

Chapter 1

AGENCY, CAUSALITY, AND HISTORY

In thinking about human potential, we often encounter the view that the future is completely open; that “we can be whatever we choose to be.” This contrasts with positions, at the opposite end of the spectrum, that rest on some sort of determinism—usually biological or theological, but sometimes encased in “natural law” or the inevitability of the “free market.” These views place inexorable limits on human possibilities, with varying degrees of severity. Curiously, the two extremes complement one another. The open-endedness of the anything-is-possible attitude provides no sense of a foundation or platform from which movement toward a future of greater human achievement and quality of life might be launched, and the *probability* of decisive progress appears low-to-minimal; this then ironically reinforces the structural pessimism of the determinists. Historical materialism—the study of human social formations and the principles that regulate their evolution, transformation and supercession—is the best starting point for an inquiry into human potential that avoids the trap set by the convergent extremes of unalterable dystopia and incalculable utopia.

There has been a resurgence of interest in the historical materialist project in recent decades. The publication of G. A. Cohen’s *Karl Marx’s Theory of History* in 1978 sparked a wide-ranging debate, which continues today, concerning the nature of directional forces in history and the role of human agency and rationality in the historical process. My own work, developed independently of Cohen’s, and published in 1984, also aroused some controversy, on a much smaller scale (see Sweezy 1986; Amin 1985; McLennan 1986). Subsequent work by Dimitris Milonakis (1993–94, 1997) and Claudio J. Katz (1994) sought to build, in different ways, on the inquiries into social evolution that in fact go back to the famous mid-century symposium on the transition from feudalism to capitalism involving Maurice Dobb, Paul M. Sweezy and others (*Science & Society* 1977; Hilton 1979). A related discussion of the work of Robert Brenner (1976) on the agrarian origins of capitalism in England (Aston and Philpin 1985) brought together historians, geogra-

phers, and demographers searching for answers to questions regarding the role of class (and other) struggles, population dynamics and markets in shaping the emergence of capitalism and the varied paths of that emergence. The focus shifted later to the world stage, with discussions of the world-system view of capitalism, and the larger historical themes in the work of Andre Gunder Frank and the Europe vs. China problematic (Wallerstein 1974, 1977; Frank 1998; Duchesne 2001–02, 2003; Wong 1997, 2003; Goldstone 2003). And most recently, a group of British social theorists has developed an approach to historical materialist theory that incorporates insights from Charles Darwin and natural selection (Blackledge and Kirkpatrick 2002). The list is undoubtedly incomplete.

In this chapter, I will first review and recapitulate the formalization of historical materialist theory that I originally developed some twenty years ago. This must of necessity be a self-contained statement that reproduces the theory “from scratch”; it would be unfair to impose on the reader the requirement to first read the original article (Laibman 1984, 1992a, ch. 13). At the same time, the reformulation arrived at here can only reflect many subsequent interventions and reconsiderations. In this chapter, the recent work of others will be addressed only implicitly. Chapter 2, with the model of this chapter fully on board, will then take up some of the recent trends and positions in more detail. Chapter 2 will return to the large themes in this discussion, above all the political implications of the several extant trends.

I. A GENERAL MODEL OF SOCIAL EVOLUTION

In theoretical approaches to history, the pendulum has continually swung between perception of commonality in the human experience, on the one hand, and insistence upon variety and the irreducible uniqueness of each historical situation, on the other (for a comprehensive survey, see Carneiro 2000). This has often been seen as a matter of classification and generalization: can “laws,” in the sense of statements proposing observed uniformity among different historical experiences, be formulated? Do, for example, all societies pass through a uniform set of stages in their development? If these historical laws are to withstand scrutiny, it must not be possible to find exceptions to the generalizations, or at least no significant ones. In periods when the theory enterprise has been approached with confidence, law-governed development is postulated; when confidence ebbs, however, empirical research chips away at these conceptual structures, and passage from a “hard” theory to a “soft” one (e.g., Gottlieb 1984)—or, in some formulations, “strong” to “weak” (Wright, et al. 1992)—takes place. And the cycle repeats, enforcing

continual alternation between a theory of evolution that appears to lack empirical foundations and to emanate only from its implied political implications, on the one hand, and a tangle of formless empirical material, acceptable to academia but with no useful broader implications, on the other.

The heart of my earlier proposal, which draws upon numerous sources in the literature (e.g., Marx 1913 (1859); Lange 1963; Dos Santos 1970) is my suggestion to overcome this dichotomy, by developing a theory that is *simultaneously* “hard” and “soft.” I posit *levels of abstraction*, and arrange these into a hierarchy, so that at the “highest” level we find the *abstract social totality* (AST) and at a “lower” level the (more) concrete *social formation* reflecting geo-climatic and developmental variation. (“Higher” and “lower” are, of course, arbitrary metaphoric devices; “inner: and “outer” would work just as well.) At still “lower” levels numerous contingent and accidental factors, including the personalities and capacities of individuals, come into play and infuse variety into the picture, which thus approaches the concreteness of the actual historical record.

If this is to be believable, so that we are not “having our cake and eating it too,” two requirements must be met. First, the higher level of the abstract social totality must inform the construction of the lower levels; it cannot operate in a vacuum, independently of the rich detail of the historical record. The “hard” theory must tell us what the relevant facts are, and help organize historical information into meaningful wholes. Second, the hard theory must be falsifiable. This is not to embrace some sort of “falsificationist” epistemology; it is simply to assert that there must be validity criteria on the basis of which the truth-value of the claims we make regarding the nature of the AST can be established. Rejecting any simple path from pristine facts to confirmation or disconfirmation of theory, I still accept the responsibility to support the theory from the evidence and experience of history. In other words, the theory cannot merely be useful, or fruitful; it must, in some sense, be true as well.

