeworld and a particular manifestation of these deep structures in discourse and "communicative action" (below) is what lets one interrogate and progressively revise parts of the background stock of inherited understandings and validity claims, thereby avoiding either relativism or the dogmatic veneration of tradition.

For Habermas the lifeworld is a reservoir of taken-for-granted practices, roles, social meanings, and norms that constitutes a shared horizon of understanding and possible interactions. The lifeworld is a largely implicit "know-how" that is holistically structured and unavailable (in its entirety) to conscious reflective control. We pick it up by being socialized into the shared meaning patterns and personality structures made available by the social institutions of our culture: kinship, education, religion, civil society, and so on. The lifeworld sets out norms that structure our daily interactions. We don't usually talk about the norms we use to

regulate our behavior. We simply assume they stand on good reasons and deploy them intuitively.

But what if someone willfully breaks or explicitly rejects a norm? This calls for discourse to explain and repair the breach or alter the norm. As a micro-level example: if someone breaks a promise then they will be asked to justify their behavior with good reasons or apologize. Such communication is also called for when norms suffer more serious breakdowns: one may question the reasons behind norms and whether they remain valid, or run into a new and complex situation where it is unclear which norms, how, to what extent, and if they apply. Regardless of how serious the norm breach or breakdown is, we need to engage in discourse to repair, refine, and replenish shared norms that let us avoid conflict, stabilize expectations, and harmonize interests. Discourse is *the* legitimate modern mechanism to repair the lifeworld; it embodies what Habermas calls “communicative action.”

Communicative action can be seen as a practical attitude or way of engaging others that is highly consensual and that fully embodies the inbuilt aim of speech: reaching a mutual understanding. In later writings Habermas distinguishes weak and strong communicative action. The weak form is an exchange of reasons aimed at mutual understanding. The strong form is a practical attitude of engagement seeking fairly robust cooperation based in consensus about the substantive content of a shared enterprise. This allows solidarity to flourish. In either form, communicative action is distinct from “strategic action,” wherein socially interacting people aim to realize their own individual goals by using others like tools or instruments (indeed, he calls this type of action “instrumental” when it is solitary or non-social). A key difference between strategic and communicative action is that strategic actors have a fixed, non-negotiable objective in mind when entering dialogue. The point of their engagement is to appeal, induce, cajole, or compel others into complying with what they think it takes to bring their objective about. In contrast, communicatively acting parties seek a mutual understanding that can serve as the basis for cooperation. In principle, this involves openness to an altered understanding of one’s interests and aims in the face of better reasons and arguments.

The contrast between communicative and strategic *action* is tightly linked to the distinction between communicative and purposive *rationality*. Purposive rationality is when an actor adopts an orientation to the world focused on cognitive knowledge about it, and uses that knowledge to realize goals in the world. As noted, it has social (strategic) and non-social (instrumental) variants. Communicative rationality is when actors also account for their relation to one another within the norm-guided social world they inhabit, and try to coordinate action in a conflict free manner. On this model of rationality, actors not only care about their own goals or following the relevant norms others do, but also challenging and revising them on the basis of new and better reasons.

Approaching rationality after action orientations is not merely stylistic. Habermas notes that while many theorists start with rationality and then analyze action, the view of action that such an order of analysis primes us to accept can tacitly smuggle in quasi-ontological connotations about the possible relations actors can have amongst themselves and to the world. Indeed, this mistake figures into Habermas' critique of Weber's account of the progressive social rationalization ushered in by modernity. Weber framed Western rationalism in terms of "mastery of the world" and then naturally assumed the rationalization of society simply meant increased purposive rationality. As is apparent from Habermas' account of social learning, this is not the only way to understand the "evolution" of societies or the species as a whole throughout history. By expanding rationality beyond purposive rationality Habermas is able to resist the Weberian conclusion that had been attractive to Horkheimer and Adorno: that modernity's increasing "rationalization" yielded a world devoid of meaning, people focused on control for their own individual ends, and that the spread of enlightenment rationality went conceptually hand-and-glove with domination. Habermas feels the notion of rationality in his *Theory of Communicative Action* resists such critiques.

The contrast between communicative and strategic action mainly concerns *how* an action is pursued. Indeed, while these action orientations are mutually exclusive when seen from an actor's perspective, the same goal can often be approached in either communicative or strategic ways. For instance, in my rural town I

may have a discussion with neighbors whereby we determine we share an interest in having snow cleared from our road, and that the best way to do this is by taking turns clearing it. This could count as an instance of communicative action. But, imagine a wealthy and powerful recluse who is indifferent to his neighbors. He could just pay a snowplow to clear the road up until his driveway. He could also use his power to manipulate or threaten others to clear the snow for him (for example, he could call the mayor and hint he may withhold a campaign donation if the snow is not cleared). Strategic action is about eliciting, inducing, or compelling behavior by others to realize one's individual goals. This differs from communicative action, which is rooted in the give-and-take of reasons and the "unforced force" of the best argument justifying an action norm.

Strategic action and purposive rationality are not always undesirable. There are many social domains where they are useful and expected. Indeed, they are often needed because communicative action is very demanding and modern societies are so complex that meeting these demands all the time is impossible. Speakers engaged in communicative action must offer justifications to achieve a sincerely held agreement that their goals and the cooperation to achieve them are seen as good, right, and true (see section 4). But, in complex and pluralistic modern societies, such demands are often unrealistic. Modern social contexts often lack opportunities for highly consensual discussion. This is why Habermas thinks weak communicative action is likely sufficient for low stakes domains where not all three types of validity claims predominate, and why strategic interaction is well-suited for other domains. For Habermas, modern societies require systematically structured social domains that relax communicative demands yet still achieve a modicum of societal integration.

Habermas takes the institutional apparatus of the administrative state and the capitalist market to be paradigmatic examples of social integration via "systems" rather than through the lifeworld. For example, if a state bureaucracy administers a benefit or service it takes itself to be enacting prior decisions of the political realm. As such, open-ended dialogue with a claimant makes no sense: someone either does or does not qualify; a law either does or does not apply. Similarly, in a clearly defined and regulated market

actors know where market boundaries lie and that everyone within the market is strategically engaged. Each market actor seeks individual benefit. It makes little sense to attempt an open-ended dialogue in a context where one supposes all others are acting strategically for profit. Both domains coordinate action, but not through robustly cooperative and consensual communication that yields solidarity. Certainly, not all large-scale and institutionalized interaction is strategic. Some social domains like scientific collaboration or democratic politics institutionalize reflexive processes of communicative action (see section 5 on democratic theory). In such fora cooperation may yield solidarity across the enterprise. Even so, the systems integration like that found in bureaucracies or markets sharply differs from integration through communicative action.

It should be stressed that these are simply paradigmatic examples, and that the same social domain can be institutionalized differently across societies. It is therefore more useful to look at the coordinative media that are typically used to interact with and steer any given institutionalized system rather than positing a fictive typology of clear social domains wherein it is assumed that either strategic or communicative action takes place. Habermas identifies three such media: speech, money, and power. Speech is the medium by which understanding is achieved in communicative action, while money and power are non-communicative media that coordinate action in realms like state bureaucracies or markets. A medium may largely be used in one social domain but that doesn't mean it has no role in others. While speech is certainly the main medium of healthy democratic politics, this doesn't mean money and power never play a role.

This all might seem to imply that there is no single correct way for system and lifeworld to jointly achieve social integration. Indeed, the complementarity between system and lifeworld laid out in *Theory of Communicative action* is broad enough to accommodate a wide range of institutional pluralism with respect to the structure of markets, bureaucracies, politics, scientific collaboration, and so on. But, the claim that there is no "one size fits all template" for social integration should not be taken as the claim that system and lifeworld have no proper relationship. Socialization into a lifeworld

precedes social integration via systems. This is true historically and at the individual level.

Moreover, Habermas claims the lifeworld has *conceptual* priority with respect to systems integration. His thinking runs as follows: the lifeworld is the codified (yet revisable) stock of mutual normative understandings available to any person for consensually regulating social interaction; it is the reservoir of communicative action. Systems integration represents carefully circumscribed realms of instrumental and strategic action wherein we are released from the full demands of communicative action. Yet the very definition and limitation of these realms always depends on communicative action regarding, for example, the types of markets or state administration a community wants to have and why. Without being rooted in the mutual understandings of the lifeworld, we would get untrammelled systems of money and power disconnected from the intersubjectively vouchsafed practical reason that Habermas thinks underpins all meaning. The organizing principles of systems themselves would stop being coherent. For instance, market competition makes sense against a backdrop of normative principles like fairness, equal opportunity to compete, rules against capitalizing on secret information, and so on. But if markets were so “no-holds-barred” that these principles no longer applied, then engaging in market activity would cease to make sense. Similarly, if markets were so regulated that there was no genuine risk or opportunity they would also start to lose coherence as an enterprise. In both these skewed hypothetical scenarios the system is rigged and thus, if there are functional alternatives, it is not worth participating in. This is a variant of his early anti-technocracy argument. Positing “objective necessities” like economic growth, social stability, national security and then circumventing communicative action veils disagreement on *what type* of growth, stability, and security is important for a given community and why. As such, systems designed to achieve these ends are primed to lose coherence and legitimacy based in widely accepted structuring principles.

