CHAPTER 6

Organization-Environment Relations

Gerald F. Davis
Northwestern University

Walter W. Powell
University of Arizona

This chapter provides a critical review of the development of theory and research on organizations and environments during the 1970s and 1980s, focusing in particular on the six most influential theoretical approaches during this period: Thompson's contingency theory, transaction cost economics, resource dependence theory, network approaches, organizational ecology, and the new institutionalism. Over this period, researchers have increasingly focused on units of analysis larger than the organization, such as the network, population, or sector, and both empirical work and theory have come to concentrate on patterns of change over time rather than on static relationships among variables. Thus, contingency theory, transaction costs, and resource dependence each characterizes the environment primarily in terms of other organizations and describes how the focal organization deals with its exchange partners; network approaches focus on how the organization's position in a larger network of information and resource exchange affects its activities; and both ecological and institutional approaches shift the figure and ground of organization-environment relations, making changes in the population or field of organizations itself the object of inquiry rather than strictly the cause of particular organizational actions. Points of conflict and convergence among these approaches are emphasized.
Introduction

IN THE FIRST edition of this Handbook, Starbuck (1976) began his excellent overview of research on organization-environment relations by noting that the task reminded him of Jonah attempting to swallow the whale. Starbuck lamented that the relevant literature had grown so extensive and heterogeneous that any effort at synthesis would be difficult. After three decades of research in this area, any attempt at synthesis is not only daunting but foolish; hence, we choose a different tack.¹

We do not present a portrait of a tidy, well-ordered field of inquiry. Purists may prefer a unified "normal science" view to the multiple, and at times conflicting, perspectives that we review and assess. But we think much is to be gained by focusing on significant ongoing streams of research and theory. Each of the perspectives we present—resource dependence, transaction cost economics, networks, institutional analysis, and population ecology—are well represented by groups of scholars doing empirical studies, making theoretical strides, and generally advancing the arguments of their respective "camps." Each of these perspectives is also represented by either important recent theoretical statements (see Hannan & Freeman, 1989; Pfeffer, 1987; Williamson, 1985) and/or edited volumes collecting the most current empirical work and theoretical exegesis (see Carroll, 1988; Mizruchi & Schwartz, 1987; Powell & DiMaggio, 1991; Singh, 1990; Wellman & Berkowitz, 1988; Williamson & Winter, 1989; Zucker, 1988). In our view, these different lines of inquiry represent the most robust current research on organization-environment relations.

We have tried to organize the various literatures in a manner that renders them commensurable and that hopefully offers insight for both newcomers and veterans to the field. Let us stress at the outset that many ostensible points of contention among the various approaches are attributable to either divergent views about the logic of action or to a focus on different levels of analysis. We highlight key concepts and working assumptions, lay out the basic research questions, and present representative works. We emphasize both the strengths and weaknesses of the approaches and conclude with brief discussions of new frontiers and points of convergence with other lines of inquiry. Our goal is to present the major arguments that scholars use to make sense of organization-environment relations.

The Nature of the Environment

Research on organizational environments began to assume salience in the 1960s due to a number of studies that illustrated how factors in an organization’s environment influenced organizational performance and design as well as turnover. Katz and Kahn’s (1966) pivotal work focused attention on the open systems nature of organizations, and this perspective gained wide influence in subsequent research. Early work fell largely into two camps: (a) studies adopting a focal organization or organization set perspective (Dill, 1958; Evan, 1966); and (b) studies that described dimensions by which environments could be analyzed (Emery & Trist, 1965; Terreberry, 1968).

The focal organization approach assumed that the environment consisted of all relevant factors external to an individual organization. Dill (1958) offered the term task environment to describe the key elements outside an organization’s boundaries. Evan (1966) used the notion of the organization set to characterize each of the organizations that a focal organization interacted with to procure inputs and to market outputs. Warren, Rose, and Bergander (1974) added an important cultural dimension to the environment, pointing to the critical role played by ideologies, political values, and professional norms in shaping the task environment.

Other scholars were persuaded that the environments of organizations were becoming
more complex (due to competition, governmental regulation, pressures from private interests, technological change, etc.) and were also changing rapidly (Emery & Trist, 1965; Terreberry, 1968). Consequently, they attempted to develop conceptual schemes for assessing patterns of variation in organizational environments. Were environments stable or turbulent? munificent or scarce? simple or complex?

These early efforts were valuable at directing attention to the role of the environment, but they had limited impact on future research. Subsequent researchers made progress by being clearer about units of analysis—is the object of theoretical interest the transaction, the focal organization, the population or sector of organizations?—and levels of analysis—does the explanation make primary reference to the actions of individuals, organizations, governments, or broader, more diffuse social forces? Eventually, methodological and theoretical innovations enabled researchers to identify more clearly the totality of actors and attributes that shape organizational behavior.

We turn now to more contemporary lines of research, organizing divergent approaches under three general rubrics: (a) dyadic models of organization-environment relations, (b) perspectives that stress the embeddedness of organizations in interorganizational networks or domains, and (c) approaches that focus on the organization of environments. We believe this partition is pedagogically useful. It reflects important differences in levels of analysis or, if you prefer, divergent lenses with which to view organizational behavior. In moving from a dyadic portrait to a more complex picture of an interorganizational network to a more macro focus on the organization of the environment, we shift levels and focus on larger units of analysis. And as we go from attention to the actions of individual organizations to a focus on organizational fields, sectors, or populations, the importance of individual action recedes. Obviously, when we look at a focal organization, management strategy looms large, but when our level of analysis is an industry or sector, the actions of individuals and single organizations assume much less salience. In the research reviewed, this change in focus involves an inevitable tradeoff of more extensive characterizations of particular environments for a “big picture” approach in which fine-grained particularities are often lost.

**Dyadic Models of the Environment: The Adaptive Organization**

Work in the late 1960s and 1970s began to treat an organization’s environment as an important determinant of organizational structure and to focus explicit attention on how variations in exchange relationships led to different patterns of organizational action. The three most important approaches to come out of this line of research and theory are (a) Thompson’s contingency theory, (b) Pfeffer and Salancik’s resource dependence theory, and (c) Williamson’s transaction costs economics.

Common to these three approaches is the centrality of the focal organization: The environment must be taken into account in order to explain the behavior of organizations, but it is still variation at the level of an individual organization that is being explained. Uncertainty is one of the most critical features of the environment in each of these approaches, and a good deal of organizational behavior consists of adaptive responses to environmental uncertainty. Resource exchange relations are taken as the primary source of uncertainty in all of these approaches. This implies that the most important elements of the environment from the focal organization’s perspective (and therefore from the perspective of the researcher who wants to explain its behavior) are those organizations that provide it with inputs or that make exchanges for its outputs. Finally, the importance of technology in
affecting organizational outcomes is downplayed in these approaches (as compared with previous work such as Woodward, 1965), and all organizations are assumed to be more or less similar—that is, acquiring resources in an uncertain world and staffed by boundedly rational managers who seek to optimize both their own and the organization’s interests.

Each of these perspectives is also distinguished by its ability to inform empirical work. Whereas much of the initial writings of the 1960s on the environment tended to be long on metaphor and short on testable implications, a large number of empirical propositions can be derived from each of these three theories.

**Thompson’s Contingency Theory**

James Thompson (1967) in his classic *Organizations in Action* portrayed the basic problem of the organization as achieving rationality in an uncertain world. Organizations are created to pursue some desired outcomes, yet they are faced with technologies and environments of varying levels of uncertainty that limit their ability to plan and execute actions to achieve desired ends. Thus, much organizational behavior can be understood as efforts to achieve a resolution of the tension between uncertainty and organizational rationality.

Thompson viewed organizations as **open systems**, fundamentally interdependent with environments over which they had only limited control, yet subject to criteria of rationality. Organizations can be thought of as taking on three levels of responsibility and control (Parsons, 1960): (a) a **technical level**, concerned with achieving the processing tasks of the organization; (b) a **managerial level**, charged with controlling and servicing the technical unit; and (c) an **institutional level**, which articulates the organization with the community and its institutions. These three levels correspond to different sources and levels of uncertainty. Thompson’s most basic hypothesis is that in order to achieve rationality and self-control, the organization seeks to seal off its technical core from environmental uncertainty by setting apart both the resource-acquisition and output-disposal functions from this technical core. Thus, we see greater uncertainty at the managerial and institutional levels; in particular, the institutional level is oriented to an environment over which it has little control, subjecting it to the highest levels of uncertainty. The managerial level is left to mediate between the technical core and the outside environment.

**Conception of the Environment.** Thompson conceives of the environment in terms of several key dimensions: (a) the organization’s **domain**, or the claims it stakes out for itself; (b) its **task environment**, that is, the elements of its environment with which it is most interdependent; and (c) the **power and dependence relations** implied by the nature of its domain and task environment. The organization’s domain is defined by the claims that the organization makes in terms of its range of products, the customers it serves, and the types of services it renders (Levine & White, 1961). When there exists **domain consensus**, that is, recognition of the organization’s claims by those within the organization and others with whom the organization interacts, the organization has a relatively well-defined role in a larger system. This provides a basis for expectations about what the organization will and will not do, hence providing the organization with a guide for future action.

The most relevant parts of this larger system from the organization’s point of view compose its task environment (Dill, 1958): (a) customers or clients; (b) suppliers of materials, labor, capital, equipment, and work space; (c) competitors for markets and resources; and (d) regulatory groups, including government agencies, unions, and interfirm associations (Thompson, 1967, pp. 27–28). These are the individuals and organizations that make a difference to the focal organization by
helping or hindering it in setting and attaining its goals. The task environment is pluralistic, composed primarily of other organizations with their own domains and task environments. But the dependence of the organization on this pluralistic environment introduces threats to its rationality—in dealing with their own networks of interdependence, organizations in the task environment may not be counted on for continuing support indefinitely. Thus, dependency on the task environment creates contingencies or potential sources of uncertainty for the organization. In addition, elements of the task environment may pose constraints on organizational action. These contingencies and constraints imposed by the task environment limit the organization’s ability to act.

Following Emerson (1962), Thompson (1967) posits that the organization’s dependence on an element of its task environment increases “(1) in proportion to the organization’s need for resources or performances which that element can provide and (2) in inverse proportion to the ability of others to provide the same resource or performance” (p. 30). Dependence is thus related to the concentration or dispersion of elements in the task environment that provide some form of support. For instance, an organization that faces a monopolistic supplier for a critical resource is highly dependent on that supplier, and, similarly, an organization that has only one buyer for its outputs (e.g., a defense contractor that sells only to the government) is also extremely dependent.

Power is simply the obverse of dependence. In our example, the government would have a great deal of power over the defense contractor. By this definition, power and dependence are not a zero-sum game: Organizations can and do become increasingly interdependent. Organizations may gain increasing power over each other or they may become increasingly dependent on each other. As we have seen, however, dependence implies a greater potential for constraint and contingencies imposed by the task environment, creating greater uncertainty for the organization. Thus, one of the key tasks for the organization seeking to achieve rationality is the management of its interdependence.

**Strategies for Achieving Organizational Rationality.** The organization seeks to manage the uncertainty imposed by its interdependence with the environment in two ways: (a) through internal strategies of adaptation and adjustment, or organizational design, and (b) through external strategies, or modes of interaction. As we have noted, the primary method the organization uses for achieving rationality is to buffer or seal off its core technologies. This can be achieved in several ways that require only minimal changes in organizational design. Organizations can stockpile both inputs and outputs, shifting the environmental uncertainty from the technical core to the resource procurement and output disposal components of the organization. This is typically costly, however, and so the organization in an unsteady environment will seek to smooth out or level inputs and outputs; for example, utilities offer lower rates during off-peak times in an attempt to level out demand over the course of the day. These two devices, stockpiling and leveling, are rarely sufficient to fully eliminate environmental fluctuations and uncertainty from the organization’s technical core. Consequently, organizations turn to more complex ways of structuring themselves to deal with environmental uncertainty.

According to Thompson, the basic issue of where to place the organization’s boundary is in large measure determined by the loci of critical contingencies in the environment. This issue of where the line between an organization and its environment is drawn is also critical to resource dependence theory and the transaction costs approach, the other two theories considered in this section. For Thompson, a costly but effective way to cope with a part of the environment that creates uncertainty for the organization is to internalize
it—to place the organization’s boundary around that element of the environment. Thus, the expansion of organizational boundaries is not a random process of growth; rather, growth will tend to be in the direction of the crucial contingencies facing the organization, that is, those aspects of the technology or task environment that are the source of the greatest uncertainty for the organization. Vertical integration is to be expected of organizations employing long-linked technologies (such as large-scale manufacturers), while those employing mediating technologies (such as banks or other services that link together clients) will tend to grow by increasing the populations served, as this is the source of the greatest environmental uncertainty. The basic point, then, is that uncertainty generates pressures for the organization to grow in order to absorb the uncertainty. This process may be limited, however, by such obstacles as government intervention (e.g., antitrust laws limiting vertical integration) or limited resources on the organization’s part.

A final aspect of organizational structure determined by environmental uncertainty is the complexity of the boundary-spanning element. The general proposition is that the complexity of the environment is reflected in the complexity of the organization’s structure, or the number and variety of units. A heterogeneous task environment presents the organization with a great number of constraints, and a dynamic task environment presents the organization with a great number of contingencies. The organization responds to this kind of uncertainty by setting up units designed to cope with specific contingencies. Thus, “the more constraints and contingencies the organization faces, the more its boundary-spanning component will be segmented” (Thompson, 1967, p. 73). This argument is extended in later theoretical work by institutional theorists (see Powell, 1988; Scott & Meyer, 1991).

In addition to these internal structural adaptations, the organization uses strategies for interaction with the task environment in order to manage its interdependence and shield the internal workings of the organization from uncertainty. Where the basic injunction for internal adaptation was to internalize the uncertainty, the basic strategy for external adaptation is to minimize dependence and to seek power. Because dependence implies constraints or contingencies, the organization will seek to minimize them first by maintaining alternatives (e.g., securing contracts with more than one supplier of a critical resource) and second by seeking power relative to those on whom the organization is dependent. One way to acquire power is to acquire prestige—establishing a positive image of the organization with relevant constituencies helps to control the organization’s dependence on these groups by making support of the organization more attractive. In more highly institutionalized sectors where organizational outputs are difficult to evaluate, these types of processes may become critically important. DiMaggio and Powell (1983) cite the case of a public television station that adopted a multidivisional organizational structure, not because of any pressing technical need, but more to enhance the organization’s credibility with corporate sponsors. Such legitimacy-driven strategies become a major theme in the institutional approach to organizations.

Power may also be acquired through the use of cooperative strategies, such as contracting, coopting, or coalescing with elements of the task environment that are potential sources of uncertainty. The organization, in effect, seeks to create a negotiated environment (Cyert & March, 1963) where the future is more predictable. Organizations may achieve power over each other through the exchange of commitments which, while reducing environmental uncertainty, also constrain the organization’s action in the future. Three types of commitments (Thompson, 1967, pp. 35-36), in increasing order of the constraint they impose, are (a) contracting, or “the negotiation of an agreement for the exchange of performances in the future”;

---

320 Davis and Powell
(b) coopting, "the process of absorbing new elements into the leadership or policy-determining structure of an organization as a means of averting threats to its stability or existence"; and (c) coalescing, "a combination or joint venture with another organization or organizations in the environment." These different strategies for managing dependence by gaining power correspond to differing levels of concentration/dispersion in the task environment.

Critique. Thompson's theory is an extremely powerful and suggestive synthesis of theoretical work on organizations originating from several fields—sociology, psychology, and economics, among others. Its direct impact on later work, such as its emphasis on interorganizational power and dependence and on strategies for linkage and interaction at the organizational level, can be seen in resource dependence theory, the next theoretical approach reviewed. Although it informed a decade of empirical work, perhaps its greatest influence has been on later theory-building: Resource dependence theory, transaction cost analysis, and institutional theory all draw heavily on Thompson's insights into the nature of organizations and their dealings with the environment.

One difficulty with Thompson's approach is that concepts are presented at a very abstract level. Uncertainty, the central concept of the theory, is vague and is not dimensionalyzed and made into a tractable, measurable concept. The relation of uncertainty to other concepts is also not as direct as Thompson would seem to imply. For example, the mere fact that elements in the environment are concentrated is not sufficient to introduce uncertainty; most organizations get all their supply of electricity, a crucial input, from a single supplier, yet this is hardly a source of great uncertainty, as this theory would tend to imply. This suggests that the uncertainty construct deserves to be unpacked and elaborated, rather than being treated as a unidimensional threat to organizational rationality.

Moreover, Thompson's theory is generic—assumed to apply to all organizations. No mention is made of sectors or industries, which can impose systematic variations in the types of constraints on member organizations and the repertoire of available responses. Nor are any patterns of historical or societal variation in environmental conditions mentioned. Particularly notable in its absence as a serious force in this theory is the government. In our example of the utility in the previous paragraph, the reason that no uncertainty arises from this relationship is because regulations guard against the utility's potential for caprice. Yet because the government's effect on this relationship is indirect (i.e., not through direct exchange), it falls outside the scope of the theory. This is a serious shortcoming that is unfortunately common to all three of the approaches reviewed in this section. Despite these limitations, however, Organizations in Action is rightly regarded as an influential and foundational work.