The central feature of human existence, as revealed at the level of the AST, is its *stadiality*. I use “stadial”—here, and throughout this book—as the adjective form of “stage”; it refers to the property of development through a series of well-defined stages. “Stage” and “stadial” seem easier to use than “periodization” and “periodization theory,” which have come into common use (see Albritton et al., 2001); but “periodization” does not have an adjective form (“periodical” obviously does not serve), and “stage” and “stadial” do the job more concisely and clearly.

Terminology established, I proceed to the central property of the AST: social evolution proceeds through a series of stages, and at the level of the

AST these stages are (to quote myself; Laibman 1992, 265) an “*absolutely determinate* ladder.” The stages are, in a terminology that I have developed more recently, *theoretical stages*, as opposed to *descriptive stages*. The latter are the familiar generalizations from observed temporal regularities: the empirical raw material for the derivation of theoretical stages, and—as noted above—the continuing basis for reconfirmation, disconfirmation, and development of the stadial conceptualization at the AST level. Theoretical stages, however, have logical as well as descriptive properties (this formulation, incidentally, suggests that logic is a category of ontology as well as epistemology). In a word, they are “chain-linked.” Each stage in a theoretical sequence requires for its existence some crucial property of the preceding stage; contains within it a crucial contradiction, or progressive insufficiency leading to increasing tension and incoherence; and establishes a crucial foundation that defines the succeeding stage. The stages are therefore not arbitrary in character or number; they are not a recapitulation of “theoryless” history, “just one damn thing after another” at a higher level of generality.

This conception must clearly be squared with human agency; we are, after all, talking about *human* history, not natural history, and the element of human consciousness must be inherent in our account. Determinacy at the AST level does not exist *in spite of* human will and agency, but rather *because of* them; they are definitional for human, as opposed to earlier and coexisting nonhuman, life. I return to this major issue soon.

The central methodological implication of levels of abstraction can now be stated. The determinate ladder of theoretical stages—assuming it exists, an assumption that cannot be separated from confirmation of its ongoing extraction from the raw material of history—occurs *only* at the level of the AST. As soon as we depart from this level and examine social formations at more concrete levels, variety in external conditions and consequent variations in pace and detailed features of development are introduced. We might say that the AST would be directly visible only on a planet with one continent, with no mountains, rivers, narrow isthmuses, or any other barriers to communication and diffusion of cultural traits, and a common climate, flora, fauna, and so on. Our world clearly diverges in momentous ways from this abstraction, and that fact determines the enormous variety of paths and rates of development of human groups evolving in relative isolation from one another (see Diamond 1997, for a thorough treatment).

Distinct social formations come to embody varying combinations of traits associated with the theoretical stages. There are blockages and by-passes: social formations leap over stages in their own development by coming into contact with other formations in which those stages have already

been completed—a “torch relay” effect (Semenov, 1980). Alternatively, external circumstances may create “low-level equilibrium traps” in which a given formation may stagnate. Conditions may even lead to retrogression and extinction along particular evolutionary pathways. In present times, when unprecedented levels of world integration have been achieved among human populations and technology contains hitherto unknown perils as well as promises, we must envision the possibility of the ultimate blockage: general nuclear or ecological extinction. A stadial evolutionary theory does not *predict* human survival! It may, however, point in the direction of the obstacles in the present that need to be overcome, if survival and progress are to be a reasonable prospect. To the extent that these obstacles, identified in theory and at the AST level, are actually effective in the lives of people, the ground or source is established for systematic agency—the consciousness-based movement that can eventually resolve a contradiction and initiate a new stage in sociohistoric evolution.

Building Blocks for a Rigorous Evolutionary Theory

To isolate the crucial elements driving the stages of the AST, we look at the most basic, and inevitable, requirements of social reproduction—what *all* societies must always do. And here we find the core of the insight that defines the Marxist tradition in social theory. Any human society that has ever existed—this is a synthetic *a priori* proposition, not an empirical generalization—has entered into a constant metabolic relationship with its external natural environment. The defining condition of human existence is *labor*: conscious interaction with nature to derive inputs and transform those inputs into forms that sustain life. In this process, since symbolic reference is inherently social and presupposes human interaction, there are two distinct but inseparable aspects: interaction of humans with nature, and interaction among human agents themselves. These are, respectively, the aspects of the social labor process that we call “productive forces” (PFs) and “production relations” (PRs). It is important to stress at the outset that the PFs and PRs are aspects of the same socially and stadially defining activity. There is nothing to gain, and much to lose, from reverting to a classificatory conception of this distinction and arguing about whether to “locate,” for example, tools in the PFs and techniques in the PRs, or whether knowledge “belongs” in one or the other category. Labor activity is intimately connected to consciousness of that activity; ideas and action are in continual symbiotic interaction and tension. All technical activity is simultaneously social, and both forces aspects and relations aspects of that activity are simultaneously present in material practice and corresponding consciousness.