Habermas thinks the lifeworld self-replenishes through communicative action: if we come to reject inherited mutual understandings embedded in our normative practices, we can use

communicative action to revise those norms or make new ones. Mechanisms of systems integration depend on this lifeworld backdrop for their coherence as enterprises achieving a modicum of social integration. The trouble is that systems have their own self-perpetuating logic that, if unchecked, will “colonize” and destroy the lifeworld. This is a main thesis in *Theory of Communicative Action*: strategic action embodied in domains of systems integration must be balanced by communicative action embodied in reflexive institutions of communicative action such as democratic politics. If a society fails to strike this balance, then systems integration will slowly encroach on the lifeworld, absorb its functions, and paint itself as necessary, immutable, and beyond human control. Current market and state structures will take on a veneer of being natural or inevitable, and those they govern will no longer have the shared normative resources with which they could arrive at mutual understandings about how they collectively want their institutions to look like. According to Habermas, this will lead to a variety of “social pathologies” at the micro level: anomie, alienation, lack of social bonds, an inability to take responsibility, and social instability.

In *Theory of Communicative Action* Habermas pins his hopes for resisting the colonization of the lifeworld on appeals to invigorate and support new social movements at the grassroots level, as they can directly draw upon the normative resources of lifeworld. This model of democratic politics essentially urges groups of engaged democratic citizens to shore up the boundaries of the public sphere and civil society against encroaching domains of systems integration such as the market and administrative state. This is why his early political theory is often called a “siege model” of democratic politics. As section 5 will show, this model was heavily revised in *Between Facts and Norms*. Before turning to that work, we must flesh out discourse ethics—an idea that figured into *Theory and Communicative Action* but which was only fully developed later.

4. Discourse Ethics

Habermas’s moral theory is called discourse ethics. It is designed for contemporary societies where moral agents encounter pluralistic

notions of the good and try to act on the basis of publically justifiable principles. This theory first received explicit and independent articulation in *Moral Consciousness and Communicative Action* (German 1983, English 1990a) and *Justification and Application* (German 1991a, English 1993), but it was anticipated by and depends on ideas in *Theory of Communicative Action*. The overview that follows draws upon these works. Much like the prior section, it only traces the broad outline of discourse ethics.

Discourse ethics applies the framework of a pragmatic theory of meaning and communicative rationality to the moral realm in order to show how moral norms are justified in contemporary societies. It could be seen as a theory that uncovers what we pragmatically *do* when we make and defend the moral validity claims underlying and manifested in our norms. Yet, we need to be careful with this characterization. Because of its cognitive commitments to *moral learning* and *knowledge* discourse ethics cannot simply be a reconstructive description of how it is we practically avoid conflicts and stabilize expectations in post-conventional social contexts. It is also an attempt to provide a formal procedure for determining which norms are *in fact* morally right, wrong, and permissible. Discourse ethics is squarely situated in the tradition of Neo-Kantian deontology in that it takes the rightness and wrongness of obligations and actions to be universal and absolute. On such a view, the same moral norms apply to all agents equally. They strictly bind one to performing certain actions, prohibit others, and define the boundaries of permissibility. There is no “relative” validity of genuinely *moral* norms even though, as we shall see, they can be embedded in social contexts that have consequences for their application. As long as these caveats are kept in mind we can understand discourse ethics by analyzing the practice of making and defending validity claims and how there are certain conditions of possibility tacitly underpinning and enabling this practice.

What are the conditions that enable this practice? As touched on above, Habermas posits certain unavoidable pragmatic presuppositions of speech which, when realized in discourse, can approximate a counterfactual ideal speech situation to greater or lesser degrees (1971; MCCA, 86). Discourse participants need to

presuppose these conditions in order for the practice of discursive justification to make sense and for arguments to be truly persuasive. Four of these presuppositions are identified as the most important: (i.) no one who could make a relevant contribution is excluded, (ii.) participants have equal chances to make a contribution, (iii.) participants sincerely mean what they say, and (iv.) assent or dissent is motivated by the strength of reasons and their ability to persuade through discursive argumentation rather than through coercion, inducement, and so on (BNR, 82; TIO, 44). The point is not that actual discourses ever realize these conditions—this is why the ideal speech situation is best understood as a counterfactual regulative ideal. Rather, the point is that the outcomes of any discourses are only reasonably taken to be “valid” (empirically true, morally right, and so forth) under the presumption that these conditions have been sufficiently met. As soon as a violation is discovered this casts doubt upon the validity of the discursive outcome.

In addition to these pragmatic presuppositions Habermas proposes his discourse principle (D). This principle is supposed to capture the type of impartial, discursive justification of practical norms required in post-conventional societies: “only those action norms are valid to which all possibly affected persons could agree as participants in rational discourse”. (BFN 107; TIO 41) While (D) was initially framed as a principle for moral discourses it was soon revised to the more general form above, as there are many practical norms concerning interpersonal interaction that are not directly moral even if they must be compatible with morality. Yet even in its broadened form it is crucial to note that (D) only applies to discourses concerning *practical norms* about interpersonal behavioral expectations, not all discourses about theoretical, aesthetic, or therapeutic concerns (which may or may not involve interpersonal social interaction). The guiding thought is that if discourses about an action norm are carried out in a sufficiently ideal manner and they yield consensus then this is a good indication the norm is valid. The principle *does not* hold that consensus reached through discourse *constitutes* validity, nor that whatever norm people coalesce around after discourse that looks sufficiently ideal is assured to be valid. Rather, (D) simply holds that consensus

about a norm can be a good test of validity if it has been achieved in the right type of discursive way. It is important to note that, because of its very broad scope, (D) mainly functions by pointing out *invalid* norms. By itself the discourse principle cannot tell us which norms are valid. It can only help us identify norms that are good *candidates* for validity.

Moreover, before the validity of an action norm can be assessed, we need more details on the types of discourse and validity claims at issue (TIO 42). Within his project of discourse ethics Habermas identifies moral, ethical, and pragmatic discourses (JA 1-17; MCCA 98). Each type deploys practical reason differently, framing and analyzing questions under the rubrics of the purposive (practical), the good (ethical), or the just (moral). The language of differing discourse “types” should not be taken to mean that norms come prepackaged in distinct kinds. Instead, any norm can be discursively thematized in any of these ways and should not be arbitrarily limited to a given type. With that caution in mind, we can begin to understand discourse types and the norms they produce.

Ethical discourses are a good place to start. For, while they are constrained by the outcomes of moral discourses and therefore not foundational, our prior discussion of the lifeworld provides an apt segue. Ethical discourses are paradigmatically about clarifying, consciously appropriating, and realizing the identity, history, and self-understanding of a group or individual. They make validity claims to authenticity rather than truth or rightness. They also involve value judgments about a particular social form or practice concerning the good life in a community. This is one reason why the outcomes of *ethical* discourses will have *relative* validity: they are meant to redeem validity claims for actors in some community or another. Another reason is that values differ from the types of generalizable or universalizable interests embodied in moral norms. While moral norms are supposed to strictly oblige agents to either do or not do some action, values admit of degree. While moral norms express principles backed by reasons, values are affective components of meaning acquired in virtue of living in a given social context. They are connected to reasons but not reducible to them. Values can orient us to goals, aid motivation, and help successfully navigate the lifeworld but cannot ground moral obligations by

themselves. Values attract or repel but do not persuade; they can provide motivation to “do the right thing”—to have the will to follow a moral insight—but they do not constitute or even always help us discern what “the right thing” is (BFN 255).

Ethical discourses are rooted in ethicality (*Sittlichkeit*), which is distinct from morality (*Moralität*). Like many philosophers, Habermas separates the realm of the right from the realm of the good. Following a loosely Hegelian terminology, he parses this as the difference between morality and ethicality. Ethicality is a way of life composed of both cognitive and affective elements as well as more structural elements that reproduce this way of life: laws, institutions, conventions, social roles, and so forth. It is particularistic in that it defines goals in terms of what is good for a group as a whole and its members. As Habermas believes in George Herbert Mead’s model of “individuation through socialization,” ethicality is deeply engrained and connected to the lifeworld. No one can simply drop their internalized ethical perspective just as no one can simply step out of the lifeworld they have inherited. Individuals are always in some sense bound up with the identity, practices, and values of their upbringing and traditions even if they come to largely reject them. But, as was clear from Habermas’ critique of Gadamer, ethical perspectives do not determine us. Ethical discourses explain how this is by mediating between inheritance and transcendence. While we inherit and internalize an ethical perspective as individuals, we can always question parts of it that we wish to challenge, refashion, or reject for lack of sufficient reasons underwriting certain norms.