Pfeffer and Salancik's Resource Dependence Theory

Resource dependence theory (Pfeffer & Salancik, 1978) builds on the insights of Thompson (1967) and work on the political economy of organizations (Zald, 1970) while extending the notion of environmental determinism even further than previous approaches. The basic premise of resource dependence theory is that organizational behavior can be explained by looking at the organization's context. If organizations are presumed to be adaptive, then given a small set of assumptions about the organization, it is sufficient to characterize the environment faced to explain the actions and outcomes of the organization. The logic is similar to the doctrine of situationism in social psychology (Bowers, 1973). Resource dependence theory argues that the organization's patterns
parallels that described by Thompson (1967): Organizations are dependent on a resource environment that can impose constraints and create uncertainty, both of which are noxious to the focal organization. According to Pfeffer and Salancik (1978),

the three most elemental structural characteristics of environments are concentration, the extent to which power and authority in the environment is widely dispersed; munificence, or the availability or scarcity of critical resources; and interconnectedness, the number and pattern of linkages, or connections, among organizations. These three structural characteristics, in turn, determine the relationships among social actors—specifically, the degree of conflict and interdependence present in the social system. Conflict and interdependence, in turn, determine the uncertainty the organization confronts. (p. 68)

The most problematic relation from the organization’s perspective is that of dependence on external social actors. Dependence is determined by three factors: (a) the importance of the resource, which depends on the relative magnitude of the exchange (the proportion of inputs or outputs accounted for by the exchange) and the criticality of the resource (the ability of the organization to function without this resource or without a market for this output); (b) the degree of discretion that the external actor has over the allocation and use of the resource; and (c) the concentration of the resource, or the degree to which the external actor has few potential competitors for supplying the resource or a substitute (cf. Emerson, 1962). When this type of dyadic dependence is asymmetric, that is, when the organization is dependent on the external actor

Conception of the Environment. The environment in resource dependence theory
more than the actor is dependent on the focal organization, then the organization is potentially subject to external control by that actor. In addition, dependence creates uncertainty—exchange partners cannot be counted on indefinitely. Thus, much organizational activity can be understood as tactics to manage external control attempts and the uncertainty created by interdependence.

Managing Dependence. Organizations that have power based on their control of critical resources can make demands on the focal organization that threatens its long-term survival by imposing constraints on future actions. The focal organization can respond to its demand environment in two ways: (a) by compliance or adaptation, or (b) by avoiding or managing influence attempts. Pfeffer and Salancik (1978, p. 44) cite ten conditions facilitating compliance. Pfeffer (1985) summarizes as follows:

When power is asymmetric, and when the external organization has the legitimate right to use that power, and further, when the behavior of the focal organization is under its own control and is observable, external control efforts will be more likely to be attempted and to be successful. (p. 418)

Although there is some evidence supporting the notion that resource dependence facilitates interorganizational influence (Pfeffer, 1972a; Salancik, 1979), empirical work in this area has been somewhat limited (Pfeffer, 1985).

Compliance is not the organization's only possible response to environmental demands, however. In fact, compliance may carry threats to the focal organization: It may be costly in the short term, and it may constrain future adaptation. Even if compliance does not carry these direct threats, it may have costs by marking the focal organization as one that can be influenced in the future. Thus, organizations try to cope with environmental demands, and the interdependence that can give rise to these demands, by managing the conditions of social control and by managing and avoiding dependence on particular exchanges. Resource dependence has made its most significant contribution by examining the use of various types of environmental linkages as strategies of managing interdependence.

As Thompson (1967) pointed out, one response to uncertainty caused by environmental interdependence is to absorb it through merger. Three types of mergers are considered in resource dependence theory: (a) vertical, (b) horizontal, and (c) conglomerate or diversification. Pfeffer and Salancik (1978) argue that each of these is undertaken by organizations in order to stabilize critical exchanges. Vertical mergers, or mergers with suppliers or buyers, are hypothesized to be responses to problematic symbiotic interdependence (i.e., situations where one actor uses the by-products of the other). From the definition of dependence above, this implies that firms should be more likely to merge with those firms with whom they have important exchange relationships—organizations that provide a large proportion of the resources the focal organization uses or those that absorb a large proportion of the focal organization's output. In the case of buyer-seller interdependence, large exchanges are most likely to be problematic when the focal organization operates in a more concentrated industry, giving the supplier greater power. Pfeffer (1972b) found results consistent with these predictions: Mergers were more frequent between industries that had significant buyer-seller relationships, and the relation between purchase interdependence and merger activity tended to be stronger in more concentrated industries.

Horizontal merger, or merger with competitors, is hypothesized to be a response to commensalistic interdependence, that is, interdependence with others who compete for the same resources. Pfeffer and Salancik (1978) argue that the uncertainty caused by competitors bears a curvilinear relationship to industry
concentration: At low levels of concentration, the actions of any single competitor have negligible influence on the focal organization's outcomes, while at the highest levels of concentration the actions of a competitor have a great deal of impact on the focal organization, but these actions can be predicted with a fair bit of confidence based on previous behavior. In addition, tacit understandings among competitors are most feasible at the highest levels of concentration. Therefore, competitive uncertainty is highest at intermediate levels of concentration: The actions of individual competitors have a significant impact, but there are too many to allow stable expectations to arise or to create shared anticompetitive agreements. Thus, organizations face the greatest uncertainty and have the greatest interest in managing competitive interdependence in industries of intermediate concentration. Pfeffer and Salancik (1978, p. 125) present industry-level results consistent with this prediction.

Diversification, the final type of growth strategy considered, is posited to be a response to high levels of dependence that cannot be managed in other ways. Organizations facing suppliers or buyers who account for a large portion of the organization's exchanges and for whom the organization has few potential substitutes are particularly vulnerable to influence attempts and uncertainty. Resource dependence theory predicts that organizations under these conditions will seek to reduce their dependence by expanding into other domains, thereby reducing their reliance on any single exchange partner. For example, firms that do a large proportion of their business with the government, a type of dependence that cannot be managed by vertical integration, should be more likely to diversify in order to manage this dependence. Pfeffer and Salancik (1978) report results consistent with this hypothesis. In general, then, resource dependence theory argues that growth through merger will tend in the direction of the organization's most problematic dependencies. Independent of its effects on profit, size tends to increase stability and reduce uncertainty; thus, growth often represents a viable strategy for coping with organizational interdependence.

In addition to merger, organizations may use less extreme strategies for managing their interdependence by establishing a negotiated environment (Cyert & March, 1963) through the use of interfirm linkages. Such linkages allow organizations to coordinate their actions to mutual benefit, creating a collective interorganizational structure that reduces uncertainty and increases access to resources. Interfirm linkages provide access to information as well as a channel for communication and may be a first step in getting ongoing support from the linked party.

Organizations coordinate in many ways—cooptation, trade associations, cartels, reciprocal trade agreements, coordinating councils, advisory boards, boards of directors, joint ventures, and social norms. Each represents a way of sharing power and a social agreement which stabilizes and coordinates mutual interdependence. (Pfeffer & Salancik, 1978, p. 144)

Such linkages do not offer the absolute control that merger affords, but they are much less costly and more flexible.

The most prolific line of research in the resource dependence tradition has focused on patterns of interlocks among boards of directors. An interlock occurs when the same individual sits on the boards of two corporations. Interlocks represent a form of cooptation—bringing a representative of some element of the environment into policy-making bodies of the organization. Following previous arguments, intraindustry interlocks should be most prevalent when uncertainty is greatest, that is, when industry concentration is intermediate. Consistent with this prediction, Pfeffer and Salancik find the number of interlocks within an industry to follow an inverse U-curve.
relationship with industry concentration. Pfeffer (1972c) also found the size and composition of the board to be predicted by the organization's need for linkage to the environment. This line of research has been pursued vigorously by Burt (1983), who has developed and formalized many of the notions of resource dependence theory and built upon better sources of data and methodological advances to produce compelling analyses of inter- and intraindustry directorate linkages across the U.S. economy. Burt's (1983) results suggest that three types of interorganizational linkage—direct interlocks among boards of directors, ownership ties, and indirect interlock ties between firms via financial institutions—tended to be overlapping or multiplex, mapping onto the same patterns of interindustry constraint. This is consistent with the results reported in Pfeffer and Salancik (1978), who found that patterns of resource dependencies predict similar patterns of organizational linkage regardless of the form of link. Burt's theoretical orientation differs from that of Pfeffer and Salancik (1978) primarily in that he places greater emphasis on the structure of relations among organizations and industries and the constraints these impose on the organization's autonomy, rather than on the strength of these relations, as did Pfeffer and Salancik (see Burt, 1980b). This conception proves to be empirically superior at predicting directorate ties (Burt, 1983, chap. 7). Additionally, Burt's research on the effects of autonomy on profitability put him in more direct contact with the concerns of economics than did previous work on resource dependence.

Critique. Resource dependence theory can be criticized on both empirical and theoretical grounds. The primary difficulty with much of the empirical work, including virtually all the research cited in Pfeffer and Salancik (1978), is that it is performed at an inappropriate level of analysis. The theory concerns the actions of individual organizations, yet most of the studies cited above were conducted at the aggregate industry level. The use of zero-order correlations for most analyses is also unfortunate, as statistical control was not available to rule out competing hypotheses. Further, in most cases the theory was not tested using all variables relevant for measuring interdependence (i.e., resource importance, concentration, and discretion). Thus, for example, the fact that the baking industry gets most of its inputs from the farming industry does not imply dependence in the absence of the other two conditions; yet such a result would be implied by the types of correlational analyses reported in Pfeffer and Salancik (1978). Finally, Pfeffer (1987) notes that the results of the original studies have not always been replicated by other investigators, bringing into question the robustness of the theory.

Despite these difficulties, resource dependence theory has proven to be one of the most fruitful frameworks for the analysis of organization-environment relations. Galaskiewicz (1985a) points out that the theory has many powerful hypotheses that have not been fully explored. For example, hypotheses concerning when organizations are likely to engage in collective action, or when the state will be invoked to deal with organizational interdependence, have thus far received little attention (but see the work of Mizruchi, cited in the section on networks). Thus, much unrealized potential still exists for this perspective. On the other hand, the theory contains unresolved difficulties stemming from some fundamental ambiguities in its formulation. Perhaps the most prominent of these is the question of whether managerial action seeks primarily to pursue greater environmental certainty, as suggested by Thompson (1967), or greater autonomy and profitability, per Burt (1983). Pfeffer and Salancik (1978) take no clear stand on this, stating at some points that organizations "are willing to bear the costs of restricted discretion for the benefits of predictable and certain exchanges" (p. 183), yet at other points
implying that the maintenance of discretion and power drives managerial action. This autonomy versus certainty dilemma is noted but not resolved in a theoretically satisfying way. As a result, it is not clear whether resource dependence theory is falsifiable: Analytically, it is no doubt true that virtually any significant organizational action increases either autonomy or certainty; thus, if actions that achieve either of these managerial goals are considered to be consistent with the theory, it is difficult to imagine a state of affairs that would contradict the theory. This problem is highlighted in Burt’s (1983) work, where considerations of profitability achieved via structural autonomy are paramount, making the revised theory consistent with the spirit of microeconomics.

Williamson’s Transaction Costs Economics

The transaction costs economics (TCE) approach to organizations (Williamson, 1975, 1981, 1985) begins with an even more fundamental question than, What do organizations do? From an economic perspective, a more puzzling question is, Why are there organizations at all? In traditional microeconomic theory, firms are merely production functions, vectors of inputs and outputs. The organizational properties of these black boxes are assumed to be essentially irrelevant to the economic functions they serve. The key question, initially posed by Coase (1937), is why these functions could not be carried out by an individual entrepreneur buying and selling her inputs and outputs in markets. Why, Coase asked, should tasks that are not inherently tied together by technology be performed in the context of a single, hierarchically organized governance structure?

The short answer to this question is because markets fail. Under some conditions, exchanges are not efficiently organized through markets. Thus, drawing on literature in economics, contract law, and organization theory, the theoretical agenda for TCE has been to trace the conditions of market failure and to relate them to the existence of organizations (or hierarchies) and the forms they take.

Several assumptions are critical for this approach. The behavioral assumptions are two: First, individuals are boundedly rational (Simon, 1957); that is, people are assumed to be intentional or goal-directed, but limited in their cognitive capacities. Second, at least some individuals are inclined to be opportunistic, or “self-interest seeking with guile” (Williamson, 1975, p. 26). This is not a general assumption about human behavior—it is only necessary that some people are prone to act opportunistically and that it is not easy to tell these people from the rest. An additional assumption is that any exchange relation can be analyzed as a contracting problem. The transaction costs framework treats contracting—implicit, explicit, simple, complex—as pervasive in economic life. This focus on the contractual aspects of organizations is one of the distinctive features of the transaction costs approach. Finally, the transaction or exchange is taken as the basic unit of analysis. According to Williamson (1981), a transaction occurs “when a good or service is transferred across a technologically separable interface” (p. 552).

Williamson’s notion is that the primary purpose served by organizations is the reduction of transaction costs, the “frictions” that accompany exchange relations. It is important to distinguish transaction costs from production costs, that is, those of capital, materials, land, and labor. Transaction costs are the costs of planning, adapting, and monitoring task completion. Initially, this central concept was defined only denotatively—as the economic equivalent of friction in physical systems, or costs that arise when parties to an exchange don’t operate harmoniously. In more recent statements, Williamson (1985) gives a relatively more fine-grained characterization of transaction costs. Transaction costs can arise during the negotiation or execution of a contract. The first type of costs includes drafting, negotiating, and safeguarding the agreement, while the
second includes maladaptation costs, the costs of haggling if the contract must be altered, the setup and running costs of governance structures that resolve disputes (including the courts and other arbitrate bodies), and bonding costs (Williamson, 1985, pp. 20–21).

Transactions can be described in terms of three dimensions, each of which affects the costs of contracting for the transaction. First, transactions differ in the amount of uncertainty they entail. Some are simple exchanges that are accomplished effortlessly (e.g., buying a dozen half-inch nuts), while others are subject to more or less uncontrollable events in the future (e.g., hiring a lab to research an exotic new weapon system using untried technology). Second, transactions differ in their frequency (e.g., a large one-time buy vs. a monthly purchase). Third, transactions differ in the degree to which the individuals or organizations involved need to invest in assets peculiar to that exchange, that is, that cannot be used in other exchanges as effectively. For example, a steel manufacturer that locates a plant next to an auto manufacturer that is its main buyer or a clerk who learns a company’s unique filing system, both have made investments that would be lost to some extent if the exchange were to stop. It is this dimension of asset specificity that is critical in understanding the existence of hierarchical organizations and the forms they take. As Williamson (1985) puts it, asset specificity is “the big locomotive to which transaction cost economics owes much of its predictive content” (p. 56).4

Asset specificity is critical because it transforms the nature of exchange relations, rendering them both valuable and vulnerable. To use the example of our steel producer who locates a plant next door to its major buyer, we can see that both the buyer and supplier have a special interest in keeping this relationship vital. The steelmaker can sell to the auto manufacturer at a reduced rate because of its lower transportation costs, saving the manufacturer money. But this relation benefits both parties: The car manufacturer saves money, and the steel producer, knowing this, can count on the manufacturer as a steady customer. Because this relation may allow the auto manufacturer to sell products for lower prices than competitors who do not have such a relationship, the relation is also an economically advantageous one. In effect, asset specificity creates a situation of small numbers, in this case bilateral supply (Williamson, 1985), where each partner becomes dependent on the other. Thus, asset specificity creates a special relation that both partners to a transaction have an incentive to protect.

The variation of transactions on the three dimensions of uncertainty, frequency, and asset specificity determines the costs of contracting and, by implication, the types of governance structures that surround exchanges, that is, market, organization, or some intermediate or hybrid form. Williamson (1985, pp. 59–60) argues that contrary to what one might expect based on Thompson (1967) or resource dependence theory, the effects of uncertainty and frequency on whether or not a transaction will be brought within an organization are contingent upon the prior existence of asset specificity. Uncertainty per se is not a sufficient condition, because without some degree of asset specificity to vest exchange partners in a particular relationship, new trading relations could be arranged relatively easily on the open market. When asset specificity is present, however, uncertainty (i.e., unpredictable events that affect contract execution) increases the number of occasions that arise to renegotiate and therefore increases the transaction costs due to haggling, misunderstandings, opportunistic behavior, or the relation being discontinued. Similarly, if transactions are infrequent, the cost of bringing the activity inside the organization would not be justified. It is only when one or both of the transacting parties has committed resources to the relationship that cannot be redeployed within the context of another exchange relation that it becomes more
efficient to arrange transactions within an organization rather than through the market.

It is worth noting here that such problems would not arise were it not for bounded rationality and opportunism. If both parties were omniscient, it would be possible to map out all future contingencies relevant to the contract in advance and write the contract with explicit arrangements for such possible eventualities. Due to bounded rationality, however, uncertainty often outstrips our cognitive capacities to write such contingent claims contracts. Similarly, even uncertainty and bounded rationality would not be a problem if both sides could be counted on to act honorably when unanticipated events come up that affect contract execution. The possibility that one side will try to take advantage of the other on these occasions, however, creates a risk that neither may be willing to take. Note that it is not necessary that one or both always act opportunistically but only that there is no easy way for each to be sure of the other's trustworthiness. Thus, it is the simultaneous occurrence of bounded rationality, opportunism (or its potential), uncertainty, and small numbers (due primarily to asset specificity) that creates the conditions for hierarchical organizations to exist as adaptive solutions to problems of market failure.

How do organizations solve these problems that cause markets to fail? According to Williamson (1981),

the advantages of firms over markets in harmonizing bilateral exchange are three. First, common ownership reduces the incentives to suboptimize. Second, and related, internal organization is able to invoke fiat to resolve differences, whereas costly adjudication is needed when an impasse develops between autonomous traders. Third, internal organization has easier and more complete access to the relevant information when dispute settling is needed. The incentive to shift bilateral transactions from markets to firms increases as uncertainty is greater, since the costs of harmonizing the interface vary directly with the need to adjust to changing circumstances. (p. 559)

Organizations may not solve these problems perfectly; indeed, TCE recognizes that running an organization is costly. The cost of constructing an organization to govern transactions will only be justified if these costs are outweighed by benefits gained by not leaving problematic exchanges to the hazards of the market.

Conception of the Environment. The environment in the transaction costs approach is conceived in a manner very similar to Thompson's framework and to resource dependence theory. Much of the focus is on the organization set, that is, those actors in the environment with whom the focal organization engages in significant exchange relations. Both suppliers and customers transact with the organization, and therefore relations with both these actors can be analyzed in contracting terms.