Two fundamental propositions characterize the relations between the PFs and PRs. The first of these will at first seem rather innocuous and obvious: the PFs and PRs of any society are not arbitrary and unrelated; on the contrary, they interact intensely and condition one another. As noted above (and below), social formations draw on elements of the AST in complex and combinatory ways, and all sorts of diverse PF and PR elements may be found coexisting within them. At the level of the AST, however, we find a well-articulated combination of a set of PFs and the PRs appropriate and congenial to that set. The PFs and PRs, therefore, form a consistent whole, to which we give the name “mode of production” (MP). Theoretical stadiality is based on a series of distinctions among the principles defining the MPs that make up the ladder of stages. For the present, the point is a relatively simple one: within a MP, the PFs and PRs are mutually shaping and conditioning. All I wish to emphasize at this point is that nothing in this formulation assigns causal primacy to one or the other pole of this relationship. I have come to believe that the quest for the basis of the “primacy of the productive forces,” initiated by G. A. Cohen, places this inquiry on a wrong path. It is not that the PFs and PRs act equally within the model, as we will see; rather, the distinctiveness of the PFs will be found in a property they possess that is not present—or not as centrally present—in the PRs. Put this way, the question of the relation between PFs and PRs does not deteriorate into a sterile discussion of which pole “causes” which, or which has the “greater” impact, in some quantitative sense, on the other. The PF–PR model (the name I give to the formalization of historical materialist theory that I am proposing) need not struggle against the obvious insight that the relations of production always either assist or constrain the PFs, and significantly shape their qualitative characteristics.

The substance of what we may call the “correspondence principle” (of PFs and PRs) must still be established, of course. Why is it not possible to combine any set of PFs with any known or possible set of PRs? (One thinks of children’s books with the pages cut horizontally in the middle, so that the upper torsos of people or animals can be combined with the lower torsos of completely different ones.) At this point, before we have elaborated the theoretical stages of the AST, we must be content with some general and illustrative answers. But the correspondence principle can only be finally confirmed (if at all) by means of the fully elaborated theory, including articulation among its different levels of abstraction.

The obvious starting point is to treat the PFs as establishing necessary (but not sufficient) conditions for a range of PRs. The coming into existence of a quantitatively significant and dependable surplus product beyond the subsistence needs of direct producers is a clear precondition for any elaboration

of social structure—ruling strata, clergy, bureaucracy, ruling classes, and so forth—beyond the simplest horizontal relations among producers and dwelling or kinship units. Note that advances in productivity are not *sufficient* for a portion of the social product to take on the character of a surplus; for this to happen, there must be a lag in the evolution of needs among the producing population, together with development of means of coercion (surplus extraction). There is a chicken-egg problem here: the means of surplus extraction must exist prior to the surplus, but those means are themselves based on (supported out of) an existing surplus. Like all chicken-egg problems, this one does not prevent chickens and eggs from coming into existence; it only points up the complexity of the process, even at this rudimentary stage. What has been said, however, is sufficient to establish an obvious negative correspondence: complex institutions and stratification cannot be present in societies that do not generate a significant, dependable, and reproducible surplus.

The correspondence principle runs deeper than this, however. Certain forms of PRs are best adapted to any given level of development of the PFs, because the nature of production requires a specific form of *incentive and control*—whether these are democratic and socially solidaristic, or whether, as in the long period lasting from the earliest emergence of social classes until the present, they are antagonistic and involve coerced, exploitative extraction of surplus. In the long prehistory of human civilization, when people existed in small, nomadic bands hunting and foraging for bare subsistence, the absence of a reproducible surplus and the precarious nature of subsistence determine PRs that are communal, egalitarian, and solidaristic—although from much anthropological evidence, this did not preclude serious antagonism and warfare between groups. With the progress of the PFs (a concept yet to be elaborated), the problems of incentive and control become more complex. Once a surplus has been captured and a ruling class solidified on its basis, the means of coercion applied to the direct producers correspond to the level of productive development, and become progressively more sophisticated as the labor activity from which a surplus is extracted itself evolves. This periodic replacement of PRs by progressively more subtle ones—more sophisticated systems of surplus extraction, if you will—enables each stage in the growth of the PFs to proceed to its full potential, and to lay necessary foundations for subsequent stages in PF–PR development. This is the basis for the stadiation conception, to be elaborated shortly.

The second proposition at the foundation of the theory is what I will call the “development principle,” and it captures (more accurately, one hopes) what is intended by “primacy” in many formulations. The motivating question is: does the mode of production have an immanent or necessary ten-

gency to change and develop, and, if so, where in the MP does this tendency reside? Proposition 2 assigns “developmentality” to the PFs, rather than to the PRs. More precisely, in the balance between stasis and transformation, the former is dominant within the PRs, and the latter within the PFs. Note that “dominant” in this wording does not require that the PRs can never acquire innovative momentum, or that the PFs must always be subject to a transformative dynamic. The assertion of the developmental property with regard to the PFs has often been countered by the claim that historical evidence does not bear it out: we know of many societies that languished, or even retreated, for long periods of time, in their technical capacities. There is also the question of the source of PF development. The view, often attributed to G. A. Cohen, that the PFs develop because people purposefully and intentionally develop them in order to reap the benefits of superior efficiency and productivity, is rejected on both empirical grounds, and on the basis of a methodological critique—that this “intentional primacy” of the PFs posits the existence of a super-historical “human nature” characterized by rational choice. My presentation of postulate 2 should therefore have something to say about both of these issues.