This dialectic between the ethicality we internalize through socialization and the way in which we wish to consciously reappropriate and (dis)own portions of such ethicality helps to explain why, in contrast to other discourse types, Habermas pays a great deal of attention to ethical discourses at *both* the individual and group levels. Ethical discourses at the individual level are called ethical-existential while ethical discourses at the group level are referred to as ethical-political discourse. For example, an individual considering a certain profession would engage in an ethical-existential discourse (for example, is this profession right for me given my character and goals?), while a polity considering whether

certain policies express their collective interest, identity, and values would engage in an ethical-political discourse (for example, does this policy align with the collective identity and commitments we have had and how we want to appropriate them moving forward?).

There are two key points about these levels. First, the outcomes of such discourses are constrained by morality irrespective of what would be authentic at individual or group levels: an individual cannot simply decide to become a serial killer just as a country cannot simply enact a policy that has patently immoral consequences (for example, for those outside it). While Habermas thinks it is important to account for the way in which morality is embedded in social contexts through ethical discourses, he is staunchly opposed to postmodern or communitarian takes on morality and justice. Second, there will often be a reflexive interplay between these two levels of ethical discourse. Discourses about what it means to genuinely inhabit a collective identity can impact the ordering and strength of the values held by individuals, and discourses about who one fundamentally is and wishes to be can, through resistance to dominant interpretations of traditions and highlighting unacknowledged injustices, impact how others in a collectivity appropriate their identity and normative practices moving forward. This interplay is bookended by broader moral discourses at both levels, thereby helping the outcomes of such discourses stay in the realm of permissibility.

Pragmatic discourses are similar to ethical discourses in that they start from the teleological perspective of an agent who already has a goal. But in contrast to the reflexive, clarifying, and potentially transformative self-realization and collective self-determination of ethical discourses, pragmatic discourses simply start with a goal of presumed value and set about realizing it. This goal may involve identity and values but it could also refer to more pedestrian concerns and interests. Because the goal is presumed to be worthwhile the values, interests, or goals at issue show up as relatively static. Pragmatic discourses simply focus on the most efficient way to realize or bring about a goal, and their claim to validity concerns whether or not certain strategies or interventions in the world are likely to produce a desired result. As Habermas puts it, pragmatic discourses correlate “causes to effects in

accordance with value preferences and prior goal determinations” so as to generate a “*relative ought*” that expresses “what one “ought” or “must” do when faced with a particular problem if one wants to realize certain values or goals” (JA 3). The “ought” is relative because it is something akin to a rule of prudence that depends on whether an agent happens to have a certain interest or find a goal worth pursuing.

Finally, we turn to what might be seen as the most important type of discourse: moral discourses. Moral discourses are broader in scope and establish stronger validity claims than either ethical or pragmatic discourses. They seek to discern and justify norms that bind *universally* rather than simply in the confines of a specific community or because an agent happens to find a goal valuable. These norms have binarily coded, unconditional validity instead of the gradated, relative validity of the outcomes produced by pragmatic and ethical discourses.

In order to discursively discern this non-relative sense of moral validity Habermas proposes a separate principle, his principle of universalization (U), for discourses about moral norms: “A norm is valid when the foreseeable consequences and side effects of its general observance for the interests and value orientations of each individual could be jointly accepted by all concerned without coercion” (TIO 42). While (U) has gone through several different formulations, the basic idea is that for whatever valid moral norms there are, such norms can be accepted by all affected persons in a sufficiently ideal discourse wherein they assert their own interests and values. (U) checks if the norms we take to be moral actually are in virtue of whether or not they are universalizable. If they are not universalizable, they cannot be moral norms. Beyond this basic characterization there are some interpretive issues with (U). Three are worth brief focus: its apparent reference to consequences, where (U) comes from, and the role of interests.

First, in the version of (U) above, it is easy to mistake the “foreseeable consequences and side effects” clause with the addition of a mild consequentialist constraint. Given Habermas’ deontological commitments this would be odd. Instead, the clause builds in a “time and knowledge index” so that it does not make impossible demands on moral agents. Fully satisfying (U) would

require discourse participants who had unlimited time, complete knowledge, and no illusions about their own interests and values; it would require participants who transcended their human condition. As (U) must be usable in the real world it can only ask that moral discourse participants attempt to account for the “anticipated typical situations” to which a norm would apply when they attempt to justify any moral norm (JA 37). The circumscribed task of (U) is key: it is only supposed to *justify* moral norms in the abstract. While this justification may point towards “typical” cases of application, it does not predetermine all applications. What about novel, atypical, or completely unforeseen situations to which the norm might unexpectedly apply?

Following Klaus Günther, Habermas claims that moral (and legal) decisions in specific cases require a logic of appropriateness found in discourses of application (Günther 1993; JA 35-37). Discourses of application look at a concrete case and survey all potentially applicable norms, relevant facts, and circumstances. They try to offer exhaustive or “complete” descriptions of a situation so as to decide among multiple, sometimes competing or only partly applicable norms that might regulate a situation. There is a division of labor between the two types of recursively related discourse: whereas discourses of justification lay out the reasons why we should endorse a norm as a general rule with reference to typical situations, discourses of application seek to apply norms to concrete cases which may be wholly new or defy expectations. As fallible agents we can make a variety of different errors in our discursive justification of a norm or fail to anticipate new situations or altered understandings of facts, values, and interests—a failure that would be revealed in application. Habermas calls this the “dual fallibilist proviso,” and it instills an awareness that moral justification is an ongoing project (TJ 259). The recursive interplay of justification and application is supposed to progressively address prior errors and oversights. New insights gleaned from application discourses or novel situations can lead us to revisit norms whose justification was taken for granted, and this refinement of our understanding regarding how and why norms are justified will help us apply them better. If we had providential foreknowledge we would not need application discourses. But since we are fallible the “foreseeable

consequences and side effects” should be seen as referring to an in-built “time and knowledge” index for the outcomes of justificatory discourses, which are then supplemented by application discourses that may impact the formulation of the initial norm.

The second interpretive issue is where (U) comes from. Habermas initially claimed that (U) could be formally deduced from a combination of the pragmatic presuppositions of discourse and (D), but weakened this claim shortly thereafter (JA 32 n17). Instead of deriving (U) from a formal deduction or informal inferences he now claims—using a term coined by Peirce—we arrive at (U) “abductively” (TIO 42). To arrive at something abductively is to suggest that we first observe a phenomenon (moral norms) and adopt a “best guess” hypothesis to explain it (the moral principle), which can then be subjected to further inductive testing (Ingram 2010, 47; Finlayson 2000a, 19). In short, (U) is now proposed as the best candidate principle for helping to explain moral normativity. To buttress the plausibility of this claim Habermas has also fallen back on his theory of social evolution and the “weak...notion of normative justification” in post-conventional contexts (TIO 45). Indeed, he now often speaks about (U) as following from the type of impartial justificatory procedure appropriate to a post-conventional condition that seeks to discern norms that are “equally in everyone’s interest,” “generalizable,” or “universalizable” (RPT 367; BFN 108, 460; TJ 265). The reference to interests leads us to the third interpretive issue with (U).

Early formulations of (U) only refer to interests (MCCA 65, 120). The inclusion of value orientations is potentially confusing. As noted above values are not necessarily cognitively grounded. As Habermas has always presented his moral theory as cognitivist it would be odd to give values such a central role. It seemed to make sense that initial formulations of (U) only included interests, as Habermas has defined interests in a cognitive fashion (on interests as “reasons to want” see Finlayson, 2000b). Bolstering an interpretation of (U) that puts priority on (cognitive) interests he has stated that “(U) works like a rule that eliminates as non-generalizable content all those concrete value orientations with which particular biographies or forms of life are permeated” (MCCA 121), and that the specific part of (U) referring to “uncoerced joint

acceptance” means that any reasons put forth in moral discourse must “cast off their agent-relative meaning and take on an epistemic meaning from the standpoint of symmetrical considerations” (TIO 43). Moreover, the interpretive secondary literature has often emphasized the centrality of interests over values and focused on how Habermas often talks about “generalizable” or “universalizable” interests as the distinctive feature that moral norms secure (Heath 2003; Finlayson 2000b; Lafont 1999). How then should the inclusion of value orientations be understood?