TCE differs from the other approaches in focusing more explicit attention on the employment relation. Whereas Thompson and Pfeffer and Salancik tend to follow the inducement-contribution notion of organizational membership outlined by March and Simon (1958) and to proceed from there, TCE treats the employment relation as problematic and takes variation in the types of employment contracts used by organizations as a matter worth explanation (see Williamson, Wachter, & Harris, 1975). On the other hand, TCE gives relatively less attention to the role of interorganizational competition. This stems in part from the efficiency orientation of the transaction costs approach, which tends to assume that the invisible hand of the market drives competition. In contrast, Pfeffer and Salancik (1978) pay more explicit attention to the role of other firms as competitors.
Other key features of the environment are treated as implicit assumptions in TCE. A crucial notion is that organizations are subject to strong selection pressures. It is assumed that the environment operates to favor organizations that are more efficient at managing transactions over those who are less so through exerting a sort of natural selection over the long run (suggested by Williamson, 1985, pp. 22–23, to typically entail a period of five to ten years). Thus, whereas the previous approaches focused primarily on the organization’s adaptive strategies, TCE maintains a much stronger role for environmental selection to operate in favor of those organizations that have effectively reduced transaction costs and against those that have not. To paraphrase Thompson, organizations adapt to efficiency pressures—or else!

Finally, the larger culture plays a subtle role in some of the work in the TCE tradition. Williamson and Ouchi (1981) note that some governance structures depend on the culture in which they are embedded. For example, soft contracting is more likely to be successful in cultures such as Japan, where dishonesty and other opportunistic behavior are subject to social sanctions. In a culture where organizations are dominated by graduates of MBA programs, on the other hand, hard contracts that take for granted the possibility of opportunism are likely to prevail. This will be reflected in the management styles observed in these different cultures.

Key Areas of Research. A large number of studies have investigated variants of the basic hypothesis that transactions are assigned to organizational forms (or governance structures) so as to reduce transaction costs. We categorize this work into two key areas: (a) research dealing with where to place an organization’s boundary (the efficient boundaries or make-or-buy problem), which has recently included consideration of intermediate forms of governance such as long-term contracts; and (b) work concerned with the internal structure of organizations, including the structure of the employment relation and the use of the multidivisional form or M-form.

The question of whether an organization should buy a component on the market or make it internally was one of the first to be examined in transaction costs terms. Organizations contain one or more units that are irreducible due to the technology involved, that is, units that cannot be further broken down into smaller units. The efficient boundaries problem asks which of these basic units should be grouped together inside a single organizational boundary. Not surprisingly, the answer turns on the level of transaction costs incurred by the alternative modes of organizing exchanges between these units, which in turn depends on the uncertainty, frequency, and asset specificity involved in the transaction. According to Williamson (1981), recurrent transactions with an intermediate level of uncertainty will be arranged as follows:

Classical market contracting will be efficacious whenever assets are nonspecific to the trading parties; bilateral or obligatory market contracting will appear as assets become semispecific; and internal organization will displace markets as assets take on a highly specific character. (pp. 558–559)

Stuckey’s (1983) detailed analysis of backward vertical integration strategies in the aluminum industry is easily the best industry-level study of the boundary problem. Williamson and Ouchi (1981, pp. 356–359) have analyzed the movement of manufacturers toward forward integration into distribution in the late nineteenth century, suggesting that variation in the degree of integration can be explained in transaction cost terms, and Monteverde and Teece (1982) show that an estimate of the engineering cost for developing a component (a type of asset specificity)
significantly predicted the degree of vertical integration of its production by General Motors and Ford. Walker and Weber (1984) also examined 60 make-or-buy decisions in a division of a large automobile manufacturer. They found that, although their measures of asset specificity and uncertainty of future demand were significantly related to vertical integration, "in general, the effect of transaction costs on make-or-buy decisions was substantially overshadowed by comparative production costs" (p. 387). Unfortunately, asset specificity was proxied with level of market competition for suppliers, a measure consistent with several alternate explanations, indicating that this study provides at best weak evidence for the importance of transaction costs in make-or-buy decisions.

Much recent work in TCE has focused on the choice among governance structures to make or buy in particular long-term contracts. Wiggins (1990) has analyzed the efficiency tradeoffs between long-term contracts and internal organization; Joskow (1985) has done extensive studies of long-term contracts for the supply of coal for coal-burning electric utilities; and Paley (1985) has detailed how informal contracts allow transportation companies to avoid regulatory constraints. These studies highlight how transaction cost economizing leads to the choice of alternative ways by which to govern exchange relations.

In addition to vertical integration, two other issues of organizational structure have received considerable attention: the structure of the employment relation and the use of the multidivisional form or M-form. Williamson (1981) argues that human assets can be described by their degree of firm specificity and by the extent to which productivity can be measured or monitored, and that variation on these dimensions will determine the type of employment contract observed. Human asset specificity comes from on-the-job training and "earning by doing," skills not highly useful to other employers. When skills are acquired that are unique to a particular employer, as with knowing an idiosyncratic and complex payroll system, the organization and the employee both have an incentive to protect the employment relation, because if it were to be severed the firm would lose an already-trained employee and have to invest in training another, and the employee would lose the additional pay that comes from already being trained. When human assets are nonspecific and it is easy to monitor individual productivity, an internal spot market will prevail; that is, organizations and employees have little investment in the relation and, therefore, the organization will not take elaborate steps to protect it. This describes the situation facing migrant farm workers, where the "contract" is in essence renegotiated each day. When human assets are nonspecific but it is difficult to meter productivity, employees will be organized as primitive teams. When human assets are somewhat firm-specific but metering tasks is easy, an obligational market will be observed, such as the use of internal labor markets, seniority, and other devices to bind the worker to the firm. Finally, when human assets are specific to the firm and it is difficult to monitor productivity, relational teams will prevail. This implies the use of socialization by the organization and a high degree of job security, as in the clan type organization described by Ouchi (1980).

The multidivisional structure, or M-form organization, has been widely adopted by large American firms (see Chandler, 1962). Williamson (1975) argues that this structure came to dominate corporate life because it economized on transaction costs: Central management was relieved of the burden of making daily operating decisions, allowing it to concentrate on "the big picture" and alleviating the loss of control that comes from an overextended functional structure. Additionally, the introduction of a quasi market into the internal structure through the creation of competition between operating divisions mutes the ability of managers to behave
opportunistically. The efficiency of this innovation has been argued for and tested by Armour and Teece (1978), who found that large firms in the petroleum industry that adopted the M-form prior to 1968 earned a significantly higher return on stockholder's equity (about two percentage points higher on average) than competitors with a functional form. Fligstein (1985) argued, following Williamson's (1975) discussion, that larger organizations and organizations experiencing greater growth should be more likely to adopt the M-form, as they stand to lose the most from retaining a functional form stretched beyond its capacity. But in his analysis, which controlled for other possible explanations, Fligstein found almost no empirical support for this hypothesis in a sample of America's largest corporations between 1919 and 1979. Taken together, these two studies suggest that the reduction of transaction costs may be a consequence of adopting the M-form, but not its cause, a topic we will touch on below.

Critique. In a relatively short time, TCE has had a major influence on research in organization theory. Recognizing that all kinds of governance structures have inherent flaws, TCE is thus a comparative exercise: One form of governance must always be compared with another. Williamson has focused the attention of organizational researchers on the question of what the critical attributes are with respect to which transactions differ. To be sure, a good deal of additional work remains to be done to operationalize the notion of transaction costs. Nevertheless, the general argument—that the comparative efficiency of alternative modes of economic organization varies systematically with the attributes of transactions—is an enduring and original contribution to the analysis of organization-environment relations.

There is no shortage of detractors from this approach, however, particularly among organizational sociologists. Perrow (1981, 1986) argues that the willful neglect of power in this approach blinds it to issues of control and domination. He also contends that the failure to define adequately the nature of transaction costs biases TCE in favor of the choice of hierarchy. More concretely, parties in contracting situations are treated as free agents voluntarily entering into agreements, yet the migrant worker in our example is hardly in the same situation at the bargaining table as his potential employer. Such processes are common in the workplace yet, thus far, have received little attention in the transaction costs approach. A second general criticism of TCE is that many of the costs that Williamson associates with market transactions are reproduced and even exacerbated by bringing transactions inside an organizational boundary (Dow, 1987). Hierarchies may provide mechanisms to resolve disputes, but organizational life can also promote more extensive and costly strife (e.g., between individuals or departments); hierarchies allow routine performance monitoring, but large-scale fraud or embezzlement against an organization requires an insider (Granovetter, 1985). To be sure, Williamson's claim is a comparative one—that hierarchy is superior in these respects to market relations. But clearly as Dow (1987), Putterman (1986), Powell (1990), and others have pointed out, he neglects the myriad alternative forms of organization, for example, networks, nonprofits, cooperatives, small firm consortia, relational contracts, and so forth, that are alternatives to hierarchical control under a wide range of conditions. Moreover, Dow (1987) makes the telling point that no attention is given to the possibility that hierarchical authority might be abused by superiors in an opportunistic fashion, rather than always serving as a device for curbing opportunism among lower level employees.

Granovetter (1985) argues that the characters who populate TCE are undersocialized, that is, their actions lack the sort of social context within which human action is in fact embedded. Exchange partners vested in a relationship due to asset specificity are described as if they were playing a Prisoner's Dilemma game, each
coolly calculating and distrustful of the other, their actions unconstrained by any sense of obligation and determined by potential payoffs rather than by membership or social affiliations or by a history of past associations. More broadly, Zald (1987) points out that macro political processes are outside the scope of this theory; thus, the application of this approach to every economic institution will both oversimplify and depoliticize the historical record.

To these criticisms we will add an observation about the scope of TCE. A key assumption is that competition and market forces exert strong selection pressures on governance structures, winnowing out unfit forms. But selection pressures are not uniformly distributed across the economy. In some sectors—such as industries dominated by oligopolies—and in some branches of the public sector and perhaps in the nonprofit sector, selection pressures operate with modest force, and thus there is no driving environmental force shaping the choice of governance structure. The assumption of efficiency is not a substitute for a well-thought-out consideration of the organization's environment. Granovetter (1985) makes the case rather pointedly that "the operation of alleged selection pressures is here neither an object of study nor even a falsifiable proposition but rather an article of faith" (p. 503). One place to start rectifying this is to consider pressures for efficiency as variable, not fixed. Additionally, it should be noted that selection pressures are not causes of organizational behavior, strictly speaking; rather, their effects shape the consequences of behavior. To the extent that selection and efficiency pressures are muted or operate only over prolonged periods, the degree to which they shape organizational behavior in the short run is problematic indeed, and as the apocalyptic line from John Maynard Keynes goes, in the long run we are all dead.

Limits to the Organization-centered Paradigm

The three approaches described above—Thompson's contingency theory, Pfeffer and Salancik's resource dependence theory, and Williamson's transaction costs economics—greatly extended the power and scope of organization theory by bringing the organization's environment into clearer focus. Fundamental issues, such as how an organization decides where to place its boundaries, received fresh insights. In addition, new issues, such as what factors determine the types of relations the organization maintains with external actors, were raised and probed. Yet each of these approaches is limited by the basic assumptions they invoke. Each assumes that organizations are rather proactive and flexible in their ability to adapt. Each focuses primary attention on dyadic exchange with members of the organization set to explain an organization's actions and structure. And each assumes a fairly high degree of rationality (albeit bounded) in the choices that organizations or their leaders make. We briefly scrutinize each of these assumptions.

Organizations are constrained in their ability to adapt by a number of factors, both external and internal. Stinchcombe (1965) long ago noted the seemingly peculiar tendency of organizational structures to be "imprinted" at the organization's birth. He argued that organizations are created using the social technology of the day, leading to cohorts of organizations sharing much in common but differing from other cohorts. Organizations are founded in "spurts," with particular structures tending to dominate waves of foundings. Stinchcombe (1965) argues that structures are preserved for three reasons: (a) because the form is (still) the most efficient; (b) because structures tend to become institutionalized, infused with value beyond any technical or economic efficiency; and (c) because organizations with particular
structures may operate in sectors with little competition from alternative, ostensibly better forms. Stinchcombe's second and third points argue against the generality of adaptive models of organizations, which presume a fair bit of fluidity and discretion on the part of organizations and their managers with respect to the organization's structure.

Hannan and Freeman (1984) build on Stinchcombe's insights with their arguments for *structural inertia* in organizations. Structural inertia exists "when the speed of reorganization is much lower than the rate at which environmental conditions change" (Hannan & Freeman, 1984, p. 151). Organizations in general succeed at what they do and survive into the future because they are able to produce outcomes with high reliability and because they accede to environmental demands for accountability. Maintaining reliability and accountability may be necessary for an organization to survive, but they are also constraining forces, generating pressures for inertia. Age and size both increase this tendency toward structural inertia. As a consequence, many organizations become trapped by their own competency.

Change efforts typically threaten vested interests within the organization. Change rarely can be implemented without the consent of those outside top management, yet this consent may not be forthcoming when it threatens the distribution of power in the organization. Those who stand to lose under a new regime typically will fight harder than those who stand to gain, and powerful vested interests may have more effective resources for stalling change compared to those who favor a new set of arrangements.

Even if organizations do manage to alter themselves, their continuing success is by no means guaranteed. Attempts at reorganization lower the organization's reliability, increasing the probability of organizational death, and this risk increases with the time it takes to accomplish reorganization. Furthermore, if a new structure is put into place successfully, the organization will lose much of its accumulated wisdom stored in the old structure, thus becoming subject to the "liability of newness" all over again (Hannan & Freeman, 1984). In a sense, organizations seeking to make significant changes in order to adapt to environmental pressures are damned if they do and damned if they don't: Failure to adapt to threatening environments may be fatal, while adaptation also carries its own risks. Yet these facts of organizational life receive scant attention from the adaptive approaches to organization-environment relations. This is an important limitation on the scope of these theories.

According to the various adaptive perspectives, exchanges with other actors in the environment are the primary means by which the environment affects organizations. But the environment is not bounded by the organization set, and exchange is not the only relation of relevance for explaining organizational outcomes. Organizations are subject to legal and regulatory sanctions at the local and national levels. Professional and trade associations, as well as political advocacy groups, can place tremendous normative pressures on organizations. Organizational strategies for coping with the uncertainty that surrounds exchange relations can be limited or facilitated by governments. Thus, analysts who focus on dyadic exchange relations of a focal organization to the exclusion of the larger political economy in which it operates (Zald, 1970) will be limited in their ability to explain organizational behavior.

Organizations are also situated in more or less elaborated networks, linked by information as well as resource exchanges, friendship ties among elites, and directorate interlocks with common partners, among others. Ties may be dense, weak, or absent; they may extend from the local to the transnational level, interacting with political processes at each step.
The dynamics of these networks have profound implications for the constituent elements, yet network effects are lost by exclusive attention to dyadic exchange relations. Thus, the models of organization-environment relations that give theoretical primacy to exchange lose some of their explanatory power by downplaying or ignoring the political economy and interorganizational networks in which organizations are embedded.

Organizations are assumed to be essentially rational actors in each of the approaches described above. But many organizational theorists contend that such an assumption is rather problematic: Organizations are limited in their ability to know and learn from their environment, and actions are often symbolic or habitual rather than prospectively and technically rational. March and Olsen (1976) argue compellingly for the limitations of rational actor models of organizational decision making. Ambiguity is pervasive in organizations: Objectives are inconsistent and ill-defined; cause-effect relationships are poorly understood, particularly linkages between organizational actions and environmental outcomes; history is difficult to recollect and interpret; and patterns of attention and participation in decision processes are extremely fluid. Decisions, rather than being the outcome of rational processes of bureaucratic procedure or political bargaining, may be the result of a "garbage can process," where problems, solutions, and participants are linked together at a particular point in time by a choice opportunity. Decisions and goals may be emergent, not the direct intention of any of the parties at the outset. And choices may be based on intuition or tradition and faith, rather than on a rational calculus linking consequences to objectives.

Organizational deviations from rational choice may be rooted in basic facts of human cognition and information processing and in the very structure and operation of organizations. Kiesler and Sproull (1982) argue that managers charged with noticing change in the environment and interpreting such change so that decisions are possible are subject to a raft of perceptual, information processing, and motivational biases. Representations of the environment tend to be causally simplistic, outdated, consistent with preexisting beliefs, and resistant to change; thus, the rationality of organizational decision processes is limited by the reigning definition of the situation. Moreover, Starbuck (1983) argues that most of the time organizational action, because of the construction of organizations out of standard operating procedures, is unreflective and nonadaptive. Most actions are not the result of conscious decision processes, but rather flow from action generators, automatic programs of behavior independent of specific stimuli. Societal standards of rationality require that actions be justified, however, and these justifications become solidified and rationalized over time. Organizations need explicit disconfirmation of the rationality of action-justification pairings before they are able to "unlearn." Thus, previous justifications constrain the ability of organizations to engage in thoughtful decision making, and this pathology worsens over time, dooming most organizations to short lives (Starbuck, 1983).

It would seem, then, that prospective rationality does not underlie many actions of individuals in organizations or of organizations themselves. This poses a serious challenge to theories of organization-environment relations that presume adaptive, boundedly rational action on the part of organizations. It is in large measure due to these weaknesses that alternative approaches to the analysis of organizations and their environments have been developed.

**Beyond Dyadic Models:**
**Network Approaches to Interorganizational Relations**

Over the past decade, new approaches combining traditional sociological concerns with
methodological advances in the analysis of networks have emerged and flourished as viable alternatives to dyadic exchange-based models. This work takes a fundamentally sociological approach to the study of organizations by viewing organizations as embedded in networks of social ties that both empower and constrain action. This line of inquiry is motivated less by a concern with organizational design and managerial practice and much more by substantive concerns with the sources of stability and change in social systems.

Network (or structural) approaches are premised on the assumption that structures of interorganizational relations are consequential for understanding the actions of organizations. This contrasts with the atomistic imagery of much of contemporary social science, in which behavior is viewed as adaptive responses to sets of incentives, and social relations are largely irrelevant, at most a drag on efficiency (see Granovetter, 1985, for an extended critique). Although network approaches build on the insights of resource dependence theory, where the most significant aspects of the organization’s environment are other organizations and where resource exchanges confer power and dependence, structural analysis places much greater weight on the means by which an organization’s position in a wider network of relations shapes its actions. Moreover, it is not simply direct relations among organizations that are significant: Both direct and indirect linkages can have an impact on individual and corporate action. Thus, for example, you are more likely to find out about a job opening through a weak tie (someone with whom you are somewhat acquainted but who travels in different circles, such as an old classmate from high school) than from a strong tie (a close friend who associates with most of the same people you do): Close friends are likely to have access to the same contacts and information that you already have and thus provide redundant information, whereas acquaintances are a bridge to contacts you would not have had otherwise (Granovetter, 1973).