Rational choice can mean the particular brand of quantified rationality that comes to characterize behavior of actors in advanced market economies; in this sense, projecting such behavior backward onto precapitalist institutions and times would indeed be a form of “presentism,” and methodologically unacceptable. However, a more general version of rational choice can be formulated, which in effect sees it as an expression of the unique human capacity (and necessity) to create symbolic meaning, manipulate symbols, and, in so doing, transform the external world (White 1969; Becker 1971; Deacon 1997; Harris 1979; Carneiro 2000). Conscious activity is purposeful activity. If there is a nonbiological human nature, involving cultural universals such as symbolic reference, language, and labor, this must imply the ever-present potential to transform the environment, a potential that eventually becomes actual. Humans thus have the capacity (and necessity) for *agency* in a way that other animal species do not, and this agency has generally been effective in transforming the ways in which humans interact with the natural environment. Amidst periods of stasis, and periodic leaps from one level and/or rate of development to another, and given the inherent capacity for diffusion of knowledge and traits among cultural groups due to the superorganic, non-instinctual determination of human behavior, the historical generalization establishing the progressiveness of the PFs throughout history is well established—indeed, one of the few non-controversial and significant generalizations possible. I have always been amused by left thinkers

who, even as they cast suspicious doubt upon all stadial theorizing on the grounds that such theorizing smacks of “determinism” and “fatalism” and amounts to a denial of human consciousness and agency, cannot see the immense role of consciousness and agency in the continual (if uneven) transformation of the productive forces.

But this transformation is indeed uneven. In fact, as noted, there are significant places and times at which it does not occur at all. Where it does not occur, however, pressure does not build up for corresponding transformation of the PRs; this scenario then describes cases of relative stagnation. For the AST all we need is the proposition that in the range of varied conditions in which production takes place throughout history, a significant number of situations will exist in which intentional action by individuals within classes or strata that have the power to act on the PFs will result in PF transformation, and that only rarely will PF changes result in *loss* of productivity or efficiency. Note that “intentional action” that *affects* the PFs does not necessarily mean action *with the purpose of* transforming the PFs.

With regard to the PFs, then, we have the following summary result. Certain societal configurations may or may not effectuate PF development. That development is the source of further progress, but this is not required or expected to appear in every historical circumstance. When the PFs do in fact develop, this is as a result of intentional action, but that intentional action may be for very different purposes than the improvement of productivity, which may in fact be an unintended consequence, and may be driven along pathways that are not intended. Rationality, in the sense of symbolic reference and the capacity for purposive action, is indeed inherent in the human condition, but its embodiment in behavior is historical and cultural.

If the PFs have the developmental property, the next question is: do the PRs similarly have this property? The PF–PR model answers this question with a qualified yes and an ultimate no! To begin with, there is a fundamental difference between the PF and PR aspects of productive activity. While the PFs are embodied in the continual practice of labor and are therefore subject to incremental change, the PRs are enshrined in a symbolic framework of meaning, including more or less complex ideological representations and the embodiment of those representations in institutions (state, church, school). In addition to the PF–PR dialectic, there is a *base-superstructure* dialectic (not adequately addressed, I think, in my earlier formulation of the model). As in the case of the PF–PR distinction, the base-superstructure distinction is not about classification of irreducible social elements; it is rather about the interrelated aspects of social processes and their reproduction in the consciousness of the actors who embody them and carry them

out. We have, then, core institutions—the production relations—surrounded by a variety of supporting institutional structures (of which the triad state-church-school is richly, although not exhaustively, descriptive). The degree of integration of surrounding with core institutions varies significantly. Both sets of institutions have corresponding ideological structures: the forms in which they are reproduced in the consciousnesses of individuals. There thus arises a two-way “relative autonomy” dynamic: between core and surround, and between the practical and ideological instances of each. Aspects of this dynamic are frequently confounded together into the unidimensional base-superstructure dichotomy. For present purposes, I want to highlight only one aspect of this situation: the cultural framework of meaning that provides the structure of a society’s ideological self-definition is strongly connected to the PRs, which after all embody the core interpersonal interactions of the society and are therefore inherently communicative, and symbolic.

The upshot of this is that the developmental property, working directly and incessantly (although unevenly) on the PFs, does not work in the same way on the PRs. Production relations tend more inherently toward stasis, for two reasons: culture, and class. Frameworks of meaning tend to preserve themselves; their very efficacy is bound up with their permanence, or with perception of that permanence infused into the consciousness of social actors. The contrast can be stated sharply: the PFs are (and are perceived to be) powerful precisely when and because they contain powerful impulses toward growth and transformation. The PRs, by contrast, are (and are perceived to be) powerful precisely insofar as they possess (and have an aura of) permanence. When the PRs are class-divided—are defined by the existence of antagonistic classes, on opposite sides of a surplus-extraction relationship—the impulse toward stasis and preservation of what exists, already immanent in the ideological and superstructural systems of definition that pervade any social system, is greatly enhanced by its functionality for the ruling class, which by definition is the site of the power to implement social policy.