Habermas has said he included value orientations in (U) so as to “prevent the marginalization of the self-understanding and worldviews of particular individuals and groups” (TIO 42). This does not mean that values are on a par with interests. Instead, his point is that interests and values are always bound together. Value orientations exert at least some indirect influence on moral discourses insofar as they subtly influence the very interpretation of our own interests (JA 90). Proceeding as if value orientations can be expunged from moral discourses may in fact introduce discursive blind spots. Indeed, candor about one’s own value-orientations may be crucial since the impartiality of (U) involves “generalized reciprocal perspective-taking” that cuts both ways: it orients participants towards “empathy for the self-understandings” of others as well as towards “interpretive interventions into the self-understanding of participants who must be willing to revise their descriptions of themselves and others” (TIO 43). The essential point is that even though “some of our needs are deeply rooted in our anthropology” and can be seen as basic generalizable interests shared by all, we must nevertheless avoid “ontologizing generalizable interests” into “some kind of given” because even “the interpretation of needs and wants must take place in terms of a public language” wherein our own self-understandings are open to revision (TJ 268; JA 90).

A final interpretive issue that merits attention is the precise status of moral rightness. Habermas has always held that morality and truth are *analogous* in that both are cognitive, binarily coded, and subject to learning processes. Moreover, he has always been sharply critical of approaches that would reduce morality to a purely subjective or relativized affair. Yet, given that rightness is not

reducible to truth and that Habermas has repeatedly disclaimed a moral realist reading of his theory, it is unclear precisely how far this analogy is supposed to extend. This is not only because there are a variety of differences between empirical and moral knowledge but also because Habermas has changed his theory of truth over the years—moving from a consensus theory that identified truth with ideal warranted assertability to a “pragmatic epistemological realism that follows in the path of linguistic Kantianism” (TJ 7). Early articulations of discourse ethics seemed to admit of interpretations wherein rightness was a justification-transcendent concept that couldn’t be captured by ideal warranted assertability. This led some interpreters to interpret Habermas’ moral theory as at least tacitly committed to some variant of internal moral realism (Davis 1994, Kitchen 1997, Lafont 1999 and 2012, Smith 2006, Peterson 2010 ms.). But, in the course of resisting this reading, Habermas has explicitly claimed that, “ideally warranted assertability is what we mean by moral validity” (TJ 258, 248). He now wishes to articulate a notion of moral rightness that can be cashed out in terms of a pragmatist constructivism that also avoids the perils of relativism and skepticism—that is, which maintains an anti-realist account of moral rightness that still resists collapsing into a form of moral consensus theory. Whether he succeeds in this endeavor is a hotly debated topic.

5. Political and Legal Theory

In post-conventional, pluralistic societies ever fewer norms can be underwritten by a shared ethos embodied in a community’s ethicality or collective identity. Moral norms cannot pick up the slack to achieve social integration and cohesion by themselves. Because moral discourse is demanding and aims at what is equally in everyone’s interest, few moral norms will be seen as justified across the world or even in a given society (JA 91, TJ 265). And, as Habermas noted in *Theory of Communicative Action*, while systems like the bureaucratic state and economy can achieve stability and coordinate expectations through money and power, this can erode mutual understandings and social solidarity; markets and bureaucracies tend to displace and colonize the lifeworld. Indeed, his political essays from this period cast democratically

created law as holding the line against system encroachments in a siege mentality (BFN 486-89, Habermas 1992b 444). This may leave us asking: What other resources exist for legitimate social integration?

In Habermas' clearest statement of political theory, *Between Facts and Norms*, modern law shows up as precisely the resource we are looking for. If law is linked to democratic political structures in the right way it confers legitimacy on legal norms, thereby fostering social integration and stability. Broadly speaking, the relation between legal legitimacy, procedural-democratic popular sovereignty, and public discourse is nested and reflexive: legitimate law must be rooted in democracy, which itself depends upon a robust public sphere. A vibrant democratic public sphere is what allows for the revision and questioning of prior law. Conceived of in this way modern law is a "transformer" that preserves the normative achievements and mutual understandings that issue from the collective self-determination of the public sphere by translating them into legitimate, binding decisions that can "counter-steer" against the logics of the state and market. As long as legal decisions are arrived at in the right type of procedural, discursive fashion there is a presumption in favor of their rationality and legitimacy. And, as long as the public sphere continues to be a robust and open forum of contestation, any prior decisions are revisable such that there is a circulation between the informal public sphere and more formal institutions of the state. This focus on the transformative, mediating nature of law revises the prior "siege" model of democratic law into a procedural "sluice" model (Habermas 2002, 243). While the prior model saw democratically generated law as a defensive dam or shield against the demands of systems, the new model sees a certain type of lawmaking as mediating the circulation between lifeworld and system in a way that produces legitimate and binding legal norms. Modern law works with systems and alongside post-conventional morality to stabilize social expectations and resolve conflicts.

We can start to understand the relation between law, democracy, and the public sphere by focusing on legal legitimacy and democracy. *Between Facts and Norms* posits a tension within law itself, as well as an internal relation between modern law and

democracy. To function, all law must demand compliance, threaten coercion, and (however tacitly) appeal to an underlying normative justification. Law is therefore characterized by a tension between “facticity” and “validity” insofar as it must be recognized as factually efficacious and normatively justified. This tension helps explain the relation between law and democracy in contemporary contexts. Pre-modern law appealed to God, nature, human reason, or shared culture for its justificatory backing. In post-conventional societies the fact that law is coercible and changeable yet merely rooted in fallible humans is laid bare. For Habermas, the underlying normative justification can now only be understood as “a mode of lawmaking that engenders legitimacy” (IO 254). The thought is that democracy is the only mode of lawmaking that is up to this legitimacy-engendering task. In light of these connections it is fruitful for present purposes to focus on “deliberations that end in legislative decision making” rather than treating political and legal legitimacy separately (BFN 171; Bohman and Rehg 1999, 36).

The democracy Habermas has in mind differs from overly populist varieties. He is clear that the legitimacy underwriting lawmaking must be twofold: law must not only express the democratic will of the community but must also be non-subordinately “harmonized” with morality (BFN 99, 106). This non-subordinate concordance of legality and discourse theoretic morality is the hardest sense of legitimacy to explain and the easiest to overlook, so it is fruitful to start there. For Habermas, “legal and moral rules...appear *side by side* as two different but mutually complementary kinds of action norms” in post-conventional societies. In order to also account for “the idea of self-legislation *by citizens*” we must avoid a “subordination of law to morality” along the lines of classical natural law theory (BFN 105-6, 120; IO 257). Yet it seems puzzling to hold that democratically determined law should be compatible with but not subordinate to discourse-theoretic morality. What about cases where law and morality seem to conflict? There are a few answers that highlight unique features in Habermas’ theory. At a general level these answers take the same shape: while there are many ways that legal systems can square with moral permissibility, there are nevertheless structural and conceptual features endogenous to processes of modern procedural-democratic popular

sovereignty that, at least at an abstract level, tend to harmonize legal norms with moral permissibility. This avoids concerns with morality trumping legality in an exogenous manner.

One reason to expect that democratically legitimate law and moral permissibility will be at least in principle commensurable is that they are both rooted in (D). We saw above how the moral principle (U) expresses the way (D) is specified for moral discourses. Habermas also proposes a principle of democratic legitimacy (L) that expresses the way (D) is specified for political discourses producing law. This principle is rooted in (D) in virtue of what Habermas calls the “legal form.” When (D) is deployed in discourses aimed at producing legal norms for regulating common life together it is understood these norms will be cloaked in the legal form: the set of formal and functional features characterizing modern positive law. Modern positive law is enacted and conventional, enforceable and coercive, rooted in institutions with some reflexivity, tailored to protect individuals through rights, and limited in scope (BFN 111-118, IO 256). If law is to function as a tool for the consensual regulation of social conflicts and the integration of society, then it needs to take on this form.

The principle of democratic legitimacy (L) is part of the normative backing that is supposed to emerge, albeit *in nuce* and very abstractly, from the historical interpenetration of (D) and the legal form that has culminated in the structures of modern democratic state. It claims, “only those statutes may claim legitimacy that can meet with the assent of all citizens in a discursive process of legislation that in turn has been legally constituted” (BFN 110; constituted is sometimes translated as organized). This principle captures how (D) is specified for political discourses so that democratic procedures underwrite the legitimacy of legal norms. Legitimacy does not arise out of formal legality alone; it needs the added normative backing of democracy. The idea of (L) is that compliance with the law must be rational and rooted in the law’s perceived legitimacy. To achieve this, political discourses must be structured in a way where formal legislative institutions accurately represent and address deliberations going on in the informal public sphere, and where there are institutionalized procedural mechanisms organized in a way to help screen out weak arguments

(BFN 340). The details of this structuring will be clarified below, particularly in relation to the process model and the relationship between democracy and the public sphere.