Indeed, in some recent work the interorganizational network itself is no longer simply “the environment” of its constituent organizations but an object of study in its own right (Mizruchi & Schwartz, 1987). This reversal of figure and ground is accompanied by a focus on emergent properties of networks. Network systems can be centralized and hierarchical like a bureaucracy, with a dominant organization at the peak (cf. Mintz & Schwartz, 1985), they can be balkanized into multiple more-or-less hierarchical clusters (cf. Roy & Bonacich, 1988), they can be disorganized and even fractious, like a highly competitive industry. Such different structures are significant both for the life chances of constituent organizations and for the explanation of organizational behavior. The implication for organizational theory is clear: Much as an individual caught in a traffic jam knows only the immediate cause of her distress, while someone surveying the scene from a helicopter above has a more comprehensive picture, a view of the organization’s interorganizational environment that extends beyond the immediate context is crucial to this approach.

Conception of the Environment

The general rubric of network analysis embraces a diversity of perspectives on organizational behavior; this work is united more by method than by theory. There is sufficient commonality among the approaches, however, to portray common elements in broad strokes. By definition, a network is composed of a set of relations or ties among actors (in this case, organizations). A tie between actors has both content (the type of relation) and form (the strength of the relation). The content of ties can include information or resource flows, advice or friendship ties, and shared personnel or members of the board of directors; indeed any type of social relation can be mapped as a tie. Thus, organizations are typically embedded in multiple, often overlapping, networks—
resource exchange networks, information or advice networks, board of director interlock networks, and so on.

To the extent that they take a focal organization perspective, network researchers focus either on the set of relations an organization has with those to which it is tied (its ego network) or on its position in the larger network system, often described in terms of its degree of centrality or prestige. Centrality describes the extent to which an actor is tied to many others in the system and (in some versions) the extent to which these others are in turn tied to many others themselves (see Bonacich, 1987). Another way to characterize network position is in terms of autonomy and constraint. Structural autonomy is the ability to pursue actions without constraint from others; firms have high structural autonomy to the extent that they operate in concentrated industries (with limited intraindustry competition) while their buyers and suppliers are competitive among themselves, thus ensuring only limited constraint from external actors (Burt, 1980a). Finally, two actors are structurally equivalent to the extent that they share similar patterns of ties to other actors in the system (e.g., firms in the same industry who have similar sets of buyers and suppliers—see White, Boorman, & Breiger, 1976).

For a variety of reasons, including historical public policy concerns and the wide availability of data, one type of network has received the lion’s share of attention from interorganizational researchers: the interlocking directorate network that is formed by having the same individuals sit on multiple boards of directors. An individual who sits on the boards of two organizations is said to create an interlock tie between those firms. A vast literature can be traced back as far as the turn of the century, when public concerns with trusts and cartels ultimately led to a section of the Clayton Act barring interlocks among competitors in 1914. Because of the position of boards of directors at the very top of the decision-making hierarchy of most organizations, the fact that the same individuals often sit on two, three, or even more boards affords a potential for common control that many have found unsettling. Interlocks among competitors, for example, could be used for collusion, and some have argued that banks control large segments of the economy by placing representatives on the boards of their subject firms (e.g., see Kotz, 1978). Interlocks among direct competitors are rather rare now (Zajac, 1988), and more recent interlock researchers have focused more on overall network position rather than specific ties (Mariolis & Jones, 1982).

Two points are worth noting regarding network research: First, as a relatively new field there remains a good deal of controversy around the merits of various methods and measures and how best to interpret them (e.g., see Burt, 1987 and Galaskiewicz & Burt, 1987, on comparisons of structural equivalence and cohesion as competing explanations for diffusion effects in networks). Second, it is crucial when evaluating this research to keep in mind the content of the network ties being considered. Although many network ties are multiplex (i.e., the ties have multiple contents), this is not necessarily so, and an organization that is central in an information exchange network may be peripheral in a resource exchange network (e.g., a trade association). Thus, network position (such as degree of centrality) is only meaningful in terms of the ties that compose the network. Our trade association has little exchange-based power over its members, which limits its ability to compel actions, but as an information broker it can be crucial for mobilizing collective action: It may be able to persuade, but it cannot force.

Key Areas of Research

Most research on interorganizational networks has proceeded from two perspectives that use similar methods to pursue somewhat different agendas. In the interorganizational
perspective, organizations are the primary actors, and individuals act as agents of these organizations, whereas in the intraclass perspective individuals are the primary actors and organizations are their tools (see Palmer, 1983, and Pfeffer, 1987, for discussions of these contrasting approaches). Common to both these approaches is an overwhelming predominance of empirical research on the interlock network.

Research from the interorganizational perspective has tended to follow the logic of resource dependence theory: Networks of interorganizational resource exchanges confer power and dependence, which in turn motivate organizations to establish ties to other organizations (such as interlocks) to reduce uncertainty. In the aggregate, these ties form a network, albeit a messy and impermanent one, that changes in response to shifts in resource exchange patterns. Interorganizational networks formed through the actions of individual organizations subsequently can be used as mechanisms of diffusion and cohesion, but they are rarely used as devices for the exercise of power, and no enduring power structure emerges to allow ongoing collective political coordination (Glasberg & Schwartz, 1983). Research in this stream tends to focus on the organizational purposes served by interlocks, such as decreased uncertainty and coopetition.

The intraclass approach has focused largely on the role of intercorporate networks in facilitating cohesion among the corporate elite. Interlocks provide a foundation for such order by linking virtually all large corporations into a single dense network, with banks in central (heavily interlocked) positions that correspond to their economically important function of directing the flows of capital (Mintz & Schwartz, 1985). Moreover, the individuals who create the interlocks (multiple directors) are argued to form an inner circle of the corporate elite that is able to act in the interest of this elite in its political dealings, thereby facilitating interorganizational political cohesion (Useem, 1984). Thus, much research in this stream focuses on the creation and maintenance of social order among corporations and the ability of network ties to facilitate collective action.

Although the motivations behind these two lines of research are somewhat different, in practice there is a great deal of commonality in both their methods and findings, and thus we treat them as a single body of work, which we cluster into three areas: (a) research on the formation, maintenance, and aggregate properties of interorganizational networks; (b) effects of networks on organizational structure, ideology, and action; and (c) consequences of networks for organizational effectiveness.

**Formation, Maintenance, and Aggregate Properties of Interorganizational Networks.** Researchers from both interorganizational and intraclass perspectives have focused much attention on the conditions under which organizations form and maintain interlocks. The transformation of the American economy from entrepreneurial to corporate at the end of the 19th century was accompanied by the elaboration of an interindustry interlock network, with the railroad, coal, and telegraph industries forming an early and enduring core among industrials; ties across industries became increasingly dense but maintained a spiked wheel pattern, with core industries tightly interlinked and peripheral industries tied to the core but not to each other (Roy, 1983). Roy found that ties across industries seemed to facilitate both resource exchange and ownership relations. In a longitudinal study, Mizruchi and Stearns (1988) found that firms were more likely to appoint representatives of financial institutions to their board when solvency and profitability were declining and when the demand for capital corresponded with macroeconomic conditions such as declining interest rates or a contraction stage of the business cycle, suggesting an
interplay between the formation of intercorporate networks and larger economic forces. Burt (1983) found that both direct interlocks and indirect financial interlocks (i.e., where members of two boards are both members of a financial institution's board) as well as common ownership of establishments trace market-based constraints on profitability. And Lang and Lockhart (1990) found that after the onset of deregulation in the airline industry, which increased competitive uncertainty, airlines focused their indirect interlocking more on direct competitors than they did before deregulation.

A second set of studies is premised on the argument that particular interlocks may be created for various reasons that have little to do with organizational purposes, such as social ties between the CEO and the board member, but that those which serve organizational purposes are likely to be reestablished when they are accidentally severed (e.g., the multiple director dies or retires; Palmer, 1983). Thus, studying the factors associated with the reconstitution of interlock should shed light on the purposes they serve. By and large, firms do not rush to reconstitute accidentally broken ties: Only about one out of six were repaired among Fortune 500 firms in the mid-1960s, suggesting that these ties are rarely used as vehicles of formal coordination between firms (Palmer, 1983). Interlocks among large American corporations in the 1960s were more likely to be reconstituted when the firms involved were linked by some mechanism of formal coordination and were headquartered in the same city, and where at least one of the firms' interlocks involved an executive of one of the firms (Palmer, Friedland, & Singh, 1986). Using a broader time frame for reconstitution and a sample that included all disrupted ties (rather than only accidentally broken ties), Ornstein (1984) found a much higher rate of reconstitution among Canadian firms; reconstitution was more likely when the two firms had multiple shared directors, when at least one tie was created by an executive of one of the firms, and when the firms were partly owned either by a third party or by each other. Finally, management-controlled firms were more likely to reconstitute ties, and ties to financial institutions were much more likely to be reconstituted, whereas long-lived ties (which are more likely to have outlived their usefulness) were actually less likely to be reconstituted, according to a long-term study of interlock reconstitution (Stearns & Mizruchi, 1986). Stearns and Mizruchi introduce the notion of functional reconstitution, where a firm replaces a broken tie with a tie to a different firm in the same industry. These authors found the factors that affected direct reconstitution of a broken tie to a financial institution differed from those affecting functional reconstitution; they argue this difference implies that direct reconstitution is more likely to reflect interorganizational power relations.

The centrality of particular firms in the overall network has also received attention. Thompson (1967) argued that organizations match the elaborateness of their boundary-spanning element to the complexity of their environment. Thus, New York banks tend to be the most well-connected business organizations, as the direction of the economy as a whole is implicated in their investment decisions (Mintz & Schwartz, 1985). Large industrial firms such as AT&T, General Electric, and IBM also cast their interlock nets broadly and are therefore better able to gather information about their environments (Davis, in press). Moreover, while particularities may come and go, as the reconstitution research indicates, overall centrality is remarkably stable (Mariolis & Jones, 1982), implying that there is an order to the intercorporate network.

At the level of the network as a whole, researchers have uncovered several regularities that characterize the network of interlocks among the largest American corporations. In a sweeping study of the power structure of business in the 1960s, Mintz and Schwartz (1985)
observed that a handful of major New York banks and insurance companies formed a stable core at the top of the interlock network. They argue that financial institutions do not use their network centrality to control other corporations directly, as previous commentators have suggested, but primarily to gather economy-wide information that guides their investment decision making. Corporations are subject to a unique dependence on capital, and thus the flow of capital both creates opportunities and places constraints on what is possible for large corporations. By directing capital flows to some areas and withholding it from others, banks bring a hegemonic order to economic life, and the interorganizational "war of all against all" that economic theories lead us to expect does not appear.

Apparently, however, the current interlock network structure did not arise fully formed: Clusters of interlocks among railroads, which were the largest and most powerful nonfinancial corporations at the turn of the century, mapped onto several balkanized communities of interest linked by common ownership rather than forming a single hierarchical network. This finding argues that interest groups based on common ownership (e.g., by the Morgan or Rockefeller families), apparently uncommon today, formed a basis for the earliest interlock networks (Roy & Bonacich, 1988).

Finally, Burt (1988) analyzed input-output tables for 77 broad industry categories to find that the boundaries and degree of structural autonomy of markets (defined by patterns of exchanges among industries) were highly stable throughout the 1960s and 1970s and accounted for enduring inequalities in profit margins in these industries.12

Effects of Networks on Organizational Structure, Ideology, and Action. In contrast to the large body of work describing the factors influencing the construction and maintenance of network ties, relatively little work has been done on the effects of interorganizational ties. More recent studies have sought to remedy this gap by examining the effects of network ties on the diffusion of organizational structures and actions and the political behavior of large corporations.

In perhaps the only study documenting network effects on changes in major aspects of organizational structure, Palmer, Jennings, and Zhou (1989) found that firms were more likely to adopt the multidivisional form (described above in the section on transaction cost economics) when they had ties to other firms that had already adopted one. Nonfamily-owned firms in the Twin Cities contributed more to local charities when the CEOs had network ties to the philanthropic elite (Atkinson & Galaskiewicz, 1988), and these organizations seemed to emulate the contributions of firms whose executives had ties to the organization's boundary spanners (Galaskiewicz & Wasserman, 1989). And Davis (1991) found that large corporations were quicker to adopt a poison pill takeover defense to the extent that they were interlocked with firms that had already adopted one. Thus, network centrality appeared to be self-reproducing: More central firms were able to gain early access to information about protective strategies that flowed across network ties, thus maintaining their centrality in the face of the threat of takeover. The effect of network ties on corporate political ideology and behavior has received substantial study since Useem's (1984) explanation of the inner circle thesis. Political sociologists have debated for decades whether corporations are able to form a united front in the pursuit of government policy or whether they are generally fractious and competitive and therefore unable to form an enduring coalition. Useem (1984) argued that directors who sit on multiple boards are uniquely placed to overcome the differences dividing corporations so that these corporations are able to pursue common political goals. Thus, more central firms made more PAC contributions
to incumbents and less to conservatives in the 1980 elections, but they were more likely to be involved with conservative policy organizations, suggesting that PAC contributions are used to further corporate interests but that policy organizations are used to pursue collective interests (Clawson & Neustadt, 1989). Mizruchi and Koenig (1986; Mizruchi, 1989) found that interindustry constraint increased the degree of similarity of PAC contributions made by firms in those industries, and indirect interlocking through financial institutions as well as common ownership by financials increased the similarity of PAC contributions, which argues that interorganizational ties facilitate political cohesion. Political cohesion among organizations does not necessarily mean that these organizations are more effective at getting what they want, however: Earlier in this century, industries with more extensive network ties to economic, political, or social organizations were no more effective than other industries at having their interests taken into account by the State Department (Roy, 1981). This finding, however, is likely to reflect its time setting (cf. Laumann & Knoke, 1987).

Consequences of Networks for Organizational Effectiveness. Interorganizational ties bear a complicated relation to different aspects of organizational effectiveness. According to Burt (1983), an industry’s profitability is strongly related to the degree of exchange-based constraint it faces but has very little relation to the interlock ties it maintains with other industries. Thus, organizations seemed to direct their interlock ties toward their most severe constraints on profitability, but these ties did not provide any obvious profit advantage. Similarly, in a more recent study of Canadian firms in the mid-1960s, Richardson (1987) found that the total number of ties between financials and nonfinancials had no effect on profitability. However, having ties that had been broken and replaced was positively related to profitability when the broken tie and the replacement were created by executives of the nonfinancial but somewhat negatively related when both ties were created by executives of the financial. Richardson also found evidence that profitability leads to replaced ties and not the other way around, which is consistent with the idea that more profitable firms are better able to maintain representation on financials’ boards, whereas less profitable firms may have to submit to having financial representatives on their boards.

Wiewel and Hunter (1985) discovered that the liability of newness (i.e., the higher failure rates faced by new organizations) was partially ameliorated by network ties to previously existing similar organizations, which can provide experienced insider personnel, access to credit, external legitimacy, and so on.

Finally, in what is undoubtedly the most comprehensive study of interorganizational networks to date, Laumann and Knoke (1987) examine the network structures of two national policy domains, energy and health, during the Carter years. They proceed from the premises that organizations are the most important actors in governmental policy formation and that the network of ties among these actors is crucial for explaining their participation and influence in policy events. This path-breaking study introduces such a large battery of new concepts and theory that no short summary would suffice. Its most distinctive contribution, however, is its conceptualization of both actors and events (in this case, events in the process of national policy decision making) having both individual properties and relational properties that aggregate in ways that are consequential for policy outcomes. Organizations are linked to other organizations by ties of information transmission, resource transactions, and boundary penetration; events are linked by their ordering in time and their institutional setting as well as by their similarity to each other; and organizations are linked to events in
an action system, which has systemic properties that arise from the properties and relations of the constituent organizations and events. Laumann and Knoke combine this framework with exotic statistical techniques to analyze the individual and network factors determining organizational participation and effectiveness in national policy making. This study sets a new standard of theoretical and methodological sophistication for organizational research.

Critique. Network approaches to interorganizational relations have provided an excellent ground for combining research on organizations with more traditional sociological concerns, such as problems of social order and the role of societal elites. Research in this area has made the most of methodological advances; indeed, network methods have far outstripped network theory (cf. Burt, 1980a), which is the greatest weakness of the research reviewed here. The formation, maintenance, and mapping of network ties has received a great deal of attention, but comparatively little work has focused on whether an organization's position in various networks in fact has any significance for understanding its behavior. Much of the focus of network researchers on interlocks among boards of directors can be attributed to the extreme data demands that network methods make, coupled with the ready public availability of data on boards of directors. But increasingly sophisticated statistical techniques continue to uncover rather modest findings.

A curious irony of network research is that despite its imperative to focus on the causal importance of structures of relations among actors rather than simply the properties of those actors, the research itself tends to treat network positions as properties themselves. Thus, studies often treat interlock network centrality as if it were a feature of an organization like size. But as we mentioned early in this section, centrality only has significance in terms of the ties from which it is derived. Given that only a small minority of interlock ties seem to trace enduring social relations among firms, it is not always clear what content these ties have and, therefore, what an organization's position in the interlock network signifies.

The obvious remedy for the apparent primacy of method over substance in network research is to bring the content of the ties, rather than merely the structure formed by these ties, back in. Social ties among organizations can be consequential, but not all of them need be. Stinchcombe (1990) suggests that the dynamic and causal theory of a structure has to be built into the analysis of the links.

We need to know what flows across the links, who decides on those flows in the light of what interests, and what collective or corporate action flows from the organization of links, in order to make sense of inter-corporate relations. (p. 381)

More recent network research, including much of the work reviewed here, has taken steps to remedy this weakness. The notion that board interlocks are used by competitors for collusion, which dates back to the earlier part of this century, has little evidence to support it today (cf. Zajac, 1988), and few researchers cling to the belief that interlocks are used by corporations such as banks to exercise direct control over their hapless corporate stooges (see Glassberg & Schwartz, 1983, for a critique). Instead, the most sophisticated work today sees the interlock network as a mechanism for the diffusion of information rather than for the exercise of explicit control (Davis, 1991; Mintz & Schwartz, 1985; Useem, 1984). The future of network approaches seems to lie with combining the methodological sophistication they have brought to the study of interorganizational relations with the substantive concerns of other areas of organizational sociology.
The Organization of the Environment: Ecological and Institutional Perspectives

In contrast to dyadic or network models of organization and environment relations, ecologists and institutionalists pay much less attention to the efforts of organizations to manage and control their environments. The central thrust of ecological and institutional thinking is on the structure and composition of the environment. Most ecological and institutional research maintains that organizational change is largely shaped by changes in the environment (through population-level demographic processes of organizational foundings and death or through broad social changes promulgated by the state or sovereign professions). Consequently, these approaches suggest that change in individual organizations contribute considerably less to large-scale social transformations.