Antagonistic and Non-Antagonistic MP Development

The two propositions—the “correspondence principle” and the “development principle”—work together very differently in the two cases of antagonistic and non-antagonistic modes of production, where the former sort contains antagonistic social classes and the latter does not. In Laibman (1984), I illustrated this difference with a pair of simple diagrams, reproduced with some adaptations alongside (Figure 1.1). The left-hand panel of this figure

represents the non-antagonistic form of development. The development principle works (sporadically, but immanently) on the PFs; it is represented by downward movement, as shown by the heavy downward arrow on the left side of the panel. The correspondence principle is represented by the horizontal double-headed arrows. Correspondence between the PFs and PRs is (roughly) preserved as development proceeds, with dominant causality running from PFs to PRs (left to right in the figure). The tendency of the PRs toward stasis is *progressively* overcome by the need for institutional evolution to fulfill the potential of developing PFs, and the tension between development of the PFs and stasis of the PRs is therefore relative, and attenuated. We may think of this version of the model as describing social evolution over long periods of “prehistory”—the era of “primitive communism”—in which the low level of PF development dictates, in the negative mode of determination explained above, a regime of general equality and social solidarity in the PRs, and the absence of class formation. The fact of PF development, however, means the gradual emergence of a potential surplus, and where the circumstances of population growth, cultural diffusion and superstructural institutions facilitate this, the potential surplus is captured by emerging differentiated social strata—warrior castes, priesthoods, nobles and commoners. This process culminates in consolidation of ruling and exploited classes. The left panel of figure 1.1, needless to say, is a condensation of many thousands of years of (pre)history, taking place in a vast variety of forms.

When class division has been consolidated, the movement is then best described by the right panel of figure 1.1. Here, class conservatism is added to the general tendency toward stasis of the PRs, and PF development presses against a given system of class relations—embodied in and supported by legal and religious institutions and ideologies, specialized institutions of administration and coercion (the state), and other social structures. The resulting tension is represented in the figure by the wavy line connecting the static PR to the developing PF; it is a useful heuristic to think of this wavy line as a rubber band, which is progressively stretched as the PFs develop further and further within the existing PRs. Three aspects of this increasing tension are worth noting. First, the PRs, which have now come to embody a particular form of class power and domination, exert increasing pressure of a particular kind, forcing the development of the PFs along a characteristic path. The actual nature of technical development in a period described by this particular PF–PR tension (i.e., this particular mode of production) is determined by the social relations within which it takes place. I have always been surprised by the desire of some participants in the historical material-

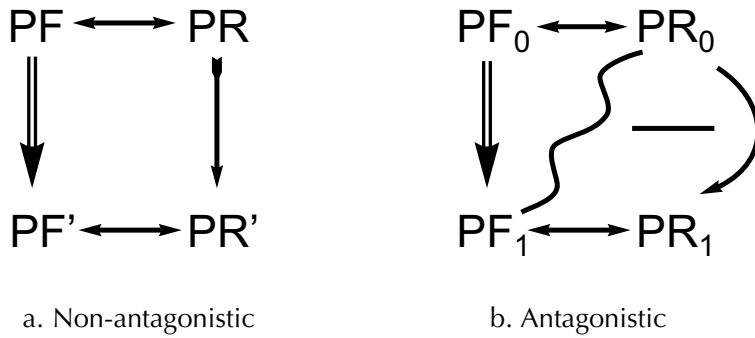


Figure 1.1. Social Evolution at the AST Level

ism discussion to characterize the PF–PR model as a form of “productive forces determinism” (which apparently means any conceptualization that takes the productive forces seriously), when in fact the model not only allows for the persistent shaping of the PFs and technical progress by social relations (the PRs), but in fact places that process at its analytical center.

The second aspect of the PF–PR tension is the clear implication that a given MP eventually and progressively restrains the PFs and blocks the ongoing development of productivity. The PRs turn from a form of development of production to “so many fetters,” to be “burst asunder,” in a classic phrase (Marx and Engels 1998 [1848], 11). The forcing of productive development along a particular line suitable for the existing PRs drives production in an increasingly problematic direction, in which the PRs are more and more revealed to be insufficient, even for existing levels and forms of output. This insufficiency may result in a crisis of surplus-extraction, that is, of the PRs; at the same time, potentials for further productive development along different lines are increasingly thwarted—although this effect may not be present in the consciousness of the actors within the existing MP, only being revealed in a subsequent departure.

The third implication of PF–PR tension is inherent in the rubber-band metaphor: when a rubber band is stretched, it eventually snaps. This is shown in the figure (right panel) by the horizontal bar blocking slow consistent development of the PRs, and the curved arrow representing (more or less) rapid transformation of the PRs. Since the PRs show a dominant tendency toward stasis, the set PF_0 – PR_0 represents a well-defined mode of production (unlike the nonantagonistic form of development, in the left panel of the figure, in which no clear distinction between two discrete MPs can be discerned through the process of evolution depicted), and the breaking point

represented by the stretched rubber band effectuates a transition to a similarly discrete, and more advanced, mode of production (PF_1-PR_1). This provides an analytical definition of *revolution*—the transition from one well-defined mode of production to another. (Needless to say, there is no simple association between this analytical revolutionary transition and the political revolutions that appear in more recent history, although some of these at least may be seen as embodying revolutionary transformation of the sort represented by a shift from one mode of production to another.) The revolution concept adduced here also serves as the basis for the *theoretical stages* that will be identified as constituting an inner model for historical analysis—the central dynamic of the AST. In this further development of the PF–PR model, we will be concerned not merely with underlying principles of social transformation, at the AST level, but rather with establishing the precise number of theoretical stages of development (modes of production) and the principles involved in their progressive supercession—what precise role each one plays in preparation for the one to follow, and how each precisely resolves the contradiction of the one preceding it.

Before we can trace out the stadial model in detail, however, a fundamental issue must be addressed: the precise nature of the break in the chain of biological evolution represented by the emergence of human, cultural existence.