However, the mere fact that (U) and (L) are rooted in (D) does little to *ensure* the commensurability of law and discourse-theoretic morality. Fortunately, there are additional reasons why we might expect such a harmonization. Habermas thinks the combination of (D) and the legal form in (L) *also* supplies us with the resources to discern the conceptual kernels of an abstract “system of rights” that will be inscribed in the core structures of any legitimate self-determining political community. The basic argument is that in order for (L) to be realized it must make reference to a concrete community engaged in self-determination through modern law. In such communities equal legal personhood takes on the role of a “protective mask,” a formal identity mainly defined by rights instead of duties, that crystalizes around individual moral persons (BFN 531, 112). This legal identity is constituted by a core of rights that secure the status and private autonomy of individuals such that they can not only live their individual lives but also genuinely deliberate (on equal footing, free from coercion, and so forth) about the terms of shared life together. Yet, these individual rights cannot be effective unless they presuppose other rights to participation and basic material provision—rights that secure public autonomy. The claim is that the legal manifestations of private and public autonomy, often expressed in the idioms of human rights and popular sovereignty, mutually presuppose one another. What results is an abstract system of rights made up of five core types. What are these right types?

First, in order to discursively engage one another people need to be reasonably secure. Therefore, rights that guarantee the status of individual persons are required. Three types of rights jointly achieve such protection: (i.) the right to equal liberties compatible with those of others, (ii.) rights of membership that determine the extent of the community, and (iii.) rights of due process that assure each person is treated the same and equally protected under the law (BFN, 133-134). These rights secure the individual private autonomy prioritized by classical liberalism. But any community engaged in specifically democratic self-determination must *also*

safeguard the ability to *actively use* the freedom afforded by this secure status to deliberate, disagree, and come to mutual understandings in concert with others. If individual rights are to be effectively used (iv.) rights of communication and political participation that formally secure equal opportunity and access to the political process are required. These rights secure the collective public autonomy prioritized by classical republicanism. They enable discourses in the public sphere as well as equal access to channels of political say and influence; they enable democratic popular sovereignty by making sure everyone can participate on fair and equal terms, and that information, innovative ideas and arguments about how to structure common life are kept freely circulating and scrutinized. Lastly, these four right types are insufficient if basic needs are threatened or go unmet. Formal guarantees of freedom and participation mean little if they amount to the freedom to starve. So, as a final step, Habermas proposes some measure of (v.) social, technological, and ecological rights securing the basic conditions of a minimally decent life. Democratic states have often done a poor job fully realizing these rights, but the claim is simply that these general right types are conceptually required if self-determination through law is to achieve the dual sense of legitimacy noted above. In this same spirit of clarification, it is also important to note that the abstract system only identifies certain right *types*, not some list of concrete rights. Communities have incredibly wide interpretive latitude when it comes to how these rights show up. Habermas often refers to rights as “unsaturated placeholders”; it is largely up to communities to “fill in” their content.

The expectation of a non-hierarchical harmonization of morality and legality may now seem less puzzling. Ideally, lawmaking discourses approximate (L) against the backdrop of an abstract system of rights inscribed in the political structures of a democratic community. This places some broad constraints on how deliberations unfold and the type of norms they can produce. Moreover, apart from these structural background constraints political discourses are also themselves unique. In contrast to moral discourses focused on “the interest of all” or ethical discourses focused on authentic self-realization, political discourses aimed at self-determination through law reference a plethora of different

concerns, and do so in an internally-structured way aimed at carving out a space (defined by rights) where moral personhood and ethical authenticity can flourish (BFN 531).

While deliberations about “political questions are normally so complex that they require the simultaneous treatment of pragmatic, ethical, and moral *aspects*” of issues, they ideally unfold along a ‘process model’ where there is a structured interplay between pragmatic, ethical, and moral concerns as well as procedurally regulated bargaining (BFN 565, 168). The basic idea is that for any provisional policy conclusion there is an obligation to respond to objections stemming from more abstract aspects of an issue or levels of discourse; discursive processes cannot be arbitrarily limited. For instance, participants in a discourse on immigration policy cannot simply consult ethical concerns regarding their community’s authentic identity but yet refuse to listen to moral discourses that bear on such policies. Any moral aspects need to be explicitly discussed, and they filter or check more particularistic issue-aspects and discourses (Cf. BFN 169 and the emendation at 565 on whether to refer to the structured interplay as between *discourses* or *aspects* of a case). The abstract system of rights and the process model mean that, within political deliberations about how to structure common life together, it will in principle always be possible for more abstract moral discourses to weakly check pragmatic and ethical-political discourses. And, this checking will be endogenous to structures of democratic self-determination.

So far the focus has been on the relation between law and democracy without much reference to the public sphere. However, it is hard to overstate the importance Habermas places on democratic deliberation rooted in the public sphere. None of the formal or structural mechanisms mentioned so far guarantee that public political discourses or laws will be specified in a given way. There is assurance neither that the abstract system of rights or (L) will be meaningfully realized, nor that the interplay of various types of concerns in political discourses will unfold along the process model. Everything hangs on the quality and institutional structuring of deliberation in the public sphere. Indeed, the primary reason why democracy confers legitimacy upon legislative outcomes is that it is rooted in a model of distinctly procedural popular sovereignty that

simultaneously expresses the will of the community and that leads to more rational outcomes. An analysis of the specific way in which democracy and the public sphere are related on Habermas' model is the best way to understand how the democratic mode of lawmaking underwrites the legitimacy of legal norms.

In *Between Facts and Norms* Habermas proposes a "two-track" model of democratic politics outlining a circulation of political power engendering legitimacy. He divides the political public sphere into informal and formal parts. The informal public sphere includes all the various voluntary associations of civil society: religious and charitable organizations, political associations, the media, and public interest advocacy groups of all varieties (BFN 355). In this sphere public political deliberation is free and unorganized. Through this open clash of views and arguments individuals and collectivities can both persuade and be persuaded, thereby contributing to the emergence of considered public opinions. In contrast, the formal public sphere includes institutionalized forums of discourse and deliberation like congress, parliament, and the judiciary as well as more peripheral administrative and bureaucratic agencies associated with state structures. This sphere is supposed to be organized in such a way that it renders decisions reflecting the considered public opinions of the informal public sphere. Formal institutionalized decision making bodies must be porous to results of the informal public sphere.

The informal public sphere is the key forum for generating a type of normative power that can integrate society through mutual understandings and solidarity rather than through money or administrative-bureaucratic power. When discourse participants in the informal public sphere freely reach mutual understandings about how to regulate the terms of shared life together "communicative power" emerges (Flynn 2004 discusses communicative power's precise locus). Communicative power arises from jointly authored norm expectations that are cognitively grounded in the force of better reasons and motivationally grounded (albeit weakly) in mutual recognition and collective ethical discourses. Cognitively speaking, free communication in the public sphere can foster "rational opinion and will-formation"

because “the free processing of information and reasons, of relevant topics and contributions is meant to ground the presumption that results reached in accordance with correct [discursive] procedure are rational” (BFN 147). This acceptance also provides weak motivation: in accepting a norm’s validity claim one accepts the background understandings and reasons underlying it which can motivate relevant circumstances. Moreover, because this mutual understanding was presumably reached through persuasive discourse where reasoned dissent was (and remains) a real possibility, norm acceptance can also motivate in a spirit of anti-paternalistic empowerment: parties recognize each other as accountable and responsible for their actions in accord with a norm until new counter-reasons are discovered. While they may be aware of counter-inclinations and motives that are not backed by good reasons, they take one another to be competent, responsible agents who can choose to act on rationally backed norms (Günther 1998). Yet, because the motivation accompanying cognitive insight is fragile and weak, communicative power must also be rooted in a community with a shared ethical-political identity and legitimate law so that motivational deficits can be met with supplemental resources of a shared life and law.

Communicative power can only arise if the informal public sphere has certain characteristics. First and foremost, it must be relatively free of distortions, coercion, and silencing social pressures so that communication can work as a filter for fostering more rational individual and collective will formation (BFN 360). The public sphere also needs to accurately function as a “context of discovery” wherein problems that affect large segments of the public are identified and taken up for discussion and resolution in discourse. Moreover, civil society must be animated by a political culture so that members actively participate in voluntary associations and public discourse about the terms of common life together (BFN 371). Normative power potentials cannot be generated if members largely retreat into private concerns or a society is internally segmented and riven with special interests (Flynn (2004) 439-444; Bohman and Rehg (1999) 41-42). Clearly, if the public sphere is to remain healthy then the media’s role in fostering accurate

information and timely mass communication will also be crucial (EFP, 138-183).

The political institutions of the formal public sphere are arranged so as to be porous to the inputs of the informal public sphere, to further refine and focus public opinion, and to make decisions. Building on the work of Bernhard Peters, Habermas maintains that modern constitutional democracies are set up so communication and decision-making flow from the “periphery” of the informal public sphere into the “center” constituted by those formal political institutions that create, enforce, clarify, or implement the law (BFN 354). In a well-functioning democratic regime there will be structural “sluices” or “floodgates” embedded in the institutions of the administrative state (legislature, judiciary, and so forth) so that the circulatory flow of power proceeds in the right direction, from the periphery to the center.