Not surprisingly, then, a common criticism of ecological and institutional research is their respective inattention to organizational change and adaptation. In large part, these criticisms are fair. Both ecologists and institutionalists emphasize the structural inertia that besets established organizations, the manner by which organizational structures and practices become valued for their own sake and organizational policies become locked in. To the extent that ecologists and institutionalists emphasize external environmental conditions and constraints, they downplay organizational innovation and adaptation. Similarly, organizational politics are not primary concerns of ecologists and institutionalists.

Initially, ecological and institutional approaches were seen as competing theories, but as the common criticisms noted above suggest, there has been a marked convergence of these two perspectives (see the discussions in Hannan & Freeman, 1989, and Powell & DiMaggio, 1991; also see empirical studies by Barnett & Carroll, 1990; Carroll & Huo, 1986; Hannan, 1986; Singh, Tucker, & Meinhard, 1991). In several key respects the two approaches are more alike than different. Both share an appreciation for the fact that history matters a great deal and both attempt to connect organizational theory with topics in general sociology. Both are increasingly animated by related questions: How do changes in institutional environments influence the survival of organizations? How do ecological processes contribute to fundamental changes in the institutional order? Still, as the reader will soon note, there are considerable points of divergence in key concepts and methods: Organizational ecologists borrow freely from population biology and have made important contributions to mathematical sociology, while institutionalists are closely linked to research on the professions and the state and have expanded the horizons of sociologists of culture.

Ecological Perspectives

The animating question for ecological research was stated by Hannan and Freeman (1977, paraphrasing the ecologist Evelyn Hutchinson). "Why are there so many kinds of organizations?" The core of the ecological approach is an effort to explain the diversity of organizational forms across the social landscape and to account for changes in the mixture of forms. This *population ecology of organizations* (Hannan & Freeman, 1977) starts with the observation that the variety and mix of organizational forms in society is fundamentally a property of aggregates of organizations. *Mix* has no analogue at the level of the individual organization (Hannan & Freeman, 1986); thus, any explanation of organizational diversity must be posed at a higher level of analysis. Moreover, whereas adaptive approaches would explain the mix of organizational forms as the result of choices by previously existing organizations, ecologists point out that diversity may also result from deaths of old, forms and births of new ones.
Thus, building on Hawley's (1950, 1968) theory of human ecology, researchers began to study dynamic processes at the population level of analysis. The fruitfulness of this approach is evidenced by burgeoning empirical studies adopting this perspective. Organizational ecology can take three levels (Carroll, 1984a): (a) an organizational level, which focuses on the demography and development of individual organizations; (b) a population level, which focuses primarily on selection processes; and (c) a community level, which emphasizes macro-evolutionary processes. While the developmental approach has spawned some extremely interesting work (we would place Langton, 1984, and Nelson & Winter, 1982, under this rubric), the ecological literature has been dominated by studies adopting the population level of analysis and, to a much lesser extent, the community level.

The Language of Organizational Ecology. Organizational ecology has introduced an entire battery of new concepts and language for describing organizations and their environments. Perhaps the most basic notion is that of a population of organizations. Intuitively, a population is simply a class of organizations facing similar environmental vulnerabilities (Hannan & Freeman, 1977), usually organizations sharing the same form. Form can be defined by internal attributes of organizations, such as the organization's blueprint for action (indicated by the formal structure or patterns of activity), or by the set of (external) relations and dependencies the organization has with its environment (Hannan & Freeman, 1989). Organizations sharing a common form constitute a population when they are bounded within a common system, usually defined by geography, political boundaries, or markets.

The shift to a population level of analysis is accompanied by a shift in the locus of causality, from the rational and adaptive organization to the environment. The economic theory of the

firm sees organizational decision makers as optimizing, but "from a population ecology perspective, it is the environment which optimizes. Whether or not individual organizations are consciously adapting, the environment selects out optimal combinations of organizations" (Hannan & Freeman, 1977, p. 939-940). Because resources such as people and money are limited, societies have a limited carrying capacity for organizations, and thus under many conditions organizations are engaged in a struggle for existence against others in their niche (i.e., those drawing on the same pool of resources, such as members of the same industry). In equilibrium, the population that survives in a niche will be the one that is "isomorphic" to the environment, whereas "that population with the characteristic less fit to environmental contingencies will tend to be eliminated" (Hannan & Freeman, 1977, p. 943). Over time, ecologists have softened their emphasis on optimization and fit in favor of more diffuse selection models (e.g., Hannan & Freeman, 1989).

The use of organizational models based on selection processes has triggered controversy. Much of this controversy stems from a misunderstanding of the uses of evolutionary theory in the social sciences. Some readers assume that evolutionary models imply either progress or superior fitness. But ecologists do not claim that selection logic implies that the organizations that have survived are more efficient or more deserving of their success. Indeed, selection models are always built on the assumption of the importance of randomness of success.

Framework of Assumptions. Ecological models of organizations obviously owe a great intellectual debt to models of biotic evolution. The traffic in intellectual technology between population biology and organizational ecology has required a rethinking of many of the assumptions previously held by organizational scholars. The basic task is to justify the implicit
and sometimes explicit analogy between (populations of) organizations and (populations of) organisms. Points of weak fit are apparent: Organizations can change their structure, organisms cannot; information is carried through nongenetic means within and among organizations, and individual organizations can expand virtually without limit, whereas a mouse cannot grow into an elephant (Hannan & Freeman, 1977). Ecologists have mustered intriguing arguments in favor of their approach, however, and these arguments have led to a revitalized debate over some of the basic assumptions of organizational analysis. The critical theoretical issues have concerned the degree of plasticity of organizations, or just how malleable organizations are, the extent to which a population is an appropriate level of analysis, and how organizational forms can be defined.

Hannan and Freeman (1977, 1984) argue for the proposition that organizations in the modern world experience strong pressures to retain their form over time rather than engage in structural change. Internal constraints on adaptation include investment in plant and equipment, informational limits, intraorganizational politics, and the institutionalization of organizational routines, while external constraints include barriers to entry and exit and legitimacy concerns. In addition, organizations are selected by the environment in the first place based on their ability to perform reliably and to account rationally for their actions. But this reliability and accountability demand that the organization's structure be reproducible from day to day, which in turn generates strong inertial pressures. Thus, structural inertia is a consequence of generalized selection pressures in society that favor reliable and accountable organizations over other types of collective actors. Structural inertia increases with the organization's age and size, while in most instances death rates decrease with both (Hannan & Freeman, 1984). Consequently, ecologists argue that the variation in observed organizational forms is attributable to the deaths of relatively stable, inert organizations and the births of new organizational forms, rather than to wholesale adaptation by existing populations. If we accept this argument, then ecological models have fairly general applicability to organizational life in the modern world.

A second issue concerns the theoretical usefulness of the population concept. The notion of a group of organizations sharing common form and similar environmental vulnerabilities is intuitively appealing, yet it has proved difficult to operationalize in a satisfying way. Baron and Bielby (1980) point out that organizations may differ greatly in structure even when they are of similar size and age and operate within the same industry and locality. This raises the question of how organizations should be classified, a crucial concern for taking a population perspective. McKelvey (1978, 1982) has approached this issue by drawing on taxonomic and specie concepts from biology. Organisms pass down specie characteristics through the gene pool. The organizational equivalent of the gene pool is the tech pool, composed of the characteristic operational technologies and managerial technologies shared by members of a group of organizations. An organizational specie is then defined as a group that shares dominant competencies, but that is sufficiently isolated from other populations by the fact that these competencies are not easily learned or transmitted across groups (McKelvey, 1982, p. 192). This isolation may be due to geographic or cultural factors. For example, the geographical isolation of firms operating in the same industry in Silicon Valley promotes personnel movement among Valley firms but isolation from firms in other industries, thereby maintaining the integrity of the population.

More recently, Hannan and Freeman (1986) have argued that rather than classifying organizations into populations based on shared formal characteristics or common patterns of
environmental dependency, a more fruitful approach may be to focus on the processes that create, sustain, and erode boundaries around populations. This would provide a first step in identifying the structure of a niche as well as the forms that occupy it. "Instead of beginning with problems of classification, this approach begins with the question: Where do organizational forms come from?" (Hannan & Freeman, 1986, p. 60). They suggest that segregating processes tend to create the conditions for greater similarity of forms within a population, while blending processes tend to blur the distinctions between forms in different populations.

Most current macro researchers are reluctant to develop an all-purpose definition of a population or an organizational field. Instead, they argue that the definition of an organizational population should be determined by the theoretical or substantive problem. 9 Hence, concerns with classification should be secondary to identifying the boundaries around forms (Hannan & Freeman, 1986).

Key Areas of Research. Organizational Birth. Stinchcombe’s (1965) treatment of several issues associated with organizational foundings has critically informed ecological research on organizational births. We consider two questions informed by Stinchcombe’s classic paper. First, how do prevailing social conditions affect the characteristics of organizations founded during a particular historical time period? Second, how do variations in the social environment affect the rate of organizational foundings?

Organizational characteristics at founding. Stinchcombe explained the peculiar fact that organizations founded at roughly the same time tend to be structurally similar to each other and dissimilar to those founded at other times by arguing that organizations are imprinted at the time of their birth and reflect the prevailing social technology of the day. One of the crucial assumptions of the ecological school is that these imprinted structures, once in place, are relatively inert. Aldrich and Mueller (1982) trace the origins of different types of organizational forms that appeared in the U.S. economy to broad-scale historical changes in the environment. They identify epochal transitions in the forms of organization that dominate the social landscape—the movement from prefactory to factory production in the early to mid-19th century, from competitive to monopoly capitalism in the late 19th and early 20th century, and from early to mature monopoly capitalism after World War I—and relate these to variations in the availability of capital, materials, labor, and infrastructure as well in the role of the state.

Tushman and Anderson (1986) show how the evolutionary logic of technological change affects the environments of organizations and the structure of industries. They argue that, within particular product classes, technological innovations that build on pre-existing competences will consolidate the positions of organizations that already dominate an industry and increase the barriers to new entrants, whereas in the rarer case of radically new technologies, such as biotechnology or superconductivity, space will be open to new organizations that are not trapped by sunk costs and skills tied to older technologies and that therefore can exploit innovations (see also Abernathy & Clark, 1985; Powell & Brantley, 1991). They found that in the airline, cement, and minicomputer industries, new entrants were more likely to initiate radical technological breakthroughs than existing firms, suggesting that a new technological regime provides both the opportunity and the means for new organizational forms to be founded based around that technology. Finally, Boeker (1988) found that the initial strategy chosen by new organizations in the semiconductor industry reflected both the functional background of the entrepreneur who founded the organization and the period in which it was founded, consistent with the imprinting hypothesis.
Rates of founding. Stinchcombe (1965) argued that certain social conditions, such as revolutions, quicken the pace at which new organizations are brought into existence by shifting social alignments and freeing resources that can be used by entrepreneurs. Of course, more pedestrian variations in the environment also affect founding rates. Ecological researchers have studied how both environments and characteristics of the population itself alter the incidence of new organizations. Pennings (1982) studied the environmental conditions that stimulate organizational births in three different industries across 70 American urban communities, finding that the factors fostering organizational fecundity differed substantially between manufacturers of plastic products, electronic components, and telecommunications equipment. McCarthy, Wolfson, Baker, and Mosakowski (1988) discovered that local citizens' organizations opposing drunk driving, such as Mothers Against Drunk Driving, appeared earlier in counties with higher education levels, greater population density, more government resources, and a higher level of grievances—thus supporting the view that social movement organizations depend primarily on the existence of human resources that can be mobilized for social action. On the other hand, in their study of trade associations, Aldrich and Staber (1988) contend that the pattern of growth of this population does not seem tied to environmental characteristics, such as a fluctuating demand for their services due to changing levels of government regulation, but rather to the diffusion of the trade association form across industries. Together these studies suggest that the existence of a social infrastructure may be a necessary but not sufficient condition for organizational founding and that processes at the level of the organizational population could play a part.

Indeed, Delacroix and Carroll (1983) show that the cycles of newspaper foundings in Argentina (1800–1900) and Ireland (1800–1925) reflect both political turbulence and population dynamics. They posit that the death of existing organizations frees resources with which to found new organizations so that the number of newspaper deaths in the recent past should increase rates of founding up to a point; high levels of death, however, signal entrepreneurs that the environment is inhospitable and thus depress birth rates, implying a curvilinear relationship between prior deaths and births. Prior foundings should also have a curvilinear effect: Up to a point, an increasing birth rate signals a munificent environment, encouraging entrepreneurial imitation and increasing the subsequent rate of foundings, but very high levels of founding will use up the resources needed for starting new organizations and will thus decrease subsequent foundings. Data from both nations supported this model. In addition, years of political turbulence were marked by an increase in the number of newspapers founded in both Argentina and Ireland. In a similar analysis of the newspaper industry in the San Francisco Bay area throughout its history, Carroll and Huo (1986) found only weak support for the hypothesized relationship between prior and subsequent foundings, but again found that in years of political turmoil the founding rate was higher than in calmer years, a finding consistent with Stinchcombe’s (1965) argument that turmoil creates or frees resources for organizational founding.²⁰

Much research has also been done on demographic effects on organizational populations. Hannan and Freeman (1987) argue that a population’s density (i.e., the mere count of organizations currently in a population) can have two effects on subsequent founding rates. Positive density dependence occurs when the number of subsequent foundings increases with the size of the population. Density increases founding rates by (a) increasing the number of organizations of a particular form, who can then use their experience to create similar organizations, and (b) increasing the legitimacy of an organizational form, as the form comes to
be taken for granted simply through its prevalence. Negative density dependence occurs because a larger population will experience increased competition for limited resources, thus depressing founding rates. Positive density dependence is argued to dominate when the population is smaller, while negative density dependence dominates in larger populations (Hannan & Freeman, 1988a). The combination of these two effects yields an inverted U-curve relation between density and founding rates. Using the population of national labor unions founded between 1836 and 1985, Hannan and Freeman (1987) tested their model of density dependence as well as Delacroix and Carroll’s (1983) hypothesis of entrepreneurial imitation, finding significant results for all the hypothesized effects. They argue that density dependence implies that the population faces a carrying capacity, that is, that there is a rough empirical limit to the population size that the environment will support. A similar analysis for founding rates in the semiconductor industry (Hannan & Freeman, 1988a) found support only for positive density dependence rather than the curvilinear effects found for unions, which the authors attribute to this industry’s relative growth and its expanding markets (which imply an environment with an expanding carrying capacity).

Organizational Death. The bulk of ecological research thus far has examined rates of mortality among organizations. Organizational death is pervasive: There were over 70,000 business bankruptcies and almost 400,000 business deaths in 1985 (Aldrich & Marsden, 1988). Death is an unequivocal indicator of organizational performance, and it is also the means by which environmental selection operates. Thus, relating mortality rates to characteristics of organizations, populations, and their environments is central to organizational ecology.

Four broad areas of research on organizational mortality can be distinguished: liabilities of age and size, specialization and generalism, internal crisis and transformation, and environmental and population characteristics. All have been studied in the context of organizational death rates. We will consider each in turn.

Liabilities of age and size. One of the truisms of organizational research is that an organization’s likelihood of failure is considerably higher during its early years than later in its life cycle. Stinchcombe (1965) attributed this liability of newness to both internal and external features of younger organizations. Members of newly founded organizations, especially those possessing a new type of structure, must learn new roles and relationships as the organization’s operations get established. This process takes time away from the organization’s “real” business and leaves it more vulnerable to failure. To compound this vulnerability, new organizations lack the external legitimacy of older organizations as well as the stable relationships with environmental constituencies that older ones have, making it more difficult to attract members and sources of support. Both these internal and external features lead to a higher death rate during the initial years.

The liability of newness hypothesis, that organizational mortality declines with age, has rapidly become perhaps the best-documented regularity in ecology. In an ambitious effort, Carroll (1983) compared three different statistical models of death rates across 52 archival datasets covering populations of retail stores, manufacturing firms, and craft, service, and wholesale organizations, finding that death rates decline with age in quite diverse organizational populations. The liability of newness hypothesis has also found support in populations of newspapers in Argentina and Ireland (Carroll & Delacroix, 1982), national labor unions and semiconductor manufacturers (Freeman, Carroll, & Hannan, 1983), newspaper organizations in seven American metropolitan areas (Carroll, 1984b), voluntary social
service organizations in metropolitan Toronto (Singh, Tucker, & House, 1986), and telephone companies in early 20th century Iowa (Barnett & Carroll, 1987). Caution is called for in interpreting these findings, however—as Freeman, Carroll, and Hannan (1983) point out, apparent age dependence could be a statistical artifact if the researcher has not been careful to control for heterogeneity in the population. If some types of organizations within a population are born feebler than others but this is not measured and controlled, it will appear as if mortality rates decline with age when what is actually happening is the early death of feeble organizations and the survival of their more robust counterparts.22

One liability that is likely to be confounded with age is smallness. Aldrich and Auster (1986) argue that compared to larger organizations, small organizations have more difficulty and pay greater costs in raising capital, face tax laws that favor selling out through merger, experience a proportionally greater load of paperwork from government regulation, and face tough competition from larger firms in securing labor, all of which increase the likelihood of death through dissolution or merger. Newer organizations tend to be smaller, and thus the liability of newness could merely reflect the liability of smallness.23

Despite the considerable evidence supporting the liability of newness hypothesis, relatively little effort has been made to unpack the ingredients of age to determine which factors are effective in warding off organizational death.24 An exception is the work of Singh, Tucker, and House (1986), which asks to what extent external factors such as lower legitimacy and weak exchange relations are responsible for the liability of newness in a population of voluntary social service organizations. They found that organizations with a high degree of support and legitimacy experienced lower death rates that declined over time, while rates for organizations with less external legitimacy did not decline. Hence, in the social service sector in Toronto, aging alone is not sufficient to overcome the liabilities of newness. This supports the notion that it is not newness per se that poses such a hazard for young organizations but rather other features, such as size, that are correlated with age.