The Biology-Culture Interface

The question of the precondition for human social evolution takes us to the interface between biology and the human sciences. In a masterful defense of science against postmodern nihilism, Barbara Ehrenreich and Janet McIntosh (1997) trace the valuable elements in the cultural and political critique of science—especially of the positivist epistemology and the presumption of scientific neutrality often adopted by the scientific establishment. They note and appreciate the substantial opposition to trends (sociobiology, eugenics, and, more recently, evolutionary psychology) that posit innate behavioral properties in human beings—from the territorial imperative to genetically determined drives for hierarchy and domination—with perverse political implications. Their concern is to decry the postmodern advance beyond this valid critique to the proposition that biology is *completely* irrelevant to human life; that we are culturally defined, and in that sense unique among life forms. The culturalist critics, they write, erroneously “dismiss the possibility that there are any biologically based commonalities that cut across cultural differences” (12). This attitude is dubbed “the new creationism”; with it, “we have gone, in the space of a decade or two, from what began as a healthy skepticism about the misuses of biology to a new form of dogma” (13).

In full solidarity with this attempt to rescue science as such and biology in particular as such (as distinct from their misuse) from withering idealist and relativistic criticism, I nevertheless point to a sweeping non sequitur that runs throughout the Ehrenreich-McIntosh argument. This is the linking of the binary biology/culture with two others: universal/particular, and necessary/contingent. It is seen in the phrase quoted above, referring to the denial of “*biologically based commonalities that cut across cultural differences.*”

The problem, of course, is that this completely ignores an entire dimension, and arguably a vital one: the *cultural universal*. It is intriguing to note that Ehrenreich and McIntosh come close to this concept in only one place in their article, and that is the one point at which they mention Marx: “. . . tacit assumptions of human similarity are embedded in the theories of even such ostensible social constructionists as Marx, whose theory of alienation assumes (in some interpretations, anyway) that there are authentic human needs that capitalism fails to meet.” Marx’s concept of *species being*, elaborated in the *1844 Manuscripts*, is a thrust toward precisely this cultural universal.

But the concept was best nailed down by the non-Marxist anthropologist Leslie A. White, whose contributions, I believe, have not been fully appreciated (White 1969; see also Becker 1971). White’s term for the *unique* human quality that defines a cultural human nature is the *symbol* (see White 1969). More general than language, the symbol nevertheless finds its major expression in language. It is the capacity to create abstract meanings and bestow them on objects and relations (and not only to respond to signs created outside the organism, a property shared by many species). Symbols are an internal filing system, through which the flow of perceptions is restricted and regulated. The system of symbolic understanding is “dense”; it cannot be restricted to a limited sphere of perception or activity. (We cannot, for example, have symbols for plants, but none for animals.) It is, to use that currently despised word, a “totality.” It provides both the possibility and the necessity of a framework of meaning, thus generating a “superorganic” realm, the realm of culture. Humans live in this realm, as well as in the external, natural world. Thus, our world is, and must be, mediated through symbols, which however do not *create* that world. This is the key to a materialist, as opposed to idealist, constructionism.

Symbols and culture are the basis for the planned, intentional action on the world that we call “labor,” as indeed labor is the basis for the emergence of the symboling capacity. (White speculates on the organic basis of symboling, and considers this to be an open question. He does not use the category “labor” in the way done here, much less refer to Engels on this point.) Labor, in fact, may be thought of as having the same relation to

human or social evolution as Darwinian selection does to biological evolution. The fundamental order of determination, however, is reversed: passive adaptation via natural selection in the natural realm, vs. conscious transformation of the external environment in the social. The social (symboling) process, of course, is far more dynamic and rapid than the biological, which requires the duration of biological time for random mutation and transmission of genetically acquired characteristics.

The concept of the symbol has a number of unifying features that cannot be explored here; I refer especially to the symbolic (abstract) conception of the self and acquisition of the ego, which occurs through language and is therefore necessarily social. Here I concentrate on the conjunction of two elements: the central role of intentional action, and the density (completeness of coverage) of the symbolic field. This conjunction carries a major implication for symboling behavior: to take effect comprehensively, the freeing of behavior from instinctual determination must be (almost) complete—leaving aside the autonomic nervous system, the startle reflex, the sucking reflex. These residues of genetically determined behavior are like the bare minimum of hard-wired programming built directly into a personal computer: the knowledge required to extract data from a disk inserted into the external drive.

Just as the existence as such of the symbolic, superorganic realm is not a matter of degree but of kind, behavior cannot be both instinctually (genetically) and symbolically (culturally) determined. The distinctively human superorganic form of existence cannot emerge until this release from biological determination is qualitatively complete. Here the massive evidence from cultural anthropology, showing the enormous variability of behavior, temperament and talents across human cultures, becomes relevant. Genetic differences in the strengths of certain generalized strivings or aptitudes among individuals, or clusters of individuals, certainly exist; these, however, remain vague potentials in the absence of cultural definition. It is noteworthy that instinctual governance of occasional behavior in other mammals, especially sexual behavior, is bound up in the transmission of pheromones through the sense of smell, and that it is precisely that sense which is highly attenuated in humans.

A Tale of Two Cats

It should be understood that symbolic reference refers to a qualitative distinction between human and (so far as we know) all other animal intelligence, and does not impugn or deny the effectiveness of the latter. A true piece of animal lore may help clarify this.

This is a tale of two extremely intelligent cats, Peri and Samir. Samir is the A-cat, and is clearly dominant—a behavior common not only to mammals, and clearly instinct-driven. (Hierarchy ensures discipline and order, and is therefore an adaptive trait.) As A-cat, however, Samir tends to monopolize the food supply, and his humans were concerned that B-cat Peri was not getting his fair share of the food they set out. The solution to this dilemma: put one of the food bowls on top of the refrigerator—a spot reachable by Peri in a single bound, something that Samir could not achieve.