The thought is that the political community should “program” and direct the institutions of the administrative complex, not the other way around (BFN 356). If the state or other powerful actors reverse this flow by simply positing new laws or rules and either demanding compliance or inducing it in some other way, then this exercise of non-communicative administrative-bureaucratic power would be neither legitimate nor stable. Habermas claims the “integrative capacity of democratic citizenship” erodes to the extent that the circulation of political power is interrupted or reversed. Only communicative power has the legitimating force needed so that a community can both author and rationally abide by the law. Democratic lawmaking is the key institution that “represents...the medium for transforming communicative power into administrative power” while preserving its normative potential (BFN 169, 81, 299). Democratically generated law ensures normative power potentials flow in the right direction and that they are maintained when implemented by institutions of the administrative state.

This account of procedural-democratic collective self-determination should not be confused with traditional national self-determination. Habermas rejects models of sovereign collective self-determination that presuppose a nation or people with a homogeneous identity and interests, as well as models where “a network of associations” stands-in for this (imaginary) collective-self (BFN 185, 486).

Instead, in modern constitutional democracies the “idea of popular sovereignty is... desubstantialized [and]...not even embodied in the heads of the associated members.” Popular sovereignty “is found in those subjectless forms of communication that regulate the flow of discursive opinion- and will- formation in such a way that their fallible outcomes have the presumption of practical reason on their side” (BFN 486). Insofar as we can speak about *the* will of a community it is an anonymous and subjectless public opinion emerging out of the discursive structures of communication themselves (BFN 136, 171, 184-186, 299, 301). This unique interpretation of popular sovereignty helps explain some final aspects of Habermas’ political theory: his views on religion and the public sphere, his constitutional patriotism, and his vision of politics beyond the nation-state.

In early writing Habermas claimed that as the rationality and pluralism of enlightenment ideals slowly took hold in modern societies the mythic explanations of religion would be less important. But, he slowly came to revise his view on religion in modern societies. At present, the way he sees religion fitting into the public sphere of a liberal democracy is what is important. In liberal democracies, untrammelled populism is held in check by not only individual rights but also the very nature of public debate: citizens collectively self-determine through persuasion and rational argumentation. To do this amidst the pluralism of modernity, the laws they make must be grounded in public reasons accessible to all. The question is what this means for religious citizens.

There have been a variety of answers. For instance, in *Political Liberalism* [John Rawls](#) held that liberal democratic citizens should ultimately only endorse policies that they can support on the basis of secular reasons. While these citizens may have religious reasons that favor a law or policy, when engaged in political debate they must eventually “translate” these reasons into terms that non-believers could accept. Habermas is sympathetic to the vision of liberal democracy animating this view of how religious citizens should act. Indeed, he criticizes thinkers like Wolterstorff who insist that religious citizens ought to be allowed to try to base coercive law on their own particularistic values and conception of the good. Nevertheless, he feels that placing the burden of “translation” onto

religious citizens alone is somewhat misguided. Such an approach underestimates the ethical-existential importance of religion in some people's lives—especially if it is bound up with the structure of their lifeworld and identity. As an alternative, Habermas proposes both religious and non-religious citizens be allowed to invoke any reasons for or against policies at the level of the *informal* public sphere, provided they take one another's claims seriously and do not dismiss them from the outset. But when it comes to the institutions of the *formal* public sphere concerning coercive lawmaking, justifications should only be based in reasons that all can accept.

This view is somewhat unsatisfying for several reasons: it simply moves the asymmetrical burden of translation “up a level,” it may run into concerns of a metaphorical split in identity, and it could even saddle non-religious citizens with undue burdens (Yates 2007, Lafont 2009). For present purposes, the most charitable reading is that Habermas assumes all democratic citizens have an obligation to adopt a thoroughly self-reflective attitude. Religious citizens must “self-modernize” insofar as they are expected to be open to things like the authority of science, the need for non-religious reasons backing coercive law, and the possible validity of claims made by other religions. But, this also means non-religious citizens must move beyond a dogmatic secularist understanding wherein it is impossible for religious claims to have any cognitive value whatsoever. Indeed, given that some fundamental moral notions—such as equal human dignity—have been inextricably tied to the history of world religions, he claims it is not always clear where the boundaries of the religious and secular are. Determining these boundaries (and what can count as publicly acceptable) may at times be a cooperative task wherein each side takes the claims of the other with some degree of seriousness (2006b, 45 and 2003b, 109). Habermas' reinterpretation of popular sovereignty also explains why he has adopted the theory of constitutional patriotism pioneered by Dolf Sternberger. Constitutional patriotism maintains that, in contrast to national identities of the past, modern political communities can base their collective identities around the unique ways they appropriate and embed the abstract, universalistic principles of democratic self-determination within their unique

histories and traditions. On such a model, political allegiance can coalesce around “a particularist anchoring of...the universalist meaning of [principles such as] popular sovereignty and human rights” (BFN 500; L’i 308; BNR 106). This particularist anchoring would presumably include the way in which a community takes up the abstract system of rights, the process model, and (L). The claim is that the specific way a political community instantiates the “abstract procedures and principles” of the modern democratic state fosters the development of a “liberal political culture” that “crystalizes” around that country’s constitutional traditions, structures, and discursive fora (IO 118; DW 78). The integrative force that emerges against this backdrop is called civic solidarity, which Habermas characterizes as “an abstract, legally mediated solidarity among citizens... a political form of solidarity among strangers” (DW 79; BNR 22). This is essentially the integrative potential of democratic citizenship when it is actively used.

One assumption here is that “culture and national politics have become...differentiated” from one another; citizens can see themselves as part of a shared political culture precisely because they no longer see the state as a vehicle for realizing a homogenous, pre-political nation. While this is a far cry from empirical realities in many parts of the world, Habermas sees the European Union as illustrative in this regard. Even in a context that was once characterized by strong national identities (where the chances for such an identity might seem slimmer than in more multicultural contexts) we can start to see how “a common political culture could differentiate itself off from the various national cultures” and how “identifications with one’s own forms of life and traditions [could be] overlaid with a patriotism that has become more abstract, that now relates... to abstract procedures and principles” (NC 261; BFN 507, 465; IO 118; BNR 327; DW 78).

Finally, Habermas sees constitutional patriotism as a normative resource that could help to expand civic solidarity across political borders and uncouple legal structures from the nation-state so they could be scaled-up into new institutions of international law. Such developments would allow new forms of democratic self-governance above the nation-state at regional and global levels (DW 79). These post-national implications are naturally produced by Habermas’

core theoretical commitments. Deliberative democracy is committed to institutionalized discourse that in some way makes it possible for law to be justified to the persons who are affected by or subjected to it. Given increasing global interdependence this obviously pushes in cosmopolitan directions. However, at the same time, it is important to remember that communicative power must be rooted in a community with a shared ethical-political identity, and that constitutional patriotism is parasitic upon a particular political culture. This rootedness means that civic solidarity and new forms of self-governance can stretch, but only so far.

This anchored cosmopolitanism yields a multi-level constitutionalization of international law that aims at some measure of global governance without government. While Habermas' account of such a multi-level system is only a sketch and many details need filling-in, the broad outline is clear. He proposes a system comprised of the "supranational" (global), "transnational" (regional), and national level political institutions with different roles. A supranational organization akin to a reformed United Nations is envisioned as securing international peace, security, and core human rights. At the mid-level, transnational authorities like the EU would tackle technical issues through coordinative efforts and political issues through negotiated bargaining among sufficiently representative regional regimes of commensurate stature. Finally, nation-states would retain their status as the locus of democratic legitimation. This would require the spread of democratic structures to each nation-state so that laws can reflect the will of the community and so that they could be reliably in line with the basic human rights secured by a supranational organization.

This vision of a multi-level political system for the constitutionalization of international law can be criticized as demanding both too much and too little. Habermas' version of cosmopolitan deliberative democracy locates the touchstone of legitimacy in the fact that "citizens are subject only to those laws which they have given themselves in accordance with democratic procedure" (CEU 14). From this perspective of democratically legitimate law, the proposed system may demand too little. Despite Habermas' insistence that negotiation between regional regimes

could take place in a way that would not “impair deliberation and inclusion,” it is hard to see how such bargaining could really constitute a process where citizens give themselves the law through democratic procedures (CEU 19). From the perspective of rootedness in political culture, the multi-level system may also demand too much with the extension of civic solidarity to transnational regimes. Habermas clearly thinks there are limits to such an extension, as “the transnational extension of civic solidarity...comes to nothing...when it is supposed to assume a global format.” However, apart from the fact that neighboring countries might be supposed to have some minimal level of shared history and culture born out of territorial proximity and an interdependency of interests, it is unclear why this extension of solidarity would reach the levels needed to underwrite the democratic legitimacy of laws within transnational units of regional governance (CEU 62).