Specialism and generalism. What kinds of environmental conditions favor specialist organizations, and when are generalists more likely to prosper? Ecologists argue that the answers to these questions can be used to explain the relative prevalence of specialist and generalist organizations, as selection pressures move populations over time in the direction of structural isomorphism. Freeman and Hannan (1983) adapted Levins’ (1968) fitness-set theory to these questions, arguing that the issue of form can be reframed as one of niche width. A population’s niche width is defined by its tolerance for changing levels of resources, its ability to resist competitors, and its response to other factors that inhibit growth. A population which has wide tolerance, meaning it can reproduce in diverse circumstances, is said to have a broad niche. Populations with more limited ranges of tolerance are called specialists... Specialist populations follow the strategy of betting all of their fitness chips on specific outcomes; generalists hedge their bets. (Freeman & Hannan, 1983, pp. 1118–1119)

A generalist must have some slack so that it can adapt to changes in the environment, while specialists tend to be leaner. Whereas adaptive theories generally argue that variable environments will favor generalists, according to niche width theory this is only true when changes are relatively infrequent, allowing generalists time to readjust; specialists are in a better position when fluctuations are both large and frequent. However, an analysis of death
rates of specialist and generalist restaurants in 18 California cities over a three-year period failed to support this particular hypothesis (Freeman & Hannan, 1983).

An alternative approach to modeling the dynamics of specialist and generalist populations is proposed by Carroll (1985, 1987) in his model of resource partitioning. He argues that generalism and specialization are fundamentally interrelated: The success of generalist organizations creates the conditions for the success of specialists. Specifically, when a population of generalist newspapers becomes dominated by a handful of firms, there will be more resources available for specialist papers that target specific audiences (e.g., ethnic or professional groups) who cannot be catered to as effectively by a large dominant daily paper. Whereas early in the history of the newspaper industry a city may have had several competing general interest dailies that appealed to different groups, the trend toward greater concentration over time shifted the dominant strategy for generalists to one of playing to the middle, leaving groups with more specific concerns to the specialists. Data on newspapers in seven American metropolitan areas supported this model: When resources were more concentrated among general interest newspapers, specialists enjoyed a lower mortality rate than specialists operating in less concentrated markets, suggesting that the internal dynamics of the population of generalists leads over time to concentration in this population, opening niche space for specialists.  

**Internal crisis and transformation.** Reorganization can be traumatic for organizations, and crises—whether anticipated or not—can be fatal. Hannan and Freeman (1977, 1984) argue that organizations that attempt to adapt to their environments by reorganizing their activities or restructuring themselves face two sorts of difficulties: First, planned change can be difficult to accomplish because of various sunk costs, political considerations, and the institutionalization of structures and procedures; and second, organizations that manage to change dramatically may recreate the liability of newness, as the procedures and role structures painstakingly worked out through trial and error are swept aside. While an organization may retain its external legitimacy in the wake of reorganization, its internal structure will no longer reflect its accumulated history and thus will be robbed of its previous survival value, increasing the chances of the organization’s failure.

This model of the effects of organizational change was compared with an adaptive model, which predicts that change will increase the organization’s viability, and a model of random organizational action, which suggests that planned change will not have a consistent effect either way (Singh, House, & Tucker, 1986). This study found that different types of changes had different effects in a population of voluntary social service organizations: Changes in structure and goals, which Hannan and Freeman (1984) posit as particularly disruptive, had no effect; changes in service area made early in life were associated with an increased risk of death; and early changes in chief executive and physical location were both linked to decreased death rates, consistent with an adaptive interpretation. On the other hand, Carroll (1984b), in a test of the succession-crisis hypothesis, found that managerial succession (in this case, the departure of the first publisher) was followed by a jump in the death rate of newspapers founded between 1800 and 1975 in seven American cities. This event was particularly precarious when the publisher who left was also the editor. These contrary findings on the effects of executive succession may be due to the different roles played by these different leaders: Founding publishers of newspapers are likely to be highly committed, intimately tied to the everyday activities of the newspaper, and personally tied to external sources of support, whereas executives who depart early in the career of a voluntary organization may
have experienced a poor fit with the organization, making their replacement an adaptive change. A revised position suggested by Scott (1987a, pp. 200-203) that could reconcile the seemingly inconsistent results on the effects of change would view changes in core features of the organization (such as the fundamental mission and values of the organization) as more problematic, consistent with the ecological or selection view, while changes in peripheral aspects of the organization (such as short-run strategies) are best described by an adaptive perspective.

**Environmental and population characteristics.** Surprisingly, relatively little work has been done concerning the effects of environmental variations on the death rates in populations of organizations. Carroll and Delacroix (1982) found that increased national economic development at the time of the organization's founding enhanced the expected life-span of newspapers in Argentina and Ireland, although apparently at a decreasing rate. They also found that Argentine newspapers born during years of political turmoil had lower life expectancies than those organized during calmer years. They speculated that such organizations are likely to be opportunists who thrive on the resources that are freed in periods following social disruption but are then out-competed in more stable resource regimes. Carroll and Huo (1986) analyzed how task and institutional environments exerted different selection pressures on the population of newspaper organizations in the San Francisco-Oakland-San Jose area. This study again found that newspapers born during political turbulence were shorter lived; the costs of raw materials and the rate of illiteracy in the populace were associated with an increased hazard of death, while population size and the density of industrial establishments in the area apparently promoted longevity. They concluded that the institutional environment is more important in explaining the birth and death rates of newspapers, whereas the effects of the task environment are more strongly felt on organizational performance.

Population-level processes as well as external environmental processes affect the mortality rates of organizational populations. The effects of density on death rates mirror those described above for birth rates: According to Hannan and Freeman (1988b), a greater population size increases a population’s legitimacy and capacity for political and legal action, thereby decreasing death rates, but at higher densities competition for limited resources is more acute, thus increasing death rates. The combination of these two opposing effects is again a curvilinear relation between density and mortality: The disbanding rate for organizations falls with increasing density to a point corresponding roughly to the population’s carrying capacity, then rises with density after that as competitive effects prevail. Analyses of populations of American national labor unions and semiconductor manufacturers supported this hypothesis (Hannan & Freeman, 1988a, 1988b). Similar effects of density dependence have been found by Barnett and Carroll (1987) for Iowa telephone companies. Singh, House, and Tucker (1986) found density to be negatively associated with the hazard of death, although apparently this is not a U-curve relation, perhaps indicating that the carrying capacity for this population has not been reached. As the only set of empirical regularities uniquely attributable to ecological research, the effects of density on organizational birth and death rates are certain to receive continued attention from ecologists (cf. Barnett, 1990; Carroll & Hannan, 1989a, 1989b; Delacroix, Swaminathan, & Solt, 1989).

**Community Dynamics.** The discovery of complex and significant relations among population density and founding and death rates, discussed above, argues for the importance of taking the emergent properties of populations into account. That characteristics of one population have been found to have patterned
impacts on those of another argues for taking the next step to the community level of analysis (Astley, 1985). The analysis of community dynamics has attracted growing attention. Brittain and Freeman (1980) describe the semiconductor industry in terms of the uncertainty, grain, and compatibility of resource states that characterize various niches within it, as well as the density of populations occupying them. They argue that r-strategists, who apply a first mover strategy by moving into new niches as they open, thrive under conditions of resource uncertainty and frequent innovation, while as the number of organizations occupying a niche increases toward its total carrying capacity, K-strategists, who thrive on efficiency, will tend to out-compete their opportunistic predecessors. Carroll (1985) found the degree of concentration of resources among generalist organizations to be related to increased life expectancies for specialists operating in the same environment. Hannan and Freeman (1987) studied the cross-effects of craft unions and industrial unions, finding that craft union founding rates decline with the density of industrial unions, but that surges in the number of industrial union foundings increase the rate of birth among craft unions. Characteristics of the craft union population had no discernible effect on the much smaller population of industrial unions.

In a study of Iowa telephone companies, Barnett and Carroll (1987) found that two subpopulations—(a) commercial, usually urban companies and (b) mutual (cooperative), usually rural companies—had complex effects on each other’s growth and death rates both within their own county and on their nonlocal counterparts. They concluded that their findings “are consistent with a hypothesis of community-level competition: Networks of mutual and commercial companies, united as interdependent communities, may have competed with other such networks” (pp. 411-412). Furthermore, these populations had symbiotic effects on each other: The probability of a mutual company’s offering long-distance service was positively related to the density of commercial companies, and the density of mutuals had a similar effect on commercials. Barnett (1990) found that the density of organizations using a complementary technology significantly decreased the hazard of death experienced by early telephone companies in Pennsylvania, indicating mutually beneficial relations among these organizations. Together, these findings support a view of communities of organizational populations variously linked by competitive and mutualistic relationships; these interdependencies must be understood to have a complete view of any single population’s natural history.

Critique. Organizational ecology differs substantially from previous adaptive approaches to organization-environment relations and as a result has faced both theoretical stumbling blocks and criticisms not encountered by other work. Defining organizational form and population in a theoretically satisfying way, specifying the role and nature of environmental selection, and delineating forms of change that shape the structure of organizational populations have all been contentious topics. We will discuss each of these issues and some potential resolutions in turn.

Defining Form and Population. Basic to defining a population is the task of determining what an organizational form is: Populations are defined as bounded sets of organizations sharing a common form. Early work on ecology defined organizational form as “a blueprint for organizational action” that can be inferred from an organization’s formal structure, patterns of activity, or normative order (Hannan & Freeman, 1977, p. 935). Yet almost without exception, empirical research within the ecological school has defined populations by their purposes or outputs, at best making rough distinctions between specialists and generalists. Semiconductor manufacturers, newspapers, labor unions, and telephone companies
are all presumed to share a common form among themselves, yet little if any inquiry is made into their formal or informal structures.

Various alternatives to defining form by internal characteristics of organizations have been suggested. Astley (1985) argues for a community ecology approach that conceptualizes "population forms in terms of their functional roles vis à vis other populations within technologically interdependent communities" (p. 225). Hannan and Freeman (1986) see two complementary techniques for discerning form: (a) by applying network models to data on resource flows among organizations, one can identify sets of organizations that share structurally equivalent positions and external dependencies and infer a common form, and (b) by locating boundaries that separate populations and determining the processes that sustain or erode them, one can begin to identify niche structures and the forms that they define. These approaches are not in conflict: Organizations that share a functional role within a community and are bounded within a common system are likely to have similar patterns of dependencies. But as critics have been quick to point out, trying to identify populations in terms of the niches they occupy involves a certain amount of circular reasoning (Young, 1988).

The Role and Nature of Environmental Selection. Selection pressures pervade the theoretical structure of organizational ecology. They are usually assumed to operate on members of populations with an even hand, selecting out those whose forms are not compatible with their environment. Yet this imagery tends to overemphasize the distinction between organization and environment and to downplay the active strategies by which organizations construct their own environment. Research on interorganizational networks points out how network position can distort the selection pressures faced by organizations. Moreover, ecologists have been rather cavalier about defining the exact nature of a competitive regime: Is it production efficiency, accountability, reliability, legitimacy, or some combination of these factors?26

Even though ecologists are not inattentive to matters of politics, we think their work would be enhanced if they were to develop arguments testing models of selection on political or network grounds. The technologically interdependent communities described by Astley (1985) are themselves embedded in a legal and political framework. State and national governments are influenced by political action committees, chambers of commerce, trade associations, and other forms of organizational collective action; in turn, governments provide the institutional structure in which organizations live and die. With the increasing mobility of business investment, competition takes place at even higher levels. State governments and local business communities vie with each other over the locations of plants and other investments, and national governments engage in a diffuse competition for business investment through their attitudes toward unions, tax laws, social welfare policies, and trade policies (Burawoy, 1985; Lindblom, 1977). Thus, competition and its effects occur simultaneously at the level of the organization and the community of organizations as well as at the state and national level, suggesting that a much broader view of selection processes is appropriate.

Forms of Change. Theorists of biological evolution have debated the nature of evolutionary change in populations for some time, and this debate has been reproduced in organizational ecology. At issue is whether the character of populations is determined gradually through selective retention or whether change occurs primarily through abrupt bursts in which old populations die and new ones are born, followed by periods of stasis (the doctrine of punctuated equilibrium). While population ecology theory is not inconsistent with a
punctuationalist view, in practice most ecological research is implicitly gradualist, taking populations as given and looking at the forces that shape them over time (Astley, 1985; see also Isaac & Griffin, 1989, for a critique of the ahistorical approach to labor history taken by ecologists). Research on the effects of technological innovation on industries (e.g., Tushman & Anderson, 1986) suggest that there is considerable merit in attempting to model punctuated equilibrium more explicitly.

A more radical problem for the ecological program is posed by current events that highlight the distinctive nature of capitalist firms within the world of organizations. Ecological researchers often proceed as if there were sufficient commonality among organizations to model the same organizational processes (such as density dependence in birth and death rates) across very different populations. It is assumed that American labor unions, semiconductor firms, early telephone companies, newspapers, and breweries in various areas of the world, and voluntary social service organizations in Toronto all share timeless causal regularities by dint of the fact that we can refer to them as organizations. Thus, one can explain changes in the mix of organizational forms over time through the births and deaths of inert forms.

We suggest that the facts of recent American corporate history pose a serious challenge to the viability of this project. Of the Fortune 500 firms in 1980, over one-quarter were subjected to a takeover or buyout attempt during the subsequent decade, most hostile, and most successful (Davis & Stout, 1990). A common post-buyout practice was to sell off units (deconglomeration) in order to pay off the financing of the buyout, leaving the organization that remained a radically restructured one. Those organizations that avoided being taken over typically underwent massive defensive restructurings, as did a large proportion of the firms that did not experience a takeover attempt (Hirsch, 1987). The financial conception of the corporation that now dominates American business (Flistein, 1990) allowed parts of organizations to be traded like baseball cards. Conversely, the firms in this population that failed through bankruptcy during the 1980s can be counted on one hand. Thus, what was perhaps the most radical remix of consequential organizations in American history occurred through processes that (a) are unique to capitalist firms that are publicly owned and (b) bear no resemblance to the births and deaths of biotic forms. Ecologists might respond that there is nothing special about the Fortune 500 (after all, there are more than 500 shoeshine stand organizations in Manhattan alone) or that we should consider such takeovers and restructurings to be the deaths of old (presumably inert) organizations and the births of new ones (cf. Hannan & Freeman, 1977); readers can decide for themselves on the merits of this position.

Future Directions. Ecological research has in many ways grown beyond the initial theoretical statements of population ecology. The jettisoning of a strict biological metaphor, which seems to have had a persistent appeal for students of organizations since the advent of open systems theories, has brought organizational ecology more in line with many of the classic concerns of organizational studies: technology and interdependence (Barnett & Carroll, 1987; Tushman & Anderson, 1986), relations of organizations with their institutional environments (Carroll & Ho, 1986), resource mobilization and social movements (McCarthy et al., 1988), and organizational demography and recruitment processes (McPherson, 1983). These developments suggest that organizational ecology may be well served by consolidating its strengths and integrating the concerns and intellectual technology of other approaches, in particular the institutional and network perspectives. But ecology's distinctiveness—its methodological rigor and dynamic approach to modeling and the explicit attention to a population perspective (Wholey
& Brittain, 1986)—must be highlighted. These tools have allowed the discovery of unexpected regularities: Curvilinear density dependence in founding and death rates and the findings of population and community dynamics are features of organizational life that would have remained unknown without taking a dynamic, population-level approach.

Ecologists have made headway on some of the thornier problems (e.g., adequate specification of organizational forms, the appropriate units and levels of analysis, the nature of organizational and population boundaries) that their perspective raises. We suspect that ecological research will remain controversial for some time precisely because its main intellectual goal—to develop theory at the population level—recognizes that initial assumptions must be made that (a) greatly simplify such key processes as competition and legitimation and (b) abstract much of the organizational detail and complexity that characterize what goes on in organizations. Ecologists clearly think this tradeoff is worth making; others may be less persuaded.

Institutional Perspectives

The label institutionalist has a long but ambiguous pedigree. Such diverse "masters" of social science as the sociologists Durkheim and Parsons and the economists Commons and Veblen were comfortable under the institutionalist flag. In contemporary social science there are, perhaps, as many new institutionalisms as there are disciplines.79 In this section we attend to the family of writings, often tagged as the new institutionalism, that have had the most impact on organization theory. Yet even within organizational studies, institutionally oriented research exhibits little of the coherence or the formalism associated with, say, population ecology or transaction cost economics. In this respect, the new institutionalism has an affinity with network research: "It is often easier to gain agreement about what it is not than about what it is" (DiMaggio & Powell, 1991).

Perhaps the most novel tenet of the institutional approach is the insistence that organizational environments must be viewed in cultural as well as technical and economic terms (Scott, 1983). Organizations and their members are embedded in cultural systems composed of rules, norms, and taken-for-granted assumptions that define the way their worlds operate. Two of the most powerful sources of cultural blueprints are the modern professions and the modern state. Consequently, the institutional approach directs attention both toward the macro level of state structures, legal systems, and the sovereign modern professions and to the micro level of everyday interactions. It is in large part at this level of individual interaction and cognition that institutional practices and beliefs are translated into both constraints on action and "tool kits" that can be used to construct and legitimate new courses of action.

Although ecological and institutional approaches differ markedly in the relative weight they assign to human volition and organizational adaptability, there is growing recognition that these two perspectives share a number of insights. Both focus on the collective organization of the environment, insisting that the environment of organizations is made up of other organizations and that the demographic and structural properties of the environment shape organizational behavior. But ecologists attend primarily to demographic processes—organizational foundings, transformations, and deaths. Institutionalists contend that key features at the environmental level influence not only demographic processes but the internal structure of organizations within a given field as well.