Samir, however, learned to jump *first* to the kitchen table, and *then* to the refrigerator top. In their next countermove, the humans placed the kitchen door ajar, into a position blocking the space between table and refrigerator. But Samir tackled this challenge as well: he learned to push the door closed with his paw (a task requiring a fair amount of sustained effort), after which the path from floor to table to refrigerator top was again clear. So the humans then did what humans are good at: escalate. They propped the door open with a doorstop. And Samir again rose to the challenge: he worked at the doorstop until it fell away; then he worked at the door; and *finally* used the table to access refrigerator top and food. Our amateur ethologists thus witnessed a sustained sequence of actions, linking cause to effect repeatedly in pursuit of a goal. The proximate moral of the story: never underestimate the intelligence and creativity of animals.

The episode's deeper message, of course, is to question the claim that human consciousness and action have a distinctive quality not shared with other living species, which I have been calling "symbolic reactivity" or "symboling"; a more graceful term is "symbolic reference" (Deacon 1997). I would like to defend the claim, and explore its properties a bit further. The distinction itself is the central issue; human uniqueness in its possession is a distinctly secondary matter. But this is all the more reason to examine Samir's behavior with care, to see if human consciousness (including problem-solving ability) is indeed unique, in a way that has implications for directionality in social evolution.

At issue are the pathways that build upon the simplest form of an organism's interaction with its environment—the unconditioned reflex, or built-in (instinct-driven) response to a stimulus (in the case of Samir, the striving for food). The unconditioned reflex is genetically encoded, and therefore the result of natural selection of an adaptive behavior, a selection that takes place in biological time. Behavior and survival chances are enriched, as is well known, by the elaboration of this mechanism into the conditioned reflex: new behaviors not present in the genome can be created, and other behaviors extinguished, by associating an intrinsic stimulus with an external one. The external stimulus is arbitrary, as in the well-known

behaviorist experiments inducing responses in animals to light or sound events, by associating these events with intrinsically pleasurable or painful stimuli (food, electric shock). The secondary stimulus is arbitrary, but the association must be imposed on the organism from outside.

In the next stage of this evolution, the last two properties are reversed: the organism itself makes the association, but between stimuli that are inherently connected in the environment. This is *learning*. Unlike the conditioned reflex, which can be observed or induced in many animal orders, learning appears to emerge only in mammals. It involves the (usually sudden) occurrence of insight: the moment at which a connection is established, between a simple tool and an outcome (chimpanzees inserting a stick into an anthill), or between a table and a refrigerator top. Learning greatly improves organismic reactivity: the organism makes the connection directly, rather than waiting for natural occurrences (or laboratory researchers) to impose it. With the underlying hunger (and, perhaps, dominance) drive in place, Samir learns one connection (table—refrigerator top); builds that connection into a conditioned response; then learns a new connection (move door—table) and builds the new connection into the conditioned chain; and finally learns a third connection (move doorstep—door), adding that one to the chain. In the absence of symbols—the exhaustive grid of meanings that enables the organism to do all this abstractly, independently of the sensory field, and prior to acting—this is an impressive intellectual achievement.

Symbolic reference completes the sequence. Symbols are arbitrary elements (constructed out of signs, largely but not exclusively linguistic) associated with objects or elements of experience by the organisms themselves. Here a shift in point of reference occurs: symbols emerge as elements of communication among individuals, and have an inherently relational quality. But with that understanding, they represent an alternative framework of meaning—a set of abstractions—that *intervenes* between *external stimuli* and *behavior*. Symbol processing thus precedes action; moreover, action takes place—and can only take place—on the basis of symbolic representations. “External” in this formulation refers to elements in the objective environment outside the organism (and cooperating organisms), but also, and crucially, to impulses coming from the genetic “interior.” The inclusiveness and exhaustiveness of the symbolic field assures that *no* behavior of the organism, other than the most elementary reflexes and the functions of the autonomic nervous system, can result from an automatic triggering source. This “freeing of [from] instinct” is the central reproductive strategy along the chain of evolution leading through the primates to homo sapiens. We are driven by, for example, hunger, or sexual desire, but we cannot avoid translating

these feelings into the superorganic realm of symbols *before* we act on them. We cannot grasp them purely emotionally, or in terms of picture-thinking; the symbolic reference field cannot be turned off.

In *African Genesis* (1967), Robert Ardrey—in total contradiction to his reductionist intentions!—provides a magnificent illustration of the distinction between symbolic and nonsymbolic reference. South African farmers, plagued by crop-stealing baboons, move into the planted fields and hide, while the baboons watch from the safety of the surrounding jungle. One farmer hides in the field while the rest retreat. The baboons will return to the field (to their peril), if they are fooled into thinking that *all* of the farmers are gone; this rests on their ability to count. Experience revealed that the baboons could count up to four or five, but got confused trying to distinguish quantities above those numbers. Imagine trying to “count” items by sheer perception, burdened with the necessity of receiving all of the rich extraneous particulars in the sensory field, that is, without *numbers* (symbols). A daunting task. By contrast, with symbols, they could, in principle, distinguish between 1,000,000 farmers and 999,999. Numbers clearly illustrate the *exhaustive* quality of symbols: *n* implies *n* + 1. There is no quantity for which there is *not* a number.