While Habermas is certainly aware of these criticisms, he is largely focused on defending his political theory in broad, systematic terms. If the broad normative outlines are correct then the overall theory will stand regardless of how the empirical details are filled in. Indeed, Habermas is rather unique among contemporary philosophers both in his systematic approach to large areas of theory and in his willingness to allow others to fill in the details of how particular claims might work. He has always insisted that philosophers do not speak from a privileged place of knowledge. The best that they can hope for is to articulate a theory that can be convincingly and rigorously tested and debated in the public sphere. We can perhaps understand not only his political theory, but several other theoretical projects in this spirit of a public intellectual putting forth a theory for testing and debate that requires further articulation by those who come after.

6. References and Further Reading

a. General Introductions to Habermas

The article presented a general and reasonably complete introduction to Habermas. However, given the breadth of his work and space constraints, the following should also be consulted:

- 1978. McCarthy, Thomas. *The Critical Theory of Jürgen Habermas*. The MIT Press.
- 1988. White, Stephen K. *The Recent Work of Jürgen Habermas*. Cambridge University Press.
- 2005. Finlayson, James Gordon. *Habermas: A Very Short Introduction*. Oxford University Press.
- 2011. Fultner, Barbara (ed.) *Jürgen Habermas: Key Concepts*. Acumen Press.
- 2014. Bohman, James and Rehg, William. *Jürgen Habermas*. Stanford Encyclopedia of Philosophy.
- Thomas Gregersen maintains an online bibliography at the Habermas Forum.

The following printed bibliographies are also useful:

- 2013. Corchia, Luca. *Jürgen Habermas. A Bibliography: Works and Studies (1952-2013)*. Pisa (IT): Arnus University Books.
- 2014. Müller-Doohm, Stephan. *Jürgen Habermas – Eine Biographie*. Berlin: Suhrkamp.

b. Introductory Books and Articles on Specific Themes

i. Biography

- 2001. Matustik, Martin Beck. *Jürgen Habermas: A Philosophical-Political Profile*. Rowman and Littlefield.
- 2010. Specter, Matthew G. *Habermas: An Intellectual Biography*. Cambridge University Press.
- 2004. Wiggershaus, Rolf. *Jürgen Habermas*. Reinbek Bei Hamburg: Rowohlt.

ii. Linguistic Turn

- 1994. Cooke, Maeve. *Language and Reason*. The MIT Press.
- 1999. Lafont, Cristina. *The Linguistic Turn in Hermeneutic Philosophy*. Jose Medina (trans.). Cambridge University Press.
- 2016. Lafont, Cristina. *Jürgen Habermas in The Blackwell Companion to Hermeneutics*, 440-445.

iii. Discourse Ethics

- 1994. Davis, Felmon John. *Discourse Ethics and Ethical Realism: A Realist Realignment of Discourse Ethics*. *European Journal of Philosophy*, 125-142.
- 1997. Rehg, William. *Insight and Solidarity: The Discourse Ethics of Jürgen Habermas*. University of California Press.
- 2000a. Finlayson, James Gordon. *Modernity and Morality in Habermas's Discourse Ethics*. *Inquiry*. 43, 319-40.

iv. Political Theory

- 1994. Bohman, James. Review: Complexity, Pluralism, and the Constitutional State: On Habermas's Faktizitat und Geltung. *Law and Society Review*, 897-930.
- 2002. *Discourse and Democracy: Essays on Habermas's Between Facts and Norms*, ed. Rene von Schomberg and Kenneth Baynes. SUNY Press.
- 2010. Hedrick, Todd. *Rawls and Habermas: Reason, Pluralism, and the Claims of Political Philosophy*. Stanford University Press

c. Works Cited

Most of Habermas's work can be found in German and English. After the original year of publication and title, see the square brackets for the English translation. For some texts a translation does not exist, only exists in part, or is divided between texts. This is denoted by an asterisk (*).

- 1953. *Mit Heidegger gegen Heidegger denken. Zur Veröffentlichung von Vorlesungen aus dem Jahre 1935*. *Frankfurter Allgemeine Zeitung*. July 25, 1953. [English: 1977]
- 1956. *Der Zerfall der Institutionen* (Arnold Gehlen). *Frankfurter Allgemeine Zeitung*. July 4, 1956.*
- 1958. *Philosophische Anthropologie: Ein Lexikonartikel*. In: A. Diemer, I. Frenzel (eds.) *Fischer-Lexikon Philosophie*. Frankfurt am Main: Fischer. Pp. 18-35.*
- 1962. *Strukturwandel der Öffentlichkeit*. Darmstadt: Luchterhand. [English: 1989]
- 1967. *Probleme einer philosophischen Anthropologie*. Lecture transcript from the 1966/7 Winter semester at the University of Frankfurt (unauthorized edition).*
- 1967. *Zur Logik der Sozialwissenschaften*. Tübingen: J.C.B. Mohr. [English: 1988a]
- 1968a. *Technik und Wissenschaft als Ideologie*. Frankfurt am Main: Suhrkamp. [English: 1970, 1973b*]
- 1968b. *Erkenntnis und Interesse*. Frankfurt am Main: Suhrkamp. [English: 1971b]
- 1969. *Protestbewegung und Hochschulreform*. Frankfurt am Main: Suhrkamp. [English, 1970]
- 1970 *Toward a Rational Society*, J.J. Shapiro (trans.) Boston: Beacon.
- 1971a. *Theorie und Praxis*. Frankfurt am Main: Suhrkamp. [English: 1973b]
- 1971b. *Knowledge and Human Interests*. J. J. Shapiro (trans.). Boston: Beacon.
- 1971c. *The Christian Gauss Lectures: Reflections on the linguistic foundations of sociology* [For published versions, see chapter 1 of German 1984b and pp.1-103 of English 2001* original translation for lecture purposes by Jeremy Shapiro; re-translated for publication by Barbara Fultner]

- 1973a. Wahrheitstheorien. In H. Fahrenbach (ed.), *Wirklichkeit und Reflexion*. Pflingen: Neske. 211-265. Reprinted as chapter 2 in 1984b.*
- 1973b. *Theory and Practice*, J. Viertel (trans.). Boston: Beacon.
- 1973c. Nachwort / Postscript to *Knowledge and Human Interests: Philosophy of the Social Sciences* [included in all subsequent printings of *Knowledge and Human Interests*].
- 1973d. *Legitimationsprobleme im Spätkapitalismus*. Frankfurt am Main: Suhrkamp. [English: 1975]
- 1975. *Legitimation Crisis*, T. McCarthy (trans.). Boston: Beacon.
- 1976a. *Zur Rekonstruktion des Historischen Materialismus*. Frankfurt am Main: Suhrkamp. [English: 1979*]
- 1976b. Was heißt Universalpragmatik? In K.-O. Apel (ed.), *Sprachpragmatik und Philosophie*. Frankfurt am Main: Suhrkamp. 174-272. [English: 1979, chap. 1]
- 1977. Martin Heidegger, on the publication of lectures from the year 1935. *Graduate Faculty Philosophy Journal* 6, no. 2: 155-180.
- 1979. *Communication and the Evolution of Society*, T. McCarthy (trans.) Boston: Beacon.
- 1981. *Theorie des kommunikativen Handelns*. Band I: Handlungsrationalität und gesellschaftliche Rationalisierung. Band II: Zur Kritik der funktionalistischen Vernunft. Frankfurt am Main: Suhrkamp. [English: 1984a and 1987]
- 1983. *Moralbewußtsein und kommunikatives Handeln*. Frankfurt am Main: Suhrkamp. [English: 1990a]
- 1984a. *The Theory of Communicative Action, Volume I: Reason and the Rationalization of Society*, T. McCarthy (trans.). Boston: Beacon.
- 1984b. *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns*. Frankfurt am Main: Suhrkamp. [English: 2001* does not include the Wahrheitstheorien essay]
- 1985. *Die Neue Unübersichtlichkeit: Kleine Politische Schriften V.ÿ* Frankfurt am Main: Suhrkamp. [English: 1991]
- 1986a. *Gerechtigkeit und Solidarität: Eine Stellungnahme zur Diskussion über Stufe 6*. In W. Edelstein and G. Nunner-Winkler (eds), *Zur Bestimmung der Moral*. Frankfurt am Main: Suhrkamp. 291-318. [English: 1990b]
- 1986b. *Entgegnung*. In A. Honneth and H. Joas (eds), *Kommunikatives Handeln*. Frankfurt am Main: Suhrkamp. 327-405. [English: 1991b]
- 1987. *The Theory of Communicative Action. Vol. II: Lifeworld and System*, T. McCarthy (trans.). Boston: Beacon.
- 1988a. *On the Logic of the Social Sciences*, S. W. Nichol森 and J. A. Stark (trans.). Cambridge, MA: MIT Press.
- 1988b. *Nachmetaphysisches Denken*. Frankfurt am Main: Suhrkamp. [English: 1992a]
- 1989. *The Structural Transformation of the Public Sphere*, T. Burger and F. Lawrence (trans). Cambridge, MA: MIT Press.