A number of new insights are suggested by the institutional approach. Environments are viewed as comprised not only of technical requirements, bundles of resources, and patterns of communication, but also of cultural
elements—symbols of legitimacy, belief systems, and professional claims (Scott, 1983). DiMaggio and Powell (1983) contend that "organizations compete not just for resources and customers, but for political power and institutional legitimacy, for social as well as economic fitness" (p. 150). Demonstrating social fitness often entails conforming to rational myths (Meyer & Rowan, 1977)—beliefs that specify what activities need to be carried out and what types of actors must be used to achieve specific social purposes. Because of their specificity and goal-directedness, such beliefs are rational. At the same time, however, these beliefs are like myths in the sense that their efficiency is presumed on the basis of their wide adoption, or their championing by groups who have been granted the right to determine such matters. Meyer and Rowan (1977) point out that there are multiple and diverse sources of rational myths: public opinion, educational systems, laws, courts, professions, ideologies, regulatory structures, certification and accreditation bodies, and governmental requirements.

Moreover, they tell us that in modern nation states, the forms and sources of widely held beliefs have themselves become more rationalized: Folkways and traditions and customs give way to laws, rules, and regulations; forms of traditional authority are replaced by the nation state, the professions, and systems of law. Thus, through its focus on processes that confer legitimacy on particular activities, the institutional approach directs attention away from material factors such as the location of physical resources or customers and toward the state and the professions, which shape organizational life both directly by imposing constraints and requirements and indirectly by creating and promulgating new rational myths.

Institutional theory combines a rejection of the assumptions of rational actor models popular in economics with an interest in institutions as independent variables, a turn toward cognitive and cultural explanations, and a concern with properties of supraindividual units of analysis that cannot be reduced to simple aggregations of the characteristics and actions of individuals. The constant and repetitive quality of much of organizational life results not from the calculated actions of self-interested individuals but from the fact that practices come to be taken for granted as "the way we do things." The model of behavior is one in which "actors associate certain actions with certain situations by rules of appropriateness" (March & Olsen, 1984, p. 741); these patterned responses are absorbed through socialization, education, on-the-job learning, or through acquiescence to convention. Individuals in organizations face choices all the time, but in making decisions they seek guidance from the experiences of others in comparable situations and by reference to standards of obligation."

The Language of Institutional Theory. As with each of the perspectives we have reviewed, practitioners tend to develop their own distinctive language in the form of a battery of widely used concepts and arguments. Three topics are highlighted in this approach: institutionalization as both a cognitive outcome and an exogenous process, organizational fields as key units of analysis, and processes of institutionalization that promote conformity within fields.

Institutionalization. What does it mean to say that something has become institutionalized? Curiously, institutional theorists seem to be of two minds on this key issue (see Jepperson, 1991, for a commendable effort to clarify the conceptual variety that characterizes this approach). A cognitively oriented line of argument stresses that practices that are institutionalized are the product of ongoing repetitive interactions that gradually acquire a rulelike, social fact quality. That is, a pattern of activity is institutionalized when it comes to be taken for granted and therefore persists without serious questioning or efforts to make it continue
(cf. Abelson, 1981, on the concepts of script and schema in psychology).

A second line of argument locates institutionalization within formal aspects of organizations, rather than as by-products of the interactions of individuals. This work associates institutional processes with the actions of the state, the professions, and other central organizations to establish a collective normative order. Research in this vein focuses on the causes and consequences of conformity to demands by such central organizations and on the ways in which the environment "interpenetrates the organization" (Meyer & Rowan, 1977).

In practice, institutionalists often invoke both meanings, thus, it is not easy to assign scholars to either label. At this point, the result is some conceptual ambiguity and methodological confusion. The ultimate goal is fairly clear, however: to understand how organizational practices and forms are developed and legitimated, come to be taken for granted, and eventually fall into disfavor. Clearly, this cycle has both micro (cognitive) and macro features to it.

**Organizational Fields.** The appropriate unit of analysis in the study of institutionalization is the organizational field or societal sector. The basic assumption is that organizations exist in socially constructed fields, composed of similar organizations that are responsible for a definable area of institutional life. An organizational field includes key suppliers, resource and product consumers, regulatory agencies, and professional associations, as well as other organizations that produce a similar service or product. The virtue of this approach is that it focuses attention not simply on competing units or on networks of organizations that directly interact with each other, but on the totality of relevant actors.

The structure of an organizational field is not easily determined a priori but must be defined on the basis of empirical investigation. Two examples might be useful here. The organizational field in the study of health-care provision might include hospitals, HMOs, insurance companies, federal and state regulatory bodies, medical and nursing schools and professional associations, and pharmaceutical and hospital-supply companies. The key boundary issue is the extent to which these varied groups take each other's behavior into account in formulating their actions. The organizational field in research on the high arts could span museums, symphony orchestras, theaters, public agencies concerned with the arts, private collectors and commercial galleries, foundations and granting agencies, as well as university departments and schools and trade and professional associations. The critical issue is the degree to which members of fields are structured into a common community.

DiMaggio and Powell (1983) argue that the process by which an organizational field comes to be structurally defined consists of four parts: (a) an increase in the extent of interaction among organizations within a field, (b) the emergence of sharply defined interorganizational structures of domination and patterns of coalition, (c) an increase in the information load with which organizations in a field must contend, and (d) the development of a mutual awareness among participants in a set of organizations that they are involved in a common enterprise.

**Processes of Institutionalization.** Assuming that organizations view themselves as members of a field or sector, what factors shape their orientation toward one another? In particular, how do organizational practices become institutionalized within a field? There are several mechanisms conducive to organizational isomorphism, for example, structural similarities among organizations within a field. Some of these processes encourage homogenization within a field directly by leading to structural and behavioral changes.
in organizations themselves. Others work indirectly by shaping the assumptions and experiences of the individuals who staff organizations. DiMaggio and Powell (1983) posit three general types of institutional pressures: (a) coercive forces that stem from political influence and problems of legitimacy, (b) mimetic changes that are responses to uncertainty, and (c) normative influences resulting from professionalization. These three mechanisms are, of course, likely to intermingle in specific empirical settings, but they tend to derive from different conditions and may lead to different outcomes. Indeed, institutional pressures may be cross-cutting and lead to conflict (Powell, 1988; Scott, 1987b).

Coercive influence results from both formal and informal pressures exerted on organizations by other organizations upon which they are dependent, as well as by strongly held cultural expectations in the society at large. In some circumstances, organizational change is a direct response to government mandate: Manufacturers adopt new pollution control technologies to conform to environmental regulations, nonprofits maintain accounts and hire accountants to meet the requirements of the tax laws, restaurants maintain minimum health standards, and organizations hire affirmative action officers to fend off allegations of discrimination.

Uncertainty is a powerful force that encourages mimetic or imitative behavior among the members of an organizational field. When organizational technologies are poorly understood, that is, when managers are unclear about the relationship between means and ends, when there is ambiguity regarding goals, or when the environment is highly uncertain, organizations often model themselves after other organizations. The modeled organization may be unaware of the modeling or may have no desire to be copied; it merely serves as a convenient source of organizational practices that the borrowing organization may use. Models may be diffused unintentionally, indirectly through employee transfer or turnover, or explicitly by organizations such as consulting firms or industry trade associations. In this view, the ubiquity of certain kinds of modern management practices is credited more to the universality of mimetic processes than to any concrete evidence that the adopted models enhance efficiency.

A third source of organizational change is normative and stems, to a considerable degree, from the culture of professionalism. Two aspects of professionalism are particularly relevant. One of these is the growth of professional communities based on knowledge produced by university specialists and legitimated through academic credentials; the second is the growth and elaboration of formal and informal professional networks that span organizations and across which innovations may diffuse rapidly. Universities and professional training institutions are important centers for the development of organizational norms among professionalized managers and staff. Professional and trade associations are another vehicle for the development and spread of normative rules about organizational and professional behavior.

Key Areas of Research. Much of the early empirical work focused on the diffusion of governmental policies (Tolbert & Zucker, 1983) and on public and nonprofit organizations in such areas as education, health care, mental health, and the arts (e.g., see the studies in Meyer & Scott, 1983, or the bulk of the illustrative examples in DiMaggio & Powell, 1983). Education is the sector that has undoubtedly received the greatest scrutiny by institutionalists (Kamens, 1977; Meyer, 1977, 1983a, 1983b, 1988; Meyer & Rowan, 1978; Meyer, Scott, & Strang, 1987; Rowan, 1982; Scott & Meyer, 1988; Tolbert, 1985). We briefly summarize this extensive literature, drawing freely from the aforementioned sources.

The 20th century has seen a wholesale expansion of the roles of the states and the federal government in American education (Meyer, Scott, & Strang, 1987). State control has
expanded in most domains of schooling (e.g., accreditation, curriculum guidelines, personnel certification, etc.). This process of state expansion is uneven across the states and is subject to conflict and debate, but the general trend is clear. Moreover, since the 1960s, the federal role has grown as well. But instead of a national educational policy, federal programs take the form of categorical or special purpose programs. This heightened complexity in the larger environment has several consequences for schools. The more highly structured policymaking becomes, the more schools focus on conforming to the official categories provided by the larger environment. But this conformity may involve only an organization’s formal structure (i.e., its organizational chart of reporting relationships, its rules and procedures, etc.), which is readily visible to the outside world. In order to be perceived as legitimate by the wider environment, educational organizations adapt their formal structure to conform to institutional norms.

In numerous studies, John Meyer, W. Richard Scott, and their colleagues found wide consensus among superintendents, principals, and teachers on formal policies—grades, curricular materials, and so forth—but very little agreement, even within schools, about teaching methodologies or substantive measures of educational effectiveness. In other words, educational organizations evince loose coupling between their formal structures and their everyday activities. For example, understanding about using grades on a scale from A to F and an annual progression from K through 12 is largely taken for granted, yet understanding of what is effective in the classroom is not. The advantages of adherence to fieldwide norms and requirements on readily visible attributes are many: increased support, legitimacy, stability, internal and external commitment, eligibility for funding, enhanced attractiveness to personnel, and protection against charges of malfeasance or negligence. Moreover, by conforming to institutionalized expectations, schools avoid close scrutiny of or control over their instructional activities.

Much of the initial theorizing about institutional environments was grounded in these studies of educational organizations (see Meyer & Rowan, 1977; Meyer, Scott, & Deal, 1981). In this research, the distinction between technical and institutional environments looms large. The former involves organizations whose success is dependent on solving technical problems, that is, achieving high standards of production, while the latter consists of organizations, such as schools, whose survival requires conformity to the normative demands of the larger environment. In technical environments, organizations are evaluated by their outputs. These firms closely monitor production and buffer their technical cores from environmental influences according to Thompson (1967). Institutional environments are composed of organizations judged more by the appropriateness of their form than by their outputs. The distinction between technical and institutional environments suggested that organizations in technical environments were rewarded for efficient production, while organizations in institutional environments were compensated for conformity and legitimacy.

Drawing on research on the arts and culture industries (DiMaggio, 1983; Powell & Friedkin, 1986), DiMaggio and Powell (1983) made a similar contrast between competitive and institutional isomorphism (see also Fennell, 1980). The concept of isomorphism is borrowed from Hawley (1968); it refers to a constraining process that forces one unit in a population to resemble other units that face the same set of circumstances. In the ecological approach, competitive isomorphism is driven by selection processes that weed out unfit organizations. The institutional approach initially emphasized accommodation, rather than competition, with environmental forces that promote sectorwide conformity. Such a view seemed apt for cultural activities, especially those in the nonprofit arena. The production of
culture relies on a weak technical base (i.e., it is hard to draw up a blueprint for a successful artistic project, just as it is difficult to articulate a widely generalizable program for rehabilitating criminals). Arts organizations typically have ambiguous or conflicting goals (e.g., publish great books and make money); and market tests for fitness are sometimes softened by public and philanthropic support. In this field, as in education, formal conformity to fieldwide expectations often insure continued survival.

But dichotomies between the technical and institutional or competitive and institutional environments proved primitive. Both approaches ceded too much terrain to competitive market forces (Powell, 1985, 1991) and failed to recognize that some organizations confront multiple, conflicting demands (Scott, 1987b). In particular, Scott has highlighted the health care sector as one in which organizations must contend with vigorous technical and competitive pressures as well as stringent regulatory and institutional demands.

Health care sectors—both medical and mental health—have proven to be another fertile ground for institutional analysis (see Alexander & Scott, 1984; Fennell, 1980; Meyer, 1986; Scott, 1985; Scott & Black, 1986; Zucker, 1987). Like education, health care has experienced an expansion in state controls and regulatory pressures and the growth of multiple layers of oversight and coordination. But the institutional environment of health care is characterized by conflict among competing jurisdictions. The state, Scott (1987b) suggests, prefers more centralized forms of control and coordination, while the medical professions opt for more decentralized systems of procedures and rules; at the same time, health care organizations compete vigorously for patients and race to adopt the latest technologies.

The research on health care, along with a burgeoning strand of research on institutional processes in the professions and in proprietary sectors, made it clear that technical and institutional pressures are not mutually exclusive; rather, they are best regarded as dimensions along which organizational environments vary. Sectors can face environments characterized by high technical pressures, high institutional pressures, neither, or both. Some sectors, such as banking or transportation, face both strong competitive and technical demands as well as pressures from various regulatory bodies and consumer groups to conform to procedural requirements. As a result, we find that the administrative structures of organizations in these fields are larger and more complex than those of organizations facing less complex environments (Powell, 1988; Scott & Meyer, 1991). In general, organizations of this type carry out tasks that combine complex technical requirements with a strong "public good" component. Schools, arts organizations, churches, and many professional service firms face strong institutional environments but relatively weak technical demands. In contrast, firms in manufacturing may experience some institutional pressures with respect to employment policies, safety standards, or pollution controls, but their main concern is production efficiency. Finally, one can imagine classes of organizations with both weak technical bases and fairly minimal institutional pressures. Exercise clubs or video stores would fit in this category.

In tandem with a broader view of the comingling of technical and institutional demands in some fields, there has been much more research on for-profit firms that falls under the general rubric of institutional theory. We should note, however, that one of the striking features of the new institutionalism is its broad diffusion into many areas of organization theory and general sociology. This popularity comes at a cost, however, because much of the work under an institutional banner employs divergent concepts, different and at times even contradictory hypotheses, and dissimilar boundary conditions. Indeed, it is not always clear what an institutional
account is opposed to: Is institutionalism an alternative to rational actor accounts or arguments that depict the collective as primarily an additive outcome of micro-level interactions? an alternative to causal models that highlight single-level explanations? an approach that is complementary with other extant perspectives but adds a needed element of context? or a residual category that purportedly soaks up unexplained variance? In practice, so-called institutional arguments have been all of these; hence, caution flags are warranted.

We selectively review some of this contemporary work, with an eye toward highlighting research on the role of the professions, patterns of diffusion or adoption of particular policies or legal agendas, and structural change in organizations. We think that despite some lack of clarity about what an institutional argument amounts to, these lines of research have proved valuable.

One of the key tenets of the institutional approach is that the professions play a key role in shaping the institutional environment. Consequently, researchers have focused their attention on both the organization of the professions, notably law and accounting, as well as on patterns of change in professional standards. There are studies of the training and promotion procedures in law firms (Tolbert, 1988; Tolbert & Stern, 1989), the uses of management-by-objectives policies in accounting firms (Covaleski, Dirschmid, & Heian, 1990), and the organization of the legal departments of multinational corporations (Miyazawa, 1986). These internal organizational studies are complemented by research on professional practice, which analyzes the championing and adoption of policies promulgated by professional bodies. Covaleski and Dirschmid (1988) examined how the use of specific budgetary rules took on symbolic value and shaped the political relations between the University of Wisconsin and state government. Their story is both fascinating and suggestive: The university had long been an advocate of more rational, professional government, only they did not expect this advocacy to reshape their own budgetary process. This is one of the few studies that examines how organizations strategically respond to institutional pressures by trying to pacify and bargain with institutional stakeholders (see also Edelman, 1990 b).

Mezias (1990) has studied a much broader diffusion process, the adoption of the investment tax credit among Fortune 200 corporations. His study has a 22-year time frame and makes a persuasive case that this change in accounting methods is best explained by an examination of the actions of the accounting profession, federal regulatory bodies, and the early adopting firms, not by the self-interested motives or characteristics of the adopting corporations. This research does two things notably well: It pushes our understanding of how the institutional environment is collectively organized (no simple dominance account would explain the interactions of the Big 8 accounting firms, the accounting standards boards, and the multiple regulatory authorities), and it incorporates the actions of large firms into the macro environment.

The diffusion of standards, laws, and even structural changes, such as the multidivisional form and matrix management, are all being studied by scholars using an institutional lens. We attend to key aspects of this line of research below, but first we note that there is considerable ambiguity about two issues in this work: the boundaries of the institutionalization process and the motives of the actors involved. Despite the use of the field or sector concept in much of the theoretical writing, the empirical studies seem to focus on important processes of institutionalization that cross fields. Sometimes the fields are bounded by region or industry, but in other studies the scope is as broad as the 100 or 200 largest U.S. corporations, or even the formation of personnel departments throughout U.S. industry. This inconsistent operationalization robs institutional explanations of some of their explanatory
power. Similarly, scholars vary widely in terms of their accounts of motives. In one of the earliest empirical studies, Tolbert and Zucker (1983) examined the introduction of civil service procedures during the period 1880 to 1935. In the early years, adoption of civil service reforms was associated with demographic characteristics of the cities (i.e., number of immigrants, city size, and working class population). In these cases, cities appeared to be adopting governmental reforms that would improve the authority of city officials and exclude immigrants and working class people from power. These are rational or self-interested motives for those in power. But after 1915, these characteristics no longer predict adoption. Tolbert and Zucker suggest that later adopters were trying to appear up to date: Civil service reforms had come to signal modern, rational city government, and thus adoption was widespread. Such an account is not fully persuasive; conditions change, power holders come and go, perhaps later incumbents were responding rationally to a new set of circumstances. Moreover, by leaving institutionalization as a residual category (i.e., failing to find that measured characteristics predict adoption is taken as evidence that adoption had become institutionalized), this research may be argued to stack the deck in favor of finding evidence for institutionalization.

In other research, motives are exempted entirely from analysis. Edelman (1990a) suggests that “it would be very difficult to distinguish empirically between such disparate motives for the formalization of due process rights as efficiency, control, and legitimacy” (p. 1403). We are bothered by both the loose boundary issue and by the limited attention to motives because they render the scope of institutional theory problematic. Nevertheless, these are initial studies of a new line of theorizing and it behooves us to examine the results.