The schematic in Table 1.1 may help in nailing down the stages in the development of organismic reactivity, or reference. (This is my own construction, but it is based on Ernest Becker’s conceptualization.)

What presumably distinguishes the dullest of human beings from the brightest of cats (Samir) is the distinction, often missed in casual discussions, between learning and symbolic reference along the bottom row of this table. The conceptual lacuna is illustrated in recent reporting on research by Frans B. M. de Waal at the Yerkes Regional Primate Research Center in Atlanta, Georgia. The findings consist of observations of behavioral variations among

Table 1.1. Classification of Levels of Reactivity/Reference

		Connection between Stimuli	
		Intrinsic	Arbitrary
Response of organism	Other-initiated	Unconditioned reflex	Conditioned reflex
	Self-initiated	Learning	Symbolic Reference

different communities of chimpanzees, as summarized in de Waal (2001). Summarizing work reported in the journal *Nature* in 1999, a reporter writes: “The researchers came up with 39 behavior patterns that fit their definition of cultural variation, meaning they were customary in some communities and absent in others, for reasons that could only be explained by learning or imitation” (*New York Times*, June 17, 1999). De Waal, in a commentary on this report, finds the evidence “so impressive that it will be hard to keep these apes out of the cultural domain without once again moving the goal posts.” I can assure him (*a*) that, on my part at least, the goal posts will move no further than the distinction between learning and symbolic reference, a distinction that does not seem to have occurred to most members of the primatological community; and (*b*) that—as indicated above—should symbolic reference indeed be found to exist among primates the “uniqueness” of human beings can readily be abandoned, so long as the qualitative singularity and dynamic inclusiveness of symbols as such are grasped.

Terrence W. Deacon (work cited, ch. 3) develops a different conceptual hierarchy, in which the transitions are from iconic reference, to indexical reference, and finally to symbolic reference. Refer to this work for massive amounts of information about both human and animal behavior, and the crucial concept of co-evolution. The upshot is the defining insight that comes down to us in many forms, from Adam eating the fruit of the Tree of Knowledge in Genesis, to Marx’s 1844 elaborations on human *species being*. This is, once again, the cultural universal—the missing link between biological universals and the cultural particularities so beloved of the “new creationists” targeted by Barbara Ehrenreich and Janet McIntosh. The universe of symbols imposes the ability (and need): to posit and manipulate contrary-to-fact situations; to be aware of space and time and their negations, that is, to know of nonexistence and death; to possess the acute self-consciousness that we call the Self, or the Ego, a symbolic abstraction that thus has language and culture as an unavoidable premise; to elaborate frameworks of meaning, cosmology, ideology; and much else.

Culture has thus effectively replaced nature, insofar as behavior and consciousness are concerned. This is a human universal, and its role is bolstered by the observation—which some may initially find surprising—that we are no longer evolving biologically. On reflection it should be clear that the built environment, the use of tools and machinery in labor, the massive medical intervention that takes place constantly, not to speak of such simple items as clothing, eyeglasses, prosthetics, hunting tools, etc., all ensure that survival of the (biologically) fittest is a thing of the past in human life (and increasingly, as we alter the physical environments and survival chances for

other species, in all life). The gene pool is no longer being cleansed of nonadaptive traits; in biological time, we are in fact ceasing to be a unified species. When the crucial turn in evolution away from natural selection to culture takes place, therefore, it signals the *replacement* of biological with socio-cultural evolution, eventually including the progressive application of genetic engineering (something we quite rightly distrust at our current stage of *moral* evolution!). None of this, of course, should be taken to imply that humans no longer have a biological existence. Our biological functions are the foundation for social existence, and set limits to our physical possibilities in obvious ways. Culture does not transcend or eliminate the survival conditions that we share with all life, even though the environment from which we draw our sustenance is increasingly a cultural product.

McIntosh, in a subsequent reply (1998–99) finds my assertion that “culture . . . has *effectively replaced* nature, insofar as behavior and consciousness are concerned” too sweeping. I am inclined to agree, and now prefer a revised formulation: the natural influences on consciousness and behavior can only take place symbolically, and therefore culturally (in the sense of the cultural universal). McIntosh wants to argue against what she calls the “sponge theory” of consciousness—that the human mind simply absorbs (soaks up) anything that culture imposes on it. The transcendent-culture view that I am defending should be clearly distinguishable from any version of what was once called “Watsonian behaviorism,” a position staked out at the “nurture” end of the “nature-nurture” spectrum. Like McIntosh, I believe that the spectrum itself should be transcended. She refers to studies that “suggest” the existence of certain universal psychological characteristics, such as a tendency (observed even in small children) to posit the existence of “insides” and “outsides” in other sentient beings (a form of “essentialism”), and a tendency to posit the existence of “unseen contaminative forces,” presumably common to all or many cultures. These may be called “innate cognitive predispositions,” as McIntosh does, but much turns on the interpretation of “innate.” Essentialism is clearly impossible without symbols, as is the positing of unseen forces, contaminative or otherwise. Essentialism, in turn, is thought to be a basis for extra-group hostility, and prejudice; McIntosh is concerned to extract the implications of innate predispositions for the existence and exercise of power in social life. This is a complex subject; I will simply note here that even if we assume that there is an innate tendency for all human beings to produce essentialist (inner/outer) conceptions, there is no reason to link this to an inherent trope toward “racial or gendered thinking”: witness the counterclaim of universal human goodness and worth, an *essence* posited by some religious philosophies, which would work against