- 1990a. *Moral Consciousness and Communicative Action*, C. Lenhardt and S. W. Nichol森 (trans). Cambridge, MA: MIT Press.
- 1990b. *Justice and solidarity: On the discussion concerning stage 6*. In: T. E. Wren (ed.), *The Moral Domain*. Cambridge, MA: MIT Press. 224-251, S. W. Nichol森 (trans.).
- 1991a. *Erläuterungen zur Diskursethik*. Frankfurt am Main: Suhrkamp. [English: 1993]
- 1991b. *The New Conservatism: Cultural Criticism and the Historians' Debate*. S. W. Nichol森 (trans.).
- 1992a. *Postmetaphysical Thinking*, W. M. Hohengarten (trans.). Cambridge, MA: MIT Press.
- 1992b. *Faktizität und Geltung. Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*. Frankfurt am Main: Suhrkamp. [English: 1996b]
- 1993. *Justification and Application*, C. P. Cronin (trans.). Cambridge, MA: MIT Press.
- 1994. *The Past as Future*. M. Pensky and P. Hohendahl (trans.). University of Nebraska Press.
- 1996a. *Die Einbeziehung des Anderen. Studien zur politischen Theorie*. Frankfurt am Main: Suhrkamp. [English: 1998a]
- 1996b. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*, W. Rehg (trans.). Cambridge, MA: MIT Press. [German, 1992]
- 1998a. *Inclusion of the Other: Studies in Political Theory*, C. Cronin and P. DeGreiff (eds). Cambridge, MA: MIT Press.
- 1998b. *Die postnationale Konstellation*. Frankfurt am Main: Suhrkamp. [English: 2001a]
- 1999a. *Wahrheit und Rechtfertigung*. Frankfurt am Main: Suhrkamp. [English: 2003a]
- 2001a. *The Postnational Constellation*, M. Pensky (trans., ed.). Cambridge, MA: MIT Press.
- 2001b. *Die Zukunft der menschlichen Natur. Auf dem Weg zu einer liberalen Eugenik?* Frankfurt am Main: Suhrkamp. [English: 2003b]
- 2001c. *Zeit der Übergänge. Kleine politische Schriften IX*. Frankfurt am Main: Suhrkamp. [English: 2004b]
- 2003a. *Truth and Justification*, B. Fultner (trans.). Cambridge, MA: MIT Press.
- 2003b. *The Future of Human Nature*, W. Rehg, M. Pensky, and H. Beister (trans.). Cambridge: Polity.
- 2004a. *Der gespaltene Westen. Kleine politische Schriften X*. Frankfurt am Main: Suhrkamp. [English: 2006a]
- 2004b. *Time of Transitions*. C. Cronin (trans.). Cambridge: Polity Press.
- 2006a. *The Divided West*. C. Cronin (trans.). Cambridge: Polity Press.
- 2006b. *Pre-political foundations of the democratic constitutional state*. In J. Habermas and J. Ratzinger *The Dialectics of Secularization: On*

Reasons and Religion. B. McNeil (trans.). San Francisco: Ignatius. 19-52.

- 2007. Kommunikative Vernunft und grenzüberschreitende Politik. Eine Replik. In: Anarchie der kommunikativen Freiheit Jürgen Habermas und die Theorie der internationalen Politik. Peter Niesen and Benjamin Herborth (eds.). Frankfurt am Main: Suhrkamp Press.
- 2008. Between Naturalism and Religion. C. Cronin (trans.). Cambridge: Polity Press.
- 2008. Ach Europa. Kleine politische Schriften XI. Frankfurt am Main: Suhrkamp. [English: 2009]
- 2009. Europe: the Faltering Project. C. Cronin (trans.). Cambridge: Polity Press.
- 2012: The Crisis of the European Union: A Response. C. Cronin (trans.). Cambridge: Polity Press.
- 2014. The Lure of Technocracy. C. Cronin (trans.). Cambridge: Polity Press.
- 2014. Entgegnung (x13) and Schlusswort. In: Habermas und der Historische Materialismus. Smail Rapic (ed.). München: Karl Alber Press.

d. Secondary Scholarship Beyond the Subject-Specific Recommendations Cited Above

- 1990. Apel, Karl-Otto. Diskurs und Verantwortung. Frankfurt am Main: Suhrkamp.
- 2002. Apel, Karl-Otto. Regarding the relationship of morality, law and democracy: on Habermas's Philosophy of Law (1992) from a transcendental-pragmatic point of view. In: Habermas and Pragmatism. Aboulafia, Mitchell, Myra Bookman and Catherine Kemp (eds.). 17-30.
- 2014. Brunkhorst, Hauke. Critical Theory of Legal Revolutions: Evolutionary Perspectives. Bloomsbury.
- 2009. Brunkhorst, Hauke, Regina Kreide and Cristina Lafont (eds.) Habermas-Handbuch. Stuttgart: JB Metzler.
- 2000b. Finlayson, James Gordon. What Are 'Universalizable Interests'? Journal of Political Philosophy. vol. 8, no. 4, 456-469.
- 2004. Flynn, Jeffrey. Communicative Power in Habermas's Theory of Democracy. European Journal of Political Theory, 433-454.
- 1993. Günther, Klaus. The Sense of Appropriateness. J. Farrell (trans.). Albany: SUNY Press.
- 1991. Honneth, Axel and Hans Joas. Communicative Action: Essays on Jürgen Habermas's Theory of Communicative Action. Gaines, Jeremy and Doris L. Jones (trans.). Cambridge, MA: MIT Press.

- 2003. Horkheimer, Max and Theodor Adorno. *Dialectic of Enlightenment*. G. Schmid Noerr (ed.), E. Jephcott (trans.). Stanford: Stanford University Press.
- 2001. Joas, Hans. Values versus Norms: a pragmatic account of moral objectivity. In: *The Hedgehog Review* 3, 42-56.
- 2012. Lafont, Cristina. Agreement and Consent in Kant and Habermas: Can Kantian Constructivism be fruitful for democratic theory? In: *The Philosophical Forum*. 43/3, 277-95.
- 2009. Lafont, Cristina. Religion and the Public Sphere: What are the deliberative obligations of democratic citizenship? *Philosophy and Social Criticism* 35, 127-150.
- 1991. McCarthy, Thomas. *Ideals and Illusions*, Cambridge, MA: MIT Press.
- 2007. Müller, Jan-Werner. *Constitutional Patriotism*. Princeton University Press.
- 2000. Müller-Doohm, Stefan (ed.). *Das Interesse der Vernunft: Rückblicke auf das Werk von Jürgen Habermas seit "Erkenntnis und Interesse"*. Frankfurt am Main: Suhrkamp.
- 2007. Niesen, Peter and Benjamin Herborth (eds.). *Anarchie der kommunikativen Freiheit-Jürgen Habermas und die Theorie der internationalen Politik*. Frankfurt am Main: Suhrkamp Press.
- 2002. Owen, David. *Between Reason and History: Habermas and the Idea of Progress*. Albany: SUNY Press.
- 2014. Rasic, Smail (ed.). *Habermas und der Historische Materialismus*. München: Karl Alber Press.
- 1989. Rockmore, Tom. *Habermas and Historical Materialism*. Bloomington and Indianapolis: Indiana University Press.
- 1998. Rosenfeld, Michel, and Andrew Arato (eds). *Habermas on Law and Democracy*, Berkeley: University of California Press.
- 1982. Thompson, John B. and David Held. *Habermas: Critical Debates*. Cambridge, MA: MIT Press.
- 1991. Wellmer, Albrecht. Ethics and dialogue: Elements of moral judgment in Kant and discourse ethics. In: A. Wellmer, *The Persistence of Modernity*, D. Midgley (trans.). Cambridge, MA: MIT Press. 113-231.
- 1995. White, Stephen K. (ed.). *The Cambridge Companion to Habermas*. Cambridge: Cambridge University Press.
- 2007. Yates, Melissa. Rawls and Habermas on religion in the public sphere. *Philosophy and Social Criticism*. 33, 880-891.

Author Information

Max Cherem

Email: Max.Cherem@kzoo.edu

Kalamazoo College

U. S. A.

An encyclopedia of philosophy articles written by professional philosophers.