Galaskiewicz (1985b, 1985c; Galaskiewicz & Wasserman, 1989) has done fascinating work on the development of corporate philanthropy in Minneapolis and St. Paul. The Twin Cities are noted for their vigorous program of corporate support for the nonprofit sector. He has shown how a general consensus on the role of philanthropy has become institutionalized among large corporations housed there and how this philanthropic mind-set is championed by corporate public affairs officers and rewarded by the social elite of the Twin Cities. This research stresses the role of inter-organizational networks: When actors are faced with uncertainty, they turn to others whom they know, trust, and admire for guidance. The boundary or field in this work is defined geographically by membership in the Twin Cities corporate philanthropy community, either as a donor or recipient.

In other studies, however, researchers have adopted much broader and more diffuse boundaries and have not attended to the actual process of information transmission. Baron, Dobbins, and Devereaux Jennings (1986) underscore the role of the federal government and the newly emerging field of personnel administration in their analysis of the proliferation of so-called modern personnel practices throughout the core of the American economy during the middle of this century. Their sample covers nearly every major manufacturing industry. Similarly, Dobbin, Edelman, Meyer, Scott, and Swidler (1988) and Edelman (1990a, 1990b) employ broad samples in their studies of the adoption of due process procedures in corporations. Edelman’s (1990a, 1990b) work is a nice combination of event history methods, commonly used by ecologists, and clear development of a theory of the legal environment. In her 1990a study, she looks at the changes in personnel practices of 52 organizations following the legislative reforms of the civil rights era. She maps these changes with measures of the intermediary role of personnel professionals as well as the corporation’s proximity to the public sphere, for example, significant regulatory or contractual linkages to government. The expansion of due process rights was rapid, going
significantly beyond direct legislative mandate. Of course, the governance of employee rights is ripe for symbolic manipulation. She does not investigate whether formal rules are closely tied to actual practices. But in a subsequent paper (Edelman, 1990b), with a much larger sample of business, educational, and government organizations, she examines equal employment opportunity laws with an eye toward how organizations mediate the adoption of these reforms. EEO is especially vulnerable to mediation, she notes, because the meaning of compliance is vague, the legal strictures focus more on procedures than outcomes, and enforcement mechanisms are weak. At the individual level, compliance with EEO is a function of internal organizational politics, balanced with considerations of industry norms and professional standards. At the collective level, organizational responses to the law shape societal and legal expectations about what constitutes good faith compliance. This is an important line of work because it shows not just the impact of the legal environment and the diffusion of new legal norms, but how the response of organizations to these new standards shapes the broader definition of what actions are legitimate and acceptable.

Another noteworthy line of inquiry addresses fundamental structural changes in organizations. Here the focus is on decisions to restructure an organization—clearly not a task undertaken for purely symbolic reasons. Thus, to the extent institutional theory is useful in explaining these reorganizations, this approach clearly speaks to core issues involving private sector firms. Still, we note again that this research also employs dissimilar notions about the boundaries of a field. Fligstein (1985) focuses on the largest 100 U.S. corporations over a 60-year period from 1919 to 1979. The critical issue is the decision to adopt a multidivisional structure. He finds that, contrary to what one would predict based on transaction cost economics, this choice is only partly driven by a firm’s size or its competitive position. A more complete understanding is gained by examining a firm’s responses to changes in federal antitrust law, as well as the tipping points at which a critical number of other firms in the industry have shifted from a functional to a multidivisional structure. He does not argue that economic calculations do not matter, but they are tempered by a corporation’s assessment of the likelihood of government regulatory action and the collective choices made by other firms that are regarded as role models. A similar line of analysis is employed in very different arenas by Burns and Wholey (1990) in their study of the adoption of matrix management by hospitals, and by Amburgy and Lippert (1989) in their research on the diffusion of management buyouts.

These diverse areas of research do not as yet have a clean cumulative payoff. Institutional theory and research has proceeded in fits and starts, and in some cases the research bears only a vague family resemblance to the purported theory. Nevertheless, there are several critical points that have been well developed theoretically and have received substantial empirical support. In particular, we know a good deal about the forms and patterns of institutionalization. Rather than simply saying that the environment matters, scholars in this area can now point to both specific collective properties of the environment (e.g., number of levels of controls and types of control systems) and to key agents (the professions and federal and state regulatory bodies) who influence the process by which organizational forms and policies become institutionalized. Moreover, we have an enhanced understanding of the relationship between environmental complexity and internal organizational structure. When environments contain multiple centers of authority and legitimacy (what Scott & Meyer, 1983, term fragmented authority structures) we find greater diversity in organizational forms, greater differentiation among individual organizations, and more levels of
administration within organizations (see the summary of this research in Meyer, Scott, & Strang, 1987; also see Powell, 1988; Scott & Meyer, 1991). In contrast, when environments are more homogeneous we find less elaborate internal structures. For example, Tolbert (1988) found that law firms that selected new associates from the same schools as those of their older partners found little need for intensive socialization programs or detailed performance reviews.

Critique. The thrust of institutionally oriented work in organizational theory has been to locate the sources of various organizational practices and structural arrangements within a broader context. Rather than viewing organizational actions as efforts to manage dependencies or reduce transaction costs, institutionalists see organizational actions as legitimacy-enhancing responses to the structure of relationships within organizational environments. In this sense, institutional theory is both more micro and more macro than other approaches. Implicit in the institutional approach is an essentially cognitive or ethnomethodological view of human action as shaped by conventions, built up from the ground level by participants in the course of interactions. To the point that much behavior takes on a taken-for-granted quality (DiMaggio & Powell, 1991; Zucker, 1977). At the same time, the institutional perspective employs a more structural focus that emphasizes the incentives created by larger vertical authority structures external to the organization and the role of professional networks in aiding in the diffusion of organizational beliefs and practices.

In some respects this approach is a novel one. It orient organizational research in cultural and normative directions that have been largely ignored. On the other hand, institutional views share a great deal with what we might term general sociology. Not surprisingly, then, one of the key shortcomings is an oversocialized and rather passive view of human agency. Where does action come from and who benefits from organizational change? Thus far, institutionalists have been rather silent on these issues, but the latest work begins to suggest the outline of a response.

Power and interests have been slighted topics in institutional analysis (DiMaggio, 1988; Perrow, 1986; Powell, 1985). Little attention has been directed toward explaining how organizational incumbents maintain their dominant positions or respond to threats during periods of crisis or instability. Nor has much work been done on how skilled institution-builders put multiple institutional logics to use to fashion strategic change. Efforts to incorporate power into institutional arguments should start with two simple observations: (a) Actors in key institutions realize considerable gains from the maintenance of those institutions, and (b) when organizational fields are unstable and established practices ill-formed, successful collective action often depends on defining and elaborating widely accepted rules of the game. Consequently, the acquisition and maintenance of power within organizational fields requires that dominant organizations continually enact strategies of control, most notably through either the socialization of newcomers into a shared world view or via the support of the state and its judicial arm.

Fligstein (1990) makes this point nicely in arguing that certain corporate strategies were favored by CEOs with marketing and finance backgrounds because they fit their interests and competencies. Successful executives developed conceptions of control that came to dominate their industries and defined appropriate standards of behavior. DiMaggio (1991) shows how early 20th century museum professionals sought radical changes in museum missions and policies, changes that also enhanced their own positions relative to those of their trustees.

Institutional theorists need to move beyond earlier statements that stressed how rules and routines created order and minimized uncertainty and examine how institutional
arrangements are also replete with conflict and contradiction (Scott, 1987a). Thus, several fundamental questions remain to be addressed: How persistent are institutions—how mutable are institutionalized practices? When do different institutional logics challenge one another? What is the role of organizational elites in maintaining existing institutions? Under what conditions are challengers and entrepreneurs able to refashion existing rules, buffer themselves from institutional demands, or create new institutional orders? And, finally, what are the tensions between arguments that emphasize the "stickiness" of institutions and approaches that assume an optimization logic, depicting institutions as the results of intentional actions or adaptive solutions to conflicting interests?

We began this section with mention of the growing rapprochement between institutionalism and ecology. Institutionalists are now much more willing to acknowledge the importance of competition and organizational selection than they once were (see Powell, 1991). Ecologists, for their part, now emphasize the importance of institutional factors and explicitly criticize Panglossian models of organizational evolution (Hannan & Freeman, 1989). Singh, Tucker, and Meinhard (1991) provide an apt example of this convergence: Using population models, they demonstrate the effects of institutional change on population dynamics and the salutary effects of institutional legitimacy on the survival rates of voluntary social service agencies in Toronto. They suggest that competition for social fitness has a decided payoff. Rather than deny the importance of competition, institutional theorists now emphasize how competition varies across historical periods and societies and stress the role of institutions in constituting these different regimes, while ecologists use institutional insights to understand how selection criteria vary in different organizational populations (e.g., Barnett & Carroll, 1990).

**Conclusion**

We began this chapter with a warning that we would not present an integrated body of thought on the topic of organization-environment relations. We chose instead to present various contemporary lines of research, and we have tried to do so candidly and fairly. We stressed that it is critical to understand the different orientations of these approaches to the issues of the relevant unit of analysis and questions of the motives that undergird organizational actions. We think many criticisms of and ostensible points of disagreement among these perspectives are often based on a misunderstanding of what these theories are actually trying to explain.

We have presented a "warts and all" view of the various theoretical camps, trying to be as critical of lines of work that we are associated with as those with which we have personal points of disagreement. But our goal has not been to provoke controversy or appear as curmudgeons. Instead we want to highlight that these are vital avenues of research with questions unanswered and relevant empirical work waiting to be done. We hope that we might encourage a few readers to join in this task.

We also have noted points of convergence among these perspectives and possible topics of affinity. We do not believe that there is a correct, all-encompassing theory of organization-environment relations. In the early stages of theory development, a school of thought may find it expedient (and professionally rewarding) to view its approach as a total causal explanation of organizational phenomena. But as tensions mature, it obviously makes more sense to reorient them in a way that allows for competing theories to contribute to our understanding of organizational behavior. Moreover, we have stressed potential points of synthesis: opportunities to employ network theory to define relevant populations for ecological analysis, possible points of contact between
transaction costs economics and networks in analyzing the durability of relationships, and potential collaboration between ecological and institutional theories on issues of how the institutional environment shapes selection processes and how ecological dynamics might result in institutional change. We close not with the perennial plea for more research, but with admiration for the accomplishments of the past two decades and anticipation of the work that is to come.

We thank Paul DiMaggio for his comments on our initial outline of this chapter, Peter Brantley, James Ranger-Moore, and Charles Perrow for comments on an earlier draft, and Marv Dunnette for his patience.

Notes

1 There are a number of recent surveys of the state of the art in macro-organizational behavior (e.g., Aldrich & Marsden, 1988; Pfeffer, 1985) as well as several useful textbooks (Hall, 1987; Perrow, 1986; Scott, 1987a), and we recommend them to interested readers.

2 Scott (1987a, chap. 6) provides a much more detailed guide to these various conceptions of the environment.

3 Many strands of more recent research draw freely on resource dependence arguments. Much of the work on directorate interlocks done from a network perspective incorporates insights from resource dependence theory, and the notion of coercive isomorphism in institutional theory (DiMaggio & Powell, 1983) builds directly on Pfeffer's work.

4 For efforts to operationalize transaction costs empirically, see Joskow (1985), Masten (1984), and Stuckey (1983).

5 See Rosen (1988) for a thoughtful review of research on internal labor markets from a transaction costs perspective.

6 Walker and Weber's (1984) results are, in fact, potentially damaging to the transaction costs explanation for vertical integration: They found that volume uncertainty (projected level of fluctuations in the buyer's future demand for a component) significantly decreased the likelihood of buying a component rather than making it, even though it should be relatively easy to write contingent claims contracts for such situations. Technological uncertainty (frequency of expected changes in specifications and probability of technological improvement in the future), on the other hand, had no effect on make-or-buy decisions. Walker and Weber suggest that this may be due in part to the fact that retooling was paid for by the buyer. But this implies that contracts may be written such that asset specificity (in this case, the costs of retooling) can be rendered unproblematic, obviating the need to bring in an internal supply.

7 See Williamson, Wachter, and Harris (1975) for a full elaboration of this argument and a discussion of the various types of employment contracts that could potentially be constructed.

8 Williamson (1985) agrees with many of these criticisms, admitting that at present "transaction cost economics is crude, it is given to instrumentalist excesses, and it is incomplete" (pp. 390–393). Presumably these are considered to be theoretical problems to be addressed, not fatal shortcomings that doom the theory.

9 In addition to interorganizational relations, network researchers have focused on intraorganizational networks (e.g., Barley, 1990; Burkhart & Brass, 1990; Nelson, 1989) and network forms of organization (e.g., Eccles & Crane, 1988; Miles & Snow, 1986; Powell, 1987). See Lincoln (1982) for a review relating organizational structures to network structures.

10 Burt (1980a) provides an extensive overview of network models, and we rely on his insights in the discussion of them.

11 Richardson (1987) notes that there have been over 100 studies of directorate interlocks, and the number is still growing. Space limitations require that this review will be rather selective, focusing primarily on the most recent work in this area. See Galaskiewicz (1985a) for a review of earlier research on interorganizational networks.

12 Using marginal rather than proportional measures of exchanges yielded broadly similar results with respect to the stability of market boundaries (Burt & Carlton, 1989).

13 But see also Zucker's (1989) stricture about the dangers of premature or surface convergence.
14 The definition of a population is a matter of some contention among ecologists, with one group allowing populations to be defined by the substantive theoretical issue at hand (Hannan & Freeman, 1977) and the other branch attributing more importance to empirically accurate taxonomies of organizational forms and populations (Mckelvey & Aldrich, 1983). See Carroll (1984a) for a comparison of these two perspectives in defining organizational populations.

15 See, for example, the critiques of ecology by Perrow, 1986, and Young, 1988.

16 James Ranger-Moore (personal communication, 1991) points out that this is less true of plant ecology (where trees keep growing until they die) than animal ecology, on which most population ecologists draw.

17 Structural inertia, like most ecological constructs, is a relative concept: An organization suffers high inertia when the rate of environmental change exceeds the rate at which the organization can change its structure to survive. Even relatively flexible organizations may be subject to high mortality in uncertain and rapidly changing environments; for example, electronics manufacturers in Silicon Valley may seem agile relative to the Post Office, yet they may barely keep pace with the dynamic computer industry.

18 “Dominant competence is defined as the combined workplace (technological) and organizational knowledge and skills...that together are most salient in determining the ability of an organization to survive” (McKelvey & Aldrich, 1983, p. 112).

19 Carroll (1984a) argues that Mckelvey and Aldrich (1983) take an overly rigid view of organizational reality by assuming that such a rigorous and empirically accurate classification scheme is possible given the almost limitless diversity and change inherent in the social world. This issue is not yet settled within the ecological school, although research has tended de facto to side with the latter perspective, which defines population by the theoretical problem at hand.

20 Surprisingly, most of the other measured aspects of the environment of this industry, such as the business cycle and the physical and social infrastructure, did not affect the rates of founding.

21 Carroll argues that, given adequate data that cover a relatively long time period, Makeham’s Law is the best-fitting model of organizational mortality and should serve as a baseline for future research. Makeham’s Law describes mathematically a process where rates decline over time to an asymptotic level, at which point rates remain essentially constant. See Tuma and Hannan, 1984, for a description of this and other dynamic models commonly used in ecological research.

22 In fact, the liability of newness hypothesis does not concern newness per se so much as liabilities of inexperience, low legitimacy, and other characteristics for which age serves as a proxy.

23 This possibility has been taken into account in some studies, however. Freeman, Carroll, and Hannan (1983) report that there is indeed a significant liability of smallness in addition to a liability of newness for American labor unions. Barnett and Carroll (1987) also separated these two variables, finding that increased age and size (number of telephone subscribers) were both associated with longer expected life-spans for early Iowa telephone companies.

24 This problem is due to data limitations, not to any lack of interest by ecologists. The best data sets cover very long periods of time, but the amount of organizational detail tends to decrease in proportion to the time span that is covered.

25 Similar arguments have been made for the book publishing (Powell, 1985) and beer (Porter, 1986) industries.

26 In their most recent work, Hannan and Freeman (1989, pp. 33–38) de-emphasize production efficiency and argue that selection pressures are multidimensional. In many circumstances, they suggest, political ties are paramount.

27 DiMaggio and Powell (1991) provide an overview of these various lines of research, noting their points of divergence and convergence.

28 The intellectual antecedents of the new institutionalism are found in the writings of Selznick (1949, 1957) and Berger and Luckmann (1967). Selznick’s studies of the TVA and the Communist Party were based on notions of commitment and cooptation. For him, organizational practices became institutionalized when they were “infused with value beyond the task at hand” (Selznick, 1957, p. 16). The new institutionalism lacks Selznick’s moral tone and his focus on informal organizational relations; instead the emphasis is on the more macro cultural understandings and socially accepted “accounts” of action. Berger and Luckman’s (1967) work stressed that humans share cognitive
understandings, which emerge through daily interaction, but these cognitive categories and belief systems come to be perceived as objective, exterior structures that define social reality. Among current institutionalists, Zucker (1977, 1983, 1987) has pursued these phenomenological insights most compellingly.

29 In this respect, institutionalists owe a considerable debt to scholars working in the Carnegie tradition (Cyert & March, 1963; March & Simon, 1958; Simon, 1945), who taught us that much of organizational behavior, particularly decision making, involves rule-following more than calculation of consequences.

30 This section and the subsequent one on processes of institutionalization borrow freely from DiMaggio and Powell (1983).

31 See Scott, 1987b, pp. 125–134, for a fuller discussion of this point.

32 There is also a growing literature on the symbolic nature of accounting practices (see Carpenter & Dirschmit, 1990; Covaleski & Dirschmit, 1990; Meyer, 1986).

References


models of organizations (pp. 33–51). Cambridge, MA: Ballinger.


organizations: Culture and environment. Cambridge, MA: Ballinger.


Williamson, O. E., & Ouchi, W. G. (1981). The markets and hierarchies program of research:


