5 Issues in the Design of Process Research

PROCESS RESEARCH DESIGNS are premised on systematic analysis of event series. They presume that understanding change and development depends on understanding dynamics. As Miller and Freisen note:

> It is very hard to draw inferences about the operations of machines by looking at snapshots of a diverse array of them. It is much more instructive to watch a few machines in motion, seeing how their parts interact while examining their inputs and outputs. We can then distinguish between moving and static parts; active and passive ones. Such knowledge is essential to any designer. Theorists do seem to realize this, at least at one level. They are constantly couching their theories in dynamic terms. Unfortunately they tend to examine only cross-sectional data to generate and to test theories. (1982, pp. 1014–1015)

Methods that rise to the challenge of process research make special demands on the researcher. To identify patterns, uncover narratives, and discriminate among developmental models requires detailed longitudinal data on a number of comparable cases, as well as systematic methods for analyzing these observations and the context in which they are embedded. As a result, process research designs require revisions and extensions of traditional research methods.

In particular, we noted in chapter 1 that longitudinal case studies can provide useful data for studying change processes because they are often based on event chronologies that can be interpreted to isolate unique or important event sequences. However, most case studies do not take advantage of the power afforded by more formalized analytical methods. Case study reports are often subject to nagging questions about the validity and generality of conclusions. It is difficult for interpretive researchers to enunciate the process by which they draw their conclusions. Hence, it is difficult to assess the validity of interpretive claims or to engage them in critical discussion. At the end of the day, the interpreter can always take refuge in the claim, "I saw it and you did not," which effectively stops all discussion.

Such arguments can be avoided if researchers present their data as distinct from their data analysis, so that different perspectives can be publicly assessed and tested. Data on a series of events in the development of organizational units can be interpreted and analyzed from a variety of perspectives. A properly designed and documented process study offers researchers the opportunity to analyze the data from a wide spectrum of interests—to discover new patterns, to test hypotheses about developmental models, to understand the rich, historical context of change, and to specify the form of developmental functions.

The foundation for this approach is the construction of a data file that systematically presents the events that were observed to occur in the organizational entities being investigated. The construction of this data file should achieve two objectives. First, it should meet the same standards of reliable and valid measurement and documentation as found in good quantitative survey studies. Second, the data file should also reflect the careful attention to qualitative nuances and detail that are often found in well-conducted case studies. In other words, process research methods should harness the sensitivity of case analyses to the power and generalizability of quantitative methods. Achieving these dual goals often requires considerable ingenuity on the part of the researcher. One way to serve both ends is to utilize a mix of methods and to bolster interpretation of quantitative results with qualitative data (Miller & Freisen, 1982).

The purpose of this chapter is to introduce a set of concepts and methods for designing process research studies that can achieve these dual objectives of qualitative richness and quantitative generality. As used here, *research design* refers not just to procedures for gathering data, but to the entire enterprise of conceiving and conducting a longitudinal study. Design decisions must be informed both by the theory and questions that direct the study and by the methods of analysis that will be employed. Design cannot be understood independently of theory or analysis. In the case of longitudinal process research, theory and analysis are important determinants of the observational schemes employed, the frequency and types of data sampled, and how observational data are interpreted and transformed for purposes of pattern discovery or theory testing.

This chapter is divided into three parts, corresponding to three important concerns in process research design: formulating the research plan, establishing and validating the observational systems, and transforming

Table 5.1 Overview of Process Research Design Topics

Formulating the Research Plan

1. Tailoring the Study to the Research Question
   - Deductive and retroductive approaches
   - Developmental motors
   - Examine a single model vs. compare alternative models
   - Direct real-time observation vs. archival studies
   - Component sources of change
2. Sampling
   - Sample size
   - Sample diversity
3. Fitting the Design to Methods of Data Analysis

Gathering Longitudinal Process Data

1. Defining Incidents and Events
   - Parsing incidents from "raw data"
   - Duration and granularity of incidents and events
2. Observational and Archival Sources of Data
3. Identifying Events from Incidents
4. Reliability and Validity of Incident Construction
5. Coding
   - Design of coding systems
   - Layered coding
   - Reliability and validity of coding systems
6. Chronicles: Event Sequence Data Files

Transforming Coded Data Into Forms Suitable for Analysis

1. Types of transformations
   - Summary data
   - Bitmaps
   - Phase maps
   - Continuous variables

observational data into useful forms for analysis. Table 5.1 provides an overview of the topics in each part. While these topics are discussed in sequential order, we emphasize that they are interdependent. Designing a particular research project requires making numerous interdependent decisions and trade-offs on these topics listed in Table 5.1.

FORMULATING THE RESEARCH PLAN

The design of any research should depend on the nature of the research questions or hypotheses to be investigated. Tailoring a study to its research questions and hypotheses entails several choices, including whether the researcher emphasizes a deductive or retroductive approach, whether to observe processes directly or rely primarily on retrospective accounts, whether the research will test a single theory or compare alternative models, and how to sort out different sources of change. Another set of issues revolves around the research sample. These include sample selection, sample size, and breadth of sampling. Finally, an adequate design must satisfy the requirements of analytical methods. These three sets of issues overlap, so the following sections will build on each other.

Tailoring the Study

Deduction, Induction, and Retroduction

Deductive and retroductive approaches define a continuum of strategies researchers may employ. While ***deduction*** is familiar to most readers, retroduction, and its relationship to the more popular term, induction, may not be. ***Induction*** refers to the inference we draw from direct observation of a phenomenon which results in assigning a probability of the likelihood of an occurrence in the future. Induction leads only to probabilistic statements. ***Retroduction***, defined by Peirce (1955), refers to the inference in which we posit a theory or substantive hypothesis to explain previously observed patterns. Such theory or hypothesis is supposed to go beyond the specific case. We believe this pattern of inference more accurately describes what occurs in social scientific research than does the pattern corresponding to induction, and so we use the term retroduction as the opposite of deduction.

If a deductive approach is taken, the basic steps in designing research might consist of adopting one or more of the basic process theories described in chapter 3, operationalizing the theory(ies) into a template, and then using this (these) template(s) to determine how closely an observed process matches the theory. One could also proceed by retroduction: observe processes of stability and change over time in a few organizational entities, sort the data into meaningful categories, and then develop a theory or hypothesis to explain the observations. This theory or hypothesis would need to be verified with a different sample or on the same sample at a different time. One could also start somewhere in between, with a partial theory and flesh it out through retroduction and induction.

In the course of a research program, most researchers will move back and forth between deduction and retroduction. Deductive studies will raise questions or adventitious observations that lead to retroductive theory building. Retroduction will generate theories that stimulate deductive research.

Identification and Testing of Developmental Motors

We noted in chapter 3 that deriving or verifying a theory of development and change is ultimately dependent on assessment of whether the conditions for one or more of the four motors is met. For example, distinguishing between a life cycle or teleological motor depends in part on whether development follows a unitary, set sequence of stages (indicating a life cycle motor) or may follow multiple sequences (suggesting, but not sufficient, for a teleological motor). Deductive research would assess the fit of the two models to the data directly, while retroductive research might examine a number of cases and note the uniformity of sequence, suggesting a life cycle. The conditions listed in Table 3.3 set up an ensemble of tests that can be employed to determine which models hold for a given process. Testing becomes more complicated when multiple, layered models hold, but it follows a similar logic.

The specific context of an organizational change process is also important to examine. For example, if initial tests indicate that a teleological model offers a plausible account for startup of international organizations (as Etzioni, 1963, posited), we might also need to conduct additional tests to ensure that it holds in the format to be expected for this context. We might test whether the particular sort of goal-setting processes to be expected in international organizations holds for this case. Only if there were support for both the general and specific conditions for a motor (or motors) to hold could we conclude with some certainty that the motor was a plausible explanation. Therefore, a critical issue in the formulation of any research plan is how to garner the proper data in a form that permits identifying the context-specific conditions when each of the four motors might be operating.

Examining Single Models versus Model Comparison

Is it better to study a single model or to compare the fit of two or more models? Working with a single model of development or change has the advantage of forcing the researcher to make definite "bets" and clear hypotheses. A study focused on a single model makes it easy to test for specification errors and to improve a model that is close to fitting.

Although most research is conducted with only a single model or theory in mind, we believe a stronger case can be made for designs that compare alternative models or theories of change. First, as Stinchcombe (1972) argued, having two or more models enables the researcher to make stronger inferences by conducting critical tests of assumptions that differentiate among the models. For example, as we noted above, one test to differentiate a life cycle from a teleological model is to determine whether events follow a single general sequence or multiple sequences. Finding a general sequence supports the assumptions of the life cycle model, whereas multiple sequences would support the teleological alternative. The advantage of comparative testing is that null results for one model do not leave the researcher in a cul-de-sac where he or she knows only what is *not* likely. Studies that compare plausible alternative models have a high probability of making a positive contribution to knowledge.

Second, most "real-world" processes are exceedingly complex to a point in which they are beyond the explanatory capabilities of any simple theory found in the literature. Hence, we suggest exploring several alternative theories that capture different aspects of the same process. This not only gives a fuller picture, but also encourages rigorous and critical appraisals of theories. As Mitroff and Emshoff (1979) observed, when scholars and practitioners work from a single perspective or theory, they are subject to the temptation to unintentionally twist and rationalize facts to fit their preconceptions. One way to counteract this temptation is to develop and juxtapose alternative theories that throw one another's assumptions and weak points into clear relief. Although results are seldom clear-cut, comparative analysis will generally enable researchers to determine which theory better explains the data or how they can be combined.

The comparative approach is particularly advantageous when a retroductive research strategy is employed. Trying out several different models facilitates exploratory investigations by giving conceptual guidance to the analysis of large bodies of field data. Consideration of several models or theories enables researchers to tack back and forth between different assumptions and premises, gauging which seem to fit the data best. Each perspective serves as a comparison point for the others. It is always difficult to rule out the possibility that the data are being selectively assimilated toward a promising model. Judicious contrast of two or more promising and workable, but distinctive theories offers perhaps the best way of keeping our minds "alive" and critical of emerging accounts. It is also consistent with the principle that knowledge advances by successive approximations and comparisons of competing, alternative theories (Diesing, 1991).

Observation of Processes in Real Time versus Reliance on Retrospective Accounts

Studying organizational change processes necessarily entails collecting longitudinal data. These data can be obtained either by observing the sequence of change events as they occur in real time or by relying on historical archival data to obtain a retrospective account of the change process. Most studies of organizational change to date have been retrospective case histories, conducted after outcomes were known. Retrospective studies provide the advantage of knowing the "big picture," how things developed, and the outcomes that ensued. This post hoc knowledge is helpful for interpreting events and for constructing a narrative of the developmental process. When researchers conduct real-time observations of a change process as it unfolds, they do not have this advantage of afterthought and may miss occurrences or events that later may be viewed as critical. Until we have the compass of the entire process, we have no way of knowing what will be important and what will not.

However, there is another side to this: prior knowledge of the outcome of an organizational change process invariably biases a study's findings. This is especially true if the final assessment valorizes the outcome as a success or failure, effective or ineffective. There is a tendency to filter out events that do not fit or that render the story less coherent. This tendency to reduce the "difficult" nature of the data may result in censorship of interesting dynamics and minority views. However, historical analysis is necessary for examining many questions and concerted efforts can be undertaken to minimize bias.

A promising approach is to initiate historical study before the outcomes of an organizational change effort become known. It is even better to observe the change process throughout its unfolding. This approach maximizes the probability of discovering short-lived factors and changes that exert important influence. As Pettigrew notes, "the more we look at present-day events, the easier it is to identify change; the longer we stay with an emergent process and the further back we go to disentangle its origins, the more likely we are to identify continuities" (1985, p. 1).

For example, if the purpose of a study is to understand how to manage the formulation or implementation of an organizational strategy, it will be necessary for researchers to place themselves into the manager's temporal and contextual frames of reference. Presumably, this would initially involve conducting a retrospective case history to understand the context and events leading up to the present strategy being investigated. However, the major focus of the study would entail conducting real-time observations of the events and activities in strategy development while they occur in time, and without knowing a priori the outcomes of these events and activities.

Regularly scheduled and intermittent real-time observations are necessary to observe if and how changes occur over time. Repetitive surveys and interviews provide comparative-static observations of the organizational unit or strategy being tracked over time. Difference scores between time periods on these dimensions would determine if and what changes occurred in the organizational unit or strategy. But to understand how these changes came about, there is a need to supplement regularly scheduled data collection with intermittent real-time data. For example, this would involve observing key committee meetings, decision or crisis events, and conducting informal discussions with key organizational participants. Thus, while difference scores on dimensions measured through regularly scheduled surveys and interviews identify *if* and *what* changes occurred, real-time observations are needed to understand *how* these changes occurred. Over the years, Chris Argyris (1968, 1985) has forcefully argued that significant new methods and skills of action science are called for to conduct this kind of longitudinal real-time research. It requires researchers to identify the principal users of their research project and to negotiate the research plan with these principal users. Typically, this entails designing the research observations from the frame of reference of these users. It further implies significant researcher commitment and organizational access, which few researchers have achieved to date. As a consequence, few developmental studies of organizational change efforts have been conducted with real-time observations.

Organizational scholars often state that the purpose of their research is to develop new scientific knowledge that will improve management practice (Van de Ven, Angle, & Poole, 1989). In other words, these researchers view managers as the principal users of their research. Obviously, researchers must obtain informed voluntary consent of participants to study any organizational change process in which they might be engaged. One reason why gaining such consent and organizational access has been problematic is because researchers seldom place themselves into the manager's frame of reference to conduct their studies. Without observing a change process from a manager's perspective, it becomes difficult (if not impossible) for an investigator to understand the dynamics confronting managers who are directing the change effort, and thereby generate new knowledge that advances the theory and practice of managing change.

Furthermore, if organizational participants do not understand the relevance of a study, there is little to motivate their providing access and information to an investigator. At issue here is *not* that researchers become consultants. The issue is one of negotiating and addressing important research questions that capture the attention and motivation of scholars and practitioners alike. Clearly, the outcomes of research on an important question may not provide immediate pay-off to practitioners or academics. Many important research questions do not have clear solutions until after the research has been conducted. If solutions are well known in advance of the research, the question may be appropriate for a consulting practice or an internal management study, but clearly not for basic scientific research. Thus, at the time of designing research and negotiating access to organizations, prospective solutions to applied problems are secondary in comparison with the importance of the research question. A good indicator of such a research question is its self-evident capability (when properly articulated) to motivate the attention and enthusiasm of scholars and practitioners alike.

For example, in launching the Minnesota Innovation Research Program (see Van de Ven et al., 1989), we found that a useful way to begin longitudinal research is to conduct meetings with small groups (eight to twelve) of managers or representatives from various organizations that were about to initiate comparable organizational change efforts or ventures in their natural organizational settings. In these hour-and-a-half meetings we introduced our research question (e.g., "How and why do innovations develop over time?"), discussed why it is important to advancing theory and practice, and outlined a longitudinal real-time research strategy for studying the research question in comparable field settings over time. Participants then shared their opinions of the research question, why it was important or useful to study the question, and how the research design might be modified to make it workable in their organizational settings. The meetings concluded by thanking participants for their useful ideas and indicating that we would contact them individually to negotiate access to study the question in their organizations. Following these meetings, the research design was modified as deemed necessary, and negotiations began with individual organizations. A substantial subset of those represented at the meetings agreed to provide access to conduct the research.

Component Sources of Change

In the study of human development, Schaie (1965, 1973; Wohlwill, 1973) discussed three common sources of temporal change:

1. *Age:* The age or temporal duration of the individual at the time of measurement. This variable represents that part of development and change that is produced by unfolding biological or institutional processes.
2. *Cohort:* The set of characteristics of all individuals who are born at the

same time and go through similar developmental processes, such as classes in school. This variable represents the common historical conditions that shape the development of a given cohort.

3. *Transient:* All the temporary or immediate and noncumulative factors that influence outcomes or the dependent variables at the time of measurement.

Schaie argued that developmental research should be designed to distinguish these three effects—those that are due to age, to external factors in the history of the developing organism (cohort), or to immediate external factors (time of measurement).

It is important to design organizational change studies so they can disentangle these three sources of change. What appears to be a developmental change due to some immanent mechanism could well be due to a cohort effect or to a unique effect at the time of measurement. For example, a sudden shift in morale compared to previous levels may result from a general improvement in social mood at the time of measurement. Interpreting this as a function of solidification of a developing culture would be incorrect, though it would be easy to see why a researcher whose attention was focused only on the organization under study might draw this conclusion. In the same vein, what appears to be a general developmental pattern might be due to cohort effects, unique events occurring only to the group of organizations which were founded in a given time and place. By this reasoning, for example, it would be risky to try to generalize principles of effective development of organizational startups in the relatively benign 1950s to organizations in the "lean and mean" 1980s, because they belong to different cohorts. They started under different resource constraints, had employees with different attitudes, and a different external environment.

This is not to imply that it is impossible to develop generalizable findings concerning development and change. Rather, it is important to consider from what source observed changes originate and to rule out alternative explanations for the ones we advance. It is also important to consider the limits of our conclusions. Taking into account age, cohort, and time of measurement, as well as organization type and context, will result in more effective research designs.

Sampling Issues

Sample Selection

There is no one best sampling scheme for process research. In determining the nature of the sample to be drawn, researchers must consider the balance among four factors: sample size, sample diversity, intensity of data gathering, and cost. In general, the cost of a study in terms of time and effort increases as sample size, sample diversity, and intensity increase. To keep costs in check, compromises must be made on one or two of the three factors.

Consider sample size. The larger the sample the more valid the study and the more generalizable the results (provided cases are drawn in a representative fashion). However, large samples create problems of analysis and interpretation (Miller & Freisen, 1982). These problems stem from the time required to gather data from a large sample and from the volume of data researchers must handle. When the data are event descriptions, as process research requires, rather than quantitative responses, the cost of data gathering increases still further. If costs are held constant, there is an inherent trade-off between the intensity of data gathering—the richness and amount of data that can be acquired for each case—and sample size. A smaller sample enables more intense data collection, but also may result in nongenerality of results.

Next consider sample diversity. Should the researcher attempt to obtain a homogeneous uniform sample or a broad, heterogeneous sample? For example, if a researcher sets out to study the process of innovation development from concept to implementation, should s/he sample units that are all pursuing the same type of innovation, such as a biomedical device, or sample a wide variety of units that are pursuing different kinds of technical and administrative innovations in different industries and sectors? A case can be made for both strategies.

A homogeneous sample has the advantage of keeping to a minimum the multitude of alternative explanations for developmental processes. This is especially advantageous in the case of lengthy sequences of events, because they are particularly vulnerable to accidental or adventitious occurrences that shift the course of development. Comparing cases that are similar in as many respects as possible facilitates identifying whether change processes are due to such transient events or to more basic developmental models, but does not control for cohort effects. A homogeneous sample also facilitates the development and investigation of very precise, focused questions or hypotheses. Hence, homogeneous sampling is useful when a well-specified theory of change or development is available.

On the other hand, a case can also be made for a broad, heterogeneous sample, because it provides a better opportunity to detect whether sources of change are due to temporal development, cohort, or transient factors. Critics have questioned the wisdom of this heterogeneous sampling of innovations, since it may result in "trying to compare apples with oranges." Our position is that researchers will never know the limits where valid comparisons end and where invalid comparisons begin unless they empirically examine the broadest possible range of cases to which our definition of innovation applies.

The comparative method is perhaps the most general and basic strategy for generating and evaluating valid scientific knowledge. This strategy involves the selection of comparison groups that differ in the scope of the population and conceptual categories of central interest to the research. As Kaplan pointed out, scientific knowledge is greatly enhanced when we divide the subject matter into concepts and cases that "carve at the joints" over the widest possible ranges, types, conditions, and consequences (1964, p. 52). In this way researchers can develop and evaluate the limits of many important propositions about the subject matter.

A broad sampling scheme also permits a researcher to make empirical links between different specialties or schools of thought that have emerged for different organizational settings in which the change process occurs. For example, because organizational structures for business creation are different in small company startups, internal corporate innovation projects, and interorganizational joint ventures, it is widely believed that the process of entrepreneurship in these organizational settings must also be different. Van de Ven et al. (1989) questioned this conventional belief and proposed the plausible alternative that creating a new business entails fundamentally the same process regardless of organizational setting. If empirical evidence is obtained to support this proposition, then significant benefits and efficiencies can be gained by applying principles for business creation from new company startups to internal corporate venturing and interorganizational joint ventures, and vice versa.

In view of the trade-offs between homogeneous and heterogeneous samples, Pettigew suggests four useful guidelines for selecting cases to study:

1. "Go for extreme situations, critical incidents and social dramas." By choosing cases that are unusual, critically important, or highly visible, researchers select cases in which the process is "transparently observable." One thing researchers should be cognizant of is that such cases may have nongeneralizable features precisely because they are uncommon. It is important to assess how typical are the conclusions derived from such cases.
2. "Go for polar types." Choose cases that seem very different in terms

of the processes under study. For example, researchers might compare successful and unsuccessful program startups. Alternatively, they might choose cases that differ from patterns in earlier cases. By successive sampling of polar types, it will eventually be possible to cover the field of possible cases.

3. "Go for high experience levels of the phenomena under study." Choose cases that have a long track record of experience with a process. For example, in the studies of innovation, choose companies with reputations as highly successful innovators. This strategy may not be possible for some cases: new program startups, for example, may best be illuminated by inexperienced agencies, since they will make the mistakes and experience the learning that highlights key requirements for successful startups.
4. "Go for more informed choice of sites and increase the probabilities of negotiating access." Selecting a case for one's sample is fruitless if one cannot obtain cooperation. Often, cases must be selected on the basis of who will cooperate, rather than on grounds of optimal sampling. As Campbell and Stanley (1963) noted long ago, this introduces a bias in sampling that should be considered in drawing conclusions from the study. (1990, pp. 275–277)

While process studies, with their rich data requirements, are costly, Paul Nutt's (1984a, 1984b, 1993) strategy of gradual expansion seems to be one way of handling resource requirements. Nutt has developed a standard data collection format that he has employed over a period of 10 years to gather narratives of strategic decisions. By dint of patient and persistent pursuit of accounts, Nutt has developed a large database of decisions. While resource limitations may reduce the number of cases we can acquire at first, if we continue our pursuit, our confidence in our results can increase over the years.

The Relationship of Data and Analytical Methods

The research plan sets the parameters for the type of analytical methods that can be employed. However, some analytical methods are likely to be better than others for exploring certain research questions or testing certain models; hence, the proposed analysis should also be a constraint on the research plan.

Two important dimensions of research design which influence the choice of analytical methods are (1) the number of cases collected and (2) the number of events that are observed in the temporal development of the average case.

Most treatments of sample size in research methodology texts focus on the number of cases (not the number of temporal units) that are selected for data collection. The larger the number of cases that are sampled from a population of interest, the more generalizable are the results (provided that the cases are drawn in a representative fashion). Furthermore, in experimental designs researchers are advised to select the number of cases needed to obtain enough power from statistical tests to equate statistical significance with practical significance in hypotheses testing (Walster & Cleary, 1970). In addition to these statistical considerations, in practice, the number of cases selected also depends on the availability of sites and the costs involved in collecting data on each case.

In longitudinal research, an equally important consideration of sample size is the number of temporal intervals or events on which data are obtained from beginning to end on each case. The number of temporal intervals or events observed depends on what constitutes the "natural" flow of experience in the organizational change cases being studied. Organizational change processes vary in temporal duration and granularity. In terms of temporal duration, some organizational change processes, such as group decision making, may occur in committee meetings lasting no more than a few hours. Other change processes, such as the development of technological and administrative innovations, may span several years. Granularity refers to the preciseness or discreteness of events that are recorded throughout the temporal duration of a case being studied.

The granularity of events may vary greatly, ranging from events of such large scope that only 5 to 20 might be observed to exhaust the period of study to events of such small scope that several thousand occur during the period under study. Event granularity typically increases with the microanalytic detail of the change process being investigated. Psychological studies of change in individuals tend to sample fine-grained events, such as speech acts, time allocation, or role behavior, whereas sociological or economic studies of organization change tend to sample more coarse-grained events, such as structural reorganizations, mergers, or stages of organizational growth.

Another consideration is the cost of coding events. Events that require a great amount of time and effort to observe and code are likely to be observed in shorter sequences than less costly ones. For example, it might require a great deal of time and effort to compile the complete record of all the transactions between researchers in the field of gallium arsenide semiconductor research; however, to get a shorter (and probably just as representative) record of some transactions by coding which researchers attended the same meetings and which labs entered into joint projects would

Table 5.2 Analytical Options for Different Types of Process Datasets

	FEW EVENTS	MANY EVENTS
Few Cases	Summary Case Studies	Summary Case Studies Phasic Case Studies Time Series Analysis Markov Analysis
Many Cases	Multivariate Analysis	Multivariate Analysis of Summary Data
	Phasic Analysis with Optimal Matching	Phasic Analysis with Optimal Matching
	Event History Analysis	Markov Analysis Time Series Analysis

be much less costly. Due to the inherent trade-offs between the temporal duration and granularity of events that can be sampled, studies of relatively brief change processes can afford to utilize categories that code fine-grained events, while studies of lengthy change processes tend to adopt categories that tap coarse-grained events.

Crossing these two dimensions of number of cases and number of events yields a 2 × 2 table, each of whose cells corresponds to a different set of appropriate data analysis methods, as shown in Table 5.2. Each method will be mentioned only briefly here, to give an overview of the relationships between sampling schemes and methods of data analysis. Data-method relations will be discussed in more depth in subsequent chapters.

Studies consisting of *few cases, few events* are not suited for most of the methods discussed in this book. However, there are many important phenomena in which the focal events occur only a few times. For example, consider a comparative study of strategic decision making where the sequence of search, screen, and choice activities is the object of investigation. As Nutt's (1984a & b) studies show, there may be relatively few instances of these activities for each decision, resulting in short event series. Other studies may focus on critical incidents, unusual or uncommon events such as conflict or key turning points in development of industries. Provided there are enough cases for systematic comparison and induction across the instances, Yin's (1984) comparative case study designs are useful in this situation.

Studies with ***many cases, few events*** enjoy more options. ***Summary measures*** for each case can be derived by collapsing the data along the time dimension (e.g., counting the number of conflicts that occur during innovation regardless of when they occurred), or through use of surrogate measures of temporal order (e.g., did the conflict occur during the first or second halves of the innovation process?). Such measures can then be treated as variables in traditional statistical methods. One method to preserve information about temporal order which clusters cases with similar sequences is *phasic analysis*. As discussed in chapter 7, once clusters of phasic sequences have been derived they can serve as the basis for variables that can be entered into traditional statistical analyses. Alternatively, Tuma and Hannan (1984) discuss how *event history analysis* can be used to determine when critical events occur, provided the length of time until they occur is recorded. Supplementary analysis can in some cases divulge causal factors underlying event occurrences (Willett & Singer, 1991).

Several additional avenues are open for studies with *few cases, many events*. Comparative analysis of *qualitative case studies* using Yin's designs are one option. Chapter 7 discusses how events can be parsed into phases that represent coherent periods of activities subsuming two or more events in sequence. These phases can then be used as bounded units to provide temporal divisions in case studies, as Holmes (1992, 1997a, 1997b) did in his studies of hostage-taking situations, and Polley and Van de Ven (1989) did in the study of a biomedical innovation. Various types of *time series* analyses can also be used when many events are available for each case. As discussed in chapter 8, these generally involve transforming the event series into some continuous form. Chapter 6 discusses the application of *stochastic modeling*, which preserves the categorical qualities of the event series and enables us to track temporal dependencies among events.

For studies with *many cases, many events* a number of powerful statistical techniques are available. As with the many cases, few events situation, simple *descriptive summaries* of the frequency with which coded events occur provide useful displays for examining stages or phases in the developmental progression. However, with such pooling of the data one can lose the temporal order of events that figure prominently in most process research studies. As discussed in chapter 7, the technique of *optimal matching* can be used to derive measures of similarity among large samples of phase sequences derived through phasic analysis for the cases. These measures can then be analyzed in at least two ways. First, they can be used as input to cluster analysis and multidimensional scaling (MDS) techniques that can identify clusters of similar sequences; the resulting clusters can then be used to define variables for causal or correlational analysis, as in Poole and Holmes

(1995). Second, these distances can be used to test for causal factors that create the differences between pairs of sequences.

Chapter 8 discusses how *trend analysis or multiple time series methods* can be used to identify patterns of change across many cases, provided the events can be used to define continuous variables. Chapter 6 also shows how *stochastic modeling* of multiple cases can provide maps of temporal dependencies among events. Causal factors leading to such dependencies can then be identified.

As noted previously, the relation between the two dimensions and analysis is a two-way street. A researcher who chooses a certain method of analysis will want to design her/his research so that the proper number of cases and events are sampled. However, in some cases the researcher is limited by what is feasible. Then it is important to select an appropriate method and find some way of creating the type of data it requires from what is there.

Concluding Comments on Formulating the Research Plan

The issues discussed here do not exhaust those that confront the researcher, but they represent several critical choices. Other good sources for the design of longitudinal organizational studies include Galtung (1967), Kimberly (1980), Miller and Freisen (1982), and the *Organization Science* special issues on longitudinal field research (1990, Volume 1, Numbers 3 and 4). Though it poses thorny design problems, longitudinal process research is a worthwhile challenge. It offers the best chance we have of evaluating the stories that hold our theories together.

At the heart of any design is data gathering. This section has considered general choices we must confront in planning how to get our data. The next section discusses in much greater detail techniques for gathering and validating data.

Data Gathering in Process Research

Consider the challenge confronting researchers who attempt to implement a typical process research design. Over one to three years of real-time field study, they use a survey to collect quantitative and qualitative data every six months, conduct interviews with key managers and technicians, attend and make direct observations of regularly scheduled (monthly or bimonthly) organizational meetings, and maintain a study diary to record and file frequent informal discussions with participants, organizational memos and reports, and stories in trade journals or newspapers about the innovation.

As studies such as these proceed, the volume of data mounts astronomically and quickly overloads the information-processing capacity of even the most insightful mind. Rigorously drawing inferential links between the data and theory require methods which go beyond subjective "eyeballing" of raw qualitative data to identify patterns. Unfortunately, data analysis methods are rarely reported in detail in published case studies or ethnographic reports. One cannot ordinarily follow how a researcher got from hundreds of pages of field observations to the final conclusions, even though the research may be sprinkled with vivid—yet idiosyncratic—quotations from organizational participants. The sheer mass of data overload our information-processing capacities and threaten us with what Pettigrew calls "data asphyxiation." Confronted with such a mass of raw data, how should the researcher convert it into a form useful for developing and testing process theories of organizational change or innovation?

Data gathering in process research can be divided into several distinct operations. First, researchers must get the *raw data*. Then they must identify *events* in this raw data that capture key aspects of the process. These events are then arrayed in a *chronicle* of the process. In moving from raw data to events to chronicle, *coding* is often employed. This part of the chapter will offer guidelines for each of these operations. This will be followed by a section discussing the final operation in the preparation of data for process analysis, its transformation into a *final form* suitable for various process analysis methods. The appendix to this chapter describes step-by-step the process of building an event sequence file in the CIP project. It covers the technical aspects of the research process that will be the focus of the remainder of this chapter.

Longitudinal observation depends on a set of categories or variables to describe the developmental process. Whether they are implicit or explicit, these concepts help to focus observation of the change process; one cannot study everything. Category systems provide the "measurement" necessary to connect theoretical models of development with empirical events. When the model(s) is known beforehand, category development proceeds as a form of operationalization of theoretical constructs. In the process of developing a category system, the constructs themselves are respecified and fine-tuned, but generally category development is a top-down process in this case. On the other hand, when study of organizational development processes is at an embryonic stage, these initial categories emerge as "sensitizing constructs" for conducting exploratory research. The categories become clearer as they are put to use, and eventually they can be codified into a formal scheme.

For example, the MIRP studies began with five "sensitizing categories"

Table 5.3 Evolution of Innovation Concepts During the Minnesota Innovation Research Project

CONCEPT	INITIAL DEFINITIONS BASED ON THE LITERATURE	WHAT THE PROJECT FOUND
Ideas	One invention to be realized	Reinvention, proliferation, reimplementation, discarding, and termination of many ideas
People	An entrepreneur with a fixed set of full-time people	Many entrepreneurs, distracted fluidly engaging and disengaging in a variety of roles over time
Transactions	Fixed network of people/firms working out the details of an innovation idea	Expanding and contracting network of partisan stakeholders converging and diverging on innovation ideas
Context	Environment provides opportunities and constraints on the innovation process	Innovation process constrained and created by multiple enacted environments
Outcomes	Final result orientation: A stable order	Final results may be indeterminate; multiple in-process assessments and spinoffs; integration of new order with old

that seemed important to innovation development: ideas, people, transactions, context, and outcomes (Van de Ven et al., 1989). As the study progressed, these assumptions and concept definitions changed substantially and became successively clearer over time. Table 5.3 compares the starting assumptions related to these concepts, as reflected in the literature at the time (summarized in the left column), with how the MIRP researchers came to view them as a result of two years of field studies (in the right column). The latter disclosed a different reality from the rather orderly and simple concepts of the former. As this example illustrates, the development of research constructs involves an iterative process of developing initial conceptual categories, observations, and progressive redefinition and refinement of categories. This iterative process underlies many of the data collection and coding steps discussed in this section.

Incidents and Events

As stated before, a theory of development consists of statements about the temporal sequence of events that explain an observed stream of incidents or occurrences. To make such a theory operational, and hence testable, Abbott (1984) emphasizes that it is important to distinguish between an *incident* (a raw datum) and an *event* (a theoretical construct). Whereas an incident is an empirical observation, an event is not directly observed; it is a construct in a model that explains the pattern of incidents. For each event one can choose any number of incidents as indicators that that event has occurred.

This definition implies a particular kind of relationship between incidents and events. Incidents are descriptions of happenings, documentary records of occurrences. Events are meaningful parsings of the stream of incidents. They are constructions based on a more-or-less systematic interpretation by the researcher of what is relevant to the process. The stream of incidents, a first-order construction, is translated into a sequence of events, a second-order construction. This implies that some incidents may be emplotted in different ways, utilized as constituents of different events. And this is not just a matter of different interpretations, that is, that incident k indicates event A versus event B, as a brief consideration of the nature of events will show.

Events may vary in several respects. First, they may differ in temporal duration. For example in the CIP case, a meeting with a potential partner is a bounded event of three to four hours duration, but the clinical trial of a design to meet FDA approval was an event that stretched over a much longer period. Events may also overlap or nest. For example, the meeting in question might occur during the period covered by the clinical trial. Or a relatively long event such as the clinical trial may be decomposed into shorter constituent events, such as a meeting to plan the trial, several different tests conducted as part of the trial, and an evaluation session, all of which are nested within the trial. Events may also differ in spatial extension. The meeting with the partner occurred in one room and stretched to a meal at a restaurant, whereas the clinical trial involved implantation in the hospital and monitoring the effectiveness of the device in the day-to-day life of the subjects. The clinical trial stretched over a much wider space and involved more people than did the meeting.

That events may differ in temporal and spatial scope suggests that incidents may well indicate more than one overlapping event. For example, the meeting with the potential partner can indicate the event "meeting with a partner," but it may also indicate a longer event, "negotiation with firm Q regarding partnership." Events may be embedded within other, different types of events of larger scope. Both levels may be important for understanding the change process, because interleaving narratives clarify it better than either narrative could on its own. Abbott gives an example from his studies of the rise of professions in society:

> I once set out to explain why there are no psychiatrists in American mental hospitals. The exodus, which dates from 1900–1930, reflects not only the rational individual mobility decisions that are specifiable annually, but also outpatient community developments that are specifiable only decadally, and changes in knowledge and social control taking place over even longer periods (1992, p. 439).

Another complication is the possibility that the incident-event relationship may change over time (Abbott, 1984). As we noted in chapter 2, the significance of events may change as the process unfolds. The same change is possible in incident-event relations. For example, the first time a potential partner is encountered may signal the expansion of a program such as CIP, whereas the sixth encounter with a potential partner may signal desperation for ideas or resources.

To sum up, events are constructs indicated by incidents. However, the indication relationship is more complicated for qualitative data than it is for quantitative scores. Psychometrics and scaling theory presume a uniformity across respondents and responses that may not be the case for the data used to define events. What quantitative analysis would relegate to the domain of "error variance" may be quite an important nuance for qualitative analysis.

A final consideration in defining events is identification of the central subject(s) that the events refer to. Narrative analysis depends on defining a central subject for the narrative. This subject may be many things—a person, group, organization, idea, product, innovation, interorganizational field, almost any social unit that develops or changes over time may serve as a central subject. As we noted in chapter 2, the nature of the central subject may change over the course of the narrative, so it is imperative to have a clear sense of who or what the central subject is. This involves defining its essential characteristics and cues that will enable the research team to track it over time.

For example, in the CIP study, the central subject was the business startup within 3M, which revolved around a technology-based product, the cochlear implant. The development of the central subject and technology were intertwined, so anything that was relevant to either was included in the incident record. Data was also considered relevant if it concerned any of the principals in the innovation unit and any competing technology with the 3M technology. The resulting categories of events and interpretations of events are detailed in the appendix to chapter 4. Decisions that data was relevant hinged on its connection to any group of people or technologies; these served as the cues for preliminary inclusion of the incident in the data file.

There is, of course, another way to interpret this example. Instead of one central subject, the CIP case had two—the startup and the technology—or even more than two—the startup, the technology, and the three or four principal movers of the innovation. Narratives with multiple central subjects are possible. However, they may become quite tangled and complicated as the multiple strands of narrative are tracked.

Defining an Incident: A Qualitative Datum

In quantitative survey research, the datum is typically assumed to be sufficiently clear to require no explicit treatment. However, this is not the case with qualitative data, where it is important to define a *datum,* which is the basic element of information that is entered into a data file for analyzing temporal event sequences in the development of organizational entities.

In survey research, a *quantitative datum* is commonly regarded to be (1) a numerical response to a question scaled along a distribution (2) about an object (the unit of analysis) (3) at the time of measurement, which is (4) entered as a variable (along with other variables on the object) into a record (or case) of a quantitative data file, and (5) is subsequently recoded and classified as an indicator of a theoretical construct. In comparison, we can define a *qualitative datum* as (1) a bracketed string of words capturing the basic elements of information (2) about a discrete incident or occurrence (the unit of analysis) (3) that happened on a specific date, which is (4) entered as a unique record (or case) in a qualitative data file, and (5) is subsequently coded and classified as an indicator of a theoretical event.

Parsing Incidents from "Raw Data"

As the definitions just stipulated indicate, the basic element of information in a qualitative datum is a bracketed string of words about a discrete incident, while in a quantitative datum the element of information is a number scaled along a predetermined distribution of a variable. Raw words, sentences, or stories collected from the field or from archives cannot be entered into a series until they are bracketed into a datum (data). Obviously, explicit decision rules are needed to bracket raw words. Many diverse types of rules are possible and their common denominator is that they should reflect the substantive purposes of the research.

In the case of CIP, the decision rule used to bracket words into a qualitative datum was the definition of an incident that occurred in the development of the innovation. An *incident* was defined as a major recurrent activity or whenever changes were observed to occur in any one of the five core

Table 5.4 Examples of Incidents and Event Codes from CIP Database

INCIDENT DATE:	INCIDENT NUMBER: 312 INCIDENT DATE: 06/01/87	INCIDENT NUMBER: 313 INCIDENT DATE: 06/01/87	INCIDENT NUMBER: 314 INCIDENT DATE: 06/18/87	INCIDENT NUMBER: 315 INCIDENT DATE: 06/25/87
Incident	MN firm executive states he will not support TAP beyond 1988. The MN firm has been investing about $4M per year in TAP, and the MA firm is only spending between $1 and $1.5M. The MN executive thinks the MA firm should contribute more. SK met with JB and offered several options such as donating modules, writing a check for $1M, or taking less royalties. JB will see iff the MA firm is still interested.	June SBU meeting canceled.	Emergency meeting conducted of MN firm's core TAP team to discuss restructuring finances as a result of recent internal management review. Items for discussion included 10–15% across the board reductions, omission of diagnostics, assumption of improved electronics by 1/1/88, 70% of sales by 1995 will come from tubesets manufactured outside of the MI firm, and no significant research beyond LDL and immune complex.	Joint administrative review of TAP by MN and MA firm executives. MN executive suggests bringing in a third partner to reduce financial burden. He suggests that TAP be spun off into a joint venture with a third partner. No conclusion reached. The MA firm executive asks "Why has my partner blinked?" He questions if the MN firm is really committed to TAP. The MN firm executive suggests that it is just an issue of financing and additional opportunities for investment.
Data Source	Phone calls with SK and JB, 6/1/87.	Phone call with SK 6/1/87.	Internal memo of 6/10/87 and 6/18/87 meeting notes.	AHV notes of 6/25/87 administrative review meeting
Core MIRP CODES	Outcome-negative Context-internal Context-external Transaction, Contraction	Context-internal Contraction	Transaction Idea-core Context-internal	Transaction Context-internal Context-external Outcome-negative

concepts in the MIRP framework: innovation ideas, people, transactions, context, and outcomes. Examples of incidents from a business startup case are shown in Table 5.4. The definitions for incidents can be found in the appendix to chapter 4, and some of the guidelines for defining these incidents can be found in the appendix of this chapter.

When each incident was identified by MIRP researchers, the bracketed string of words required to describe it included: date of occurrence, the actor(s) or object(s) involved, the action or behavior that occurred, the consequence (if any) of the action, and the source of the information. As with any set of decision rules, some further subjective judgments were involved in defining innovation incidents in an operationally consistent manner. These were resolved in discussions among the researchers that took additional information and special circumstances into account.

Duration and Granularity of Incidents

Decision rules may vary in the level of specificity and the temporal duration of incidents they construct. Some rules specify fine-grained definitions of incidents that interpret each action as a separate incident; others adopt coarse-grained definitions that require longer episodes for incidents.

The proper granularity of incidents should depend on the rates of development of various kinds of innovations, and the differing research questions associated with these rates. For example, two MIRP researchers working on a different innovation (Knudson & Ruttan, 1989) found that the rate of hybrid wheat development is governed by biological laws that require several decades in order to move from basic research through technology development to market introduction. They observed that hybrid wheat's innovation process has been following this "biological time clock" since the late 1950s. In studies of biomedical innovations such as CIP (see also Garud & Van de Ven, 1989), the rate of development appears to be governed by an "institutional regulation time clock," in which the design, testing, and commercial release of devices entailed extensive review and approval steps by the U.S. Food and Drug Administration, sometimes lasting five years. Finally, rates of development of other processes, such as group decision making (Poole & Roth, 1989) or the development of novel administrative programs (Bryson & Roering, 1989; Roberts & King, 1989) are more rapid and appear to be limited only by entrepreneurial time and attention. As these variations suggest, researchers need to develop operational procedures for tracking developmental processes that are congruent with the temporal scope of development and the corresponding granularity of incident detail appropriate to the organizational entities being examined.

Observational and Archival Sources of Data

In an earlier section we discussed planning research to study change processes in real time or in retrospect. Here we expand on this by addressing the advantages and disadvantages of gathering data from direct observation and from secondary archival sources.

Incidents can be identified either through direct observation or through archival research. By direct observation we mean study of the process *as it unfolds,* through participant observation, interviews with key principals and informants, and study of emerging records and documents. The hallmark of direct observation is that the research is carried out contemporaneously with the unfolding process. By archival research we mean study of the process *after it has occurred,* through analysis of documents and records, retrospective interviews, bibliometric analysis, and other historical-reconstructive methods. The two data-gathering methods are often used together, as when researchers conduct a historical case study prior to a direct observational study in order to better understand the context and meaning of current incidents.

With direct observation, researchers have the opportunity to judge immediately how adequate their data is and to follow up on questions or uncertain areas. Direct observation also grants flexibility to the research team. If necessary, researchers can expand or contract data collection activities. They can alter procedures to solve emerging problems and respond to unique opportunities. Being close to the process is also likely to give researchers a special feel for what is immediately significant at any given moment. By experiencing the process with key actors, researchers have a chance to gauge its emotional tenor and impact on participants far better than researchers removed in time and space.

There are also disadvantages to direct observation. Adjusting data collection procedures later in the game means that earlier data are incomplete. Researchers are then faced with the task of reconstructing earlier observations. The very immediacy of the process in direct observation may also blind researchers to significant patterns because they cannot see the forest for the trees. Often it is only after the fact, when the process can be understood as a whole, that key drivers and turning points can be discerned. Finally, studying a whole process may be very time-consuming. Direct observation can in some cases require a commitment of several years, with no date certain for an end to the process.

The quality of data from direct observation depends on several factors:

1. *Access* to important sources, activities, and documents is critical. Being available to study a process in real time means little if researchers are not able to observe critical events. Private discussions, secret actions, and classified documents may contain the key to understanding the process. To the extent that researchers miss or are not aware of these, their conclusions are limited. Of course, the catch-22 in this admonition is that in many cases researchers are not aware of what they are denied access to.

2. *Distortions and biases* may be introduced by those with a vested interest in the process. In some cases they are intentional, designed to make the informant look good or to protect him or her. In more cases, biases and distortions are simply the product of cognitive and social processes that simplify, consolidate, and assimilate to the expectations of subjects' recall and reportage.

3. *Data may be so bountiful* that researchers cannot capture it properly. For example, a sleepy meeting may suddenly turn into a major decision-making session, and a single researcher lulled into complacency may be unprepared to record key arguments or conclusions. Information overload may cause loss of data or inaccurate recording into conceptual categories. Investigators can take these factors into account when designing observational research. If these problems cannot be counteracted, they can at least be factored into the conclusions.

Archival research, the other data-gathering strategy, also has several advantages. Researchers have the benefit of hindsight, which lends valuable perspective to their efforts. Others' judgments of the process and its outcomes can inform the researchers' analyses. Provided adequate records are available, the data can be combed and recombed, coded in multiple layers, and otherwise interrogated until the full story emerges. A second advantage of archival research is that it takes much less of the researchers' time: very lengthy processes can be investigated in comparatively brief periods. Researchers with access to a good record may be able to explore a process lasting decades in a year's research, as Knudson and Ruttan (1989) did in their study of the emergence of hybrid wheat strains.

A major disadvantage of archival research is its limited flexibility. Researchers must make do with what has been preserved, either materially or in informants' memories. There is no possibility of adding measures or observations; data that is lost is gone forever. Then too, knowing how things turn out can bias one's perceptions and interpretations. A researcher studying a failed startup may well have a tendency to find events that seem to lead "downhill" more salient than those which are commonly thought to lead to success. The former may find a prominent place in the chronology, while the latter are passed over or explained away. Distance in time and space also decreases the researchers' ability to identify with and to empathize with participants. While much has been written about procedures for hermeneutic reconstruction of activities and events, this is a poor substitute for experiencing or witnessing them first hand (or even second hand, if contemporaneously).

The quality of data used in archival studies depends on several factors:

1. *The extent to which relevant records are kept* varies greatly. Some organizations and people are fastidious about keeping records, whereas others are not. Some activities lend themselves to recording better than others. For example, there are often legal requirements that minutes of board meetings be kept, but there are no such requirements for dyadic conversations among key actors.

2. Even if relevant records are kept, there may be *loss of data due to archival practices.* Confronted with a mass of historical data, archivists must appraise what is worth keeping and what is not. They may either discard (the worst case) or not catalog or organize data (a bad case, but sometimes recoverable), depending on whether they judge the information and its creator to be significant and worth study (see Baer, 1997, for some of the theories that might guide archivists confronted with organizational records).

3. *Contradictions among records* also must be reconciled. Different observers may have different perceptions of the same event, and one set of records may yield a different set of incidents than another. Researchers differ in their attitudes toward contradictions. Historians have long worked to refine methods for adjudicating and resolving inconsistencies in the historical record (Walsh, 1967). Pettigrew (1990) exemplifies a different tack: Acknowledge the validity of different perspectives and accept the fact that there are multiple layers of socially constructed reality. He argues that our understanding will be enriched by taking into account the alternative views that invariably develop, because they indicate the differential views that shape the development of multiply determined processes.

4. *Biases and self-serving distortions* occur in records just as they do in direct observation. Records may be intentionally destroyed or falsified, and they may be unintentionally neglected. Key events may not be recorded at

all. There is the same catch-22 for archival data as well. Researchers dealing with incomplete data may never get a clue that it is incomplete.

5. *Inconsistent quality of records* is a common problem in archival research. For some events excellent, clear, detailed records survive, while for others the records are spotty and confusing. As a result, there is variation in the degree of confidence that can be placed in the accounts of different incidents or events in a historical record.

Archival research can plan to minimize these problems, or at least, to acknowledge and allow for them. Some good sources of advice on archival research strategies are Baer (1997), Hill (1993), Elder, Pavalko, and Clipp (1993), and especially O'Toole (1997).

Reliability and Validity of Incident Construction

It is important to establish the reliability of classifying raw data into incidents. An equally important, though often neglected issue, is the validity of this bracketing procedure (Folger, Hewes & Poole, 1984; Poole, Folger, & Hewes, 1987). Researchers often assume that the meaning of incidents is clear and that establishing reliability is equivalent to showing clear meaning of codings. However, attaining reliability among coders simply indicates that the meaning of incidents is clear to the particular group of researchers who designed the coding system. Of course, this does not mean the classifications correspond to the way participants see them. Thus, it is important to distinguish between classifications meaningful to researchers and those meaningful to organizational participants. These two types of classifications may not overlap, and researchers must be clear about what sorts of claims they make about the meaning of the incidents they record. It is necessary to test empirically whether researchers' classifications are consistent with practitioners' common perceptions of events. If the evidence indicates inconsistency, then no claims about the meaning of events to the participants are valid. Researchers can still sustain claims about the meaning of the incident from their theoretical position, but no claims about the "social reality" of the event are appropriate.

Two basic procedures are useful to enhance the reliability and validity of the incidents entered into the qualitative data file. First, the entry of incidents from raw data sources into a data file can be performed by at least two researchers. Consensus among these researchers increases the consistency of interpretations of the decision rules used to identify incidents. Second, the resulting list of incidents can be reviewed by key informants. It is useful to ask these informants if any incidents that occurred in the development of the organizational change process are missing or incorrectly described. Based on this feedback, revisions in the incident listings can be made if they conform to the decision rules for defining each incident. Typically, these two steps result in a more complete listing of incidents about an organizational change process.

In getting this information from subjects it is important to be sensitive to where the brackets are being put, that is, whether the incidents divide up the flow of "raw data" into units which are sensible and natural to participants. Studies of the parsing of social behavior into units indicate that people are capable of attaining high agreement on their unitizing of interaction (Planalp & Tracy, 1980). However, these studies also show that unitizing depends on the goals of the actor, so it is probably a good idea to check, with several informants if possible, whether incidents represent "wholes" or should be subdivided or combined. Unitizing is likely to be more difficult in cases when the investigator has to identify incidents directly from ongoing interaction or occurrences than in the case of incidents defined through interviews, which are likely to already come in "predigested" bits.

Finally, it is important to recognize that the resulting list of incidents, no matter how painstakingly wrought, is only a sample of occurrences in the development of an organizational entity. In the case of studies based on records and interviews, the sample is limited to what informants know and can recall and what the records contain. Even with real-time field observations, it is not humanly possible to observe and record all possible incidents. Researchers are limited to what they can observe and the particular layers of meaning they can pick up. Thus, as in classical test theory of item sampling (Lord & Novick, 1968), the incidents represent a *sample* of indicators of what happened over time in an organizational change process. This has important implications for event identification.

Identifying Events from Incidents

As they stand, the incident listings are only one step above the raw material for analysis, because each incident is just a qualitative summary or indicator of what happened. The next step is to identify theoretically meaningful events from the incident data. We will discuss the coding procedures that can be used to translate incidents into event indicators in some detail in the next section. Before describing these procedures, we will discuss a prerequisite issue of moving across levels of abstraction between indicators and theoretical constructs.

Indicators can correspond to a theoretical construct in several different ways. As is the case in psychometrics, some indicators are usually better representatives of the construct than others. Some indicators are also easier to measure or detect than others, resulting in differential reliability across indicators. Sampling error also presents a problem, with the result that some indicators will not be included in a given sample, while others will be oversampled. In the study of ongoing processes, there is the troubling possibility of "right censorship," that is, that relevant incidents are not being sampled because they have not yet occurred. So we have better knowledge of the beginning of an event than of its end in many cases. Such problems suggest a need to think through the indicator-construct relationship carefully.

Several options exist to map indicators to constructs (Abbott, 1984). The first and most straightforward strategy is to give all indicators of an event equal weight. For temporally extended events, this means that the event commences when the first indicator is observed and continues until the last occurs. For example, a "resource controller intervention" into a new business startup might begin when the resource controller first contacts the unit and continue through all contacts until the outcome of the intervention for the business is noted. This event would then stretch across a number of incidents. For short-term events, this means that they occur as many times as indicators are observed. For example, idea development might be defined as occurring each time a change in an idea occurs; there would be as many events as there were incidents in which a previously held idea was changed or a new idea was advanced. One weakness of this strategy is that it does not allow for error; it assumes every incident is observed reliably and validly. Another weakness is that the strategy assumes each indicator is equally critical to the event, which is at odds with the narrative approach. The advantage of this strategy is that it takes the data at face value and does not attach any particular assumptions to incidents. It also avoids the uncertainties and possible biases involved in trying to weight incidents in terms of their importance to the event.

A second strategy is to make judgments concerning whether indicators signal an event on a case-by-case basis. This is the common approach in historical studies, where the researcher establishes whether an event occurred by considering the indicators in context. The researcher uses her or his judgment and contextual knowledge to determine the occurrence and duration of events, as well as other characteristics, such as intensity and impact. For example, resource controller interventions could be defined by having one or more researchers read a string of incidents connected with the resource controller and, based on their knowledge of the case and its context, make a judgment of when the event began and ended. This judgment may not include some of the incidents that seem to relate to the resource controller in the definition of the event. One advantage of this strategy is that it gives a much more nuanced reading of the process. A large number of attributes of an event can be discerned, including narrative properties such as how much impact it had on subsequent events. This approach assumes the researcher is able to make holistic judgments based on context much more adroitly than could any algorithm. One possible problem stems from the biases introduced by this holistic knowledge and the value judgments that may come with it. A second problem is the difficulty of making such interpretations consistently across a large sample of processes. Carried to an extreme, this approach could reduce an event sequence to the researcher's preferred story or byline.

A *third strategy is to use indicators* as *some measure of central tendency*, such as the median occurrence of an incident, the mean time when incidents occurred on the timeline, or the mean number of times an incident occurred within a given time segment. For example, resource controller intervention might be defined as an event at the midpoint (median) of contact between the business startup and the resource controller on a given topic. This enables a "single resource controller intervention" event to be pinpointed in the sequence. Even more subtle rules are possible, such as "the first time two medical schools are founded within ten years of each other will indicate that medical education is institutionalized" (Abbott, 1984). The advantage of these indicators is that they allow for error in the data. Except for the degenerate case when only one or two indicators are observed, they correct for unreliability and sampling error by relying on composite indicators. Just as multiple items make a scale more reliable, so multiple indicators correct for error. This strategy is only workable, however, when multiple indicators are available.

One workable option is to employ strategies 1 or 3 for initial classifications and then use strategy 2, interpretive judgment, to adjust the event sequence. Using conceptual and practical reasoning to correct for the bluntness of classifications that employ objective rules can greatly improve the accuracy of process analyses. Insofar as possible, this should be done at the "local" level, avoiding the use of knowledge about the overall sequence or outcomes to guide reclassifications. This will minimize the intrusion of biases and value judgments into the event classifications.

The process of classifying incidents into events is also likely to spark a return to the "raw data" by raising questions about the incident list. This may stimulate the collection of more data to fill in gaps in the account or to answer troubling queries. It may also lead to a revision of the incident list itself, as problems with event identification lead researchers to rethink their incident descriptions. Handled carefully, this circling from data to incidents to events and back again can greatly enhance the fidelity of the data in process analysis.

The appendix to chapter 4 provides examples of event categories developed by MIRP researchers. These were defined after repeated discussions over a period of about a year among MIRP researchers. A preliminary set of categories was identified, then tried out, then refined, and tried again, in five cycles of this process until the final set of categories emerged.

Coding: A Key Method in Process Research

Design of Coding Systems

There are several approaches for tacking back and forth among theory, category, and data to develop coding schemes such as the one presented in the appendix to chapter 4. *Inductive* approaches go first to the data—the incidents—and sift through the various instances, deriving categories from the ground up, using the constant comparative method. This is quite time-consuming and may make it difficult to link observation with current theory. However, it is likely to lead to interesting innovations. *Deductive* approaches use theory to specify expected categories, which are then written into rules.

In practice, the two approaches are often combined using a *retroductive* approach that often includes a literature search to derive a synthetic category scheme that seems to fit what the researcher sees in the data, then adjustment of categories in view of what is workable and informative after trying them out on the data. Poole used this approach to derive his decision functions coding system (Poole & Roth, 1989a & b). Another retroductive approach is to generate a set of categories based on theory and then refine and adjust them as they are applied to data. This permits the theoretically driven scheme to grow and to adapt in response to the exigencies of the data. Bales (1950) used this approach in developing Interaction Process Analysis.

Key choices in generating coding systems for longitudinal process research include (1) the type of unit to be used; (2) the type of coding to be made; (3) the latitude of judgment accorded to the coders; (4) univocal versus multifunctional coding; and (5) the domain of meaning the classifications are meant to tap.

1. *Type of unit.* This involves two issues, selection of a "natural" unit versus an "artificial" unit, and the granularity of the unit. Natural units are those whose bounds are set in the phenomenon itself, such as a speaker's turn, or a quarter's performance in a firm, or a meeting. Artificial units are those specified by the researcher, such as a 30-second period of discussion or a summary of climate concerns for one-month periods. Artificial units are easier to delimit and are useful if real time is the central metric of the analysis, but they may require subjects to distort the meaning of events to them. As stated previously, the granularity of the unit may vary greatly as well, since events differ in duration and scale. Granularity probably needs to be established independently for each phenomenon studied.

2. *Type of coding.* Several different types of codings are possible. The most common, of course, is to code events into the qualitative categories typical of most classification systems. However, it is also possible to have coders assign a numerical rating to an incident, for example, coding it on a 5-point scale, reflecting intensity of conflict exhibited in the event. Such global judgments can be just as reliable and valid as classifications, though they tend to gloss details and lead to distortions if the unit to be coded is a large or lengthy one (Poole et al., 1987).

3. *Latitude of judgment accorded to coders.* Coding systems incorporate the coder's knowledge of language and social context into the measurement process. Though the researcher would probably not want to eliminate use of this background knowledge, there are variations in the degree to which the coder is allowed to exercise this knowledge in making judgments. At one end of the scale are mechanical devices and computer software that greatly constrain human judgment. For example, content analysis programs based on dictionaries automatically classify sentences. A less-stringent alternative specifies a choice tree that presents a complete set of classifications and a series of simple binary questions that "lead" the coder to the proper classification (e.g., Anderson, 1983). The most common method emphasizes utility and pragmatic impact rather than logical completeness. Researchers compile as complete a list of categories as possible (or necessary), write enough rules to enable coders to recognize and distinguish the categories, and rely on coders' native knowledge and skills for the rest. This strategy is advantageous for complex meanings for which it is difficult, if not impossible, to develop complete classification rules. However, reliance on coder judgment makes the procedure harder to control and may result in inconsistent classifications if categories are not sufficiently defined.

4. *Univocal versus multifunctional coding.* Many traditional sources recommend that coding categories be mutually exclusive and exhaustive. But assigning a single code to each incident (mutual exclusivity) may be problematic if more than one social function is served by an act or if more than one meaningful thing happens in the same incident. In view of the fact that social life has multiple layers of meaning and that more than one consequence can be taken from any incident, multifunctional coding may be preferred for many applications. For example, the CIP study utilized multifunctional coding. As the appendix to chapter 4 indicates, incidents could be assigned a number of different codes and the same incident could receive more than one code from a given category.

The notion of multifunctional coding fits well with Abbott's concept of processes as multiple, intertwining narratives (chapter 2). To track the participation of incidents or events in more than one narrative simultaneously requires multiple codings.

5. *Domain of meaning to be coded.* The last issue involves determining what sort of meaning the classification scheme is intended to capture. Though finer distinctions are possible, Poole et al. (1987) distinguished observer-privileged meanings from subject-privileged meanings. Observer-privileged meanings are those accessible to outside observers, whereas subject-privileged meanings are those understandings that insiders and participants would have of the same incidents. Clearly a coding scheme designed to pick up subject-privileged meaning is harder to design than an observer-privileged system. In the section on validation below, we briefly discuss how to test whether a system can get at subject-privileged meanings.

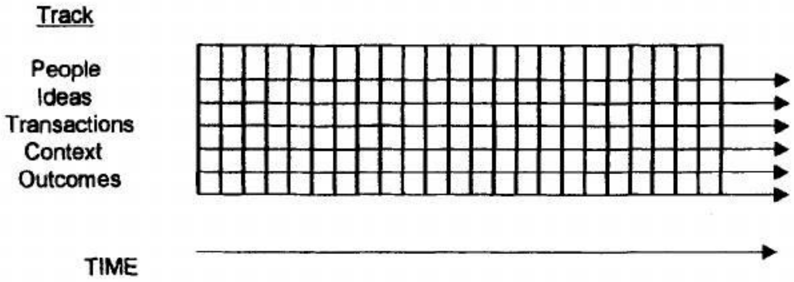
Layered Coding

The five issues just discussed cover many important issues that researchers confront in designing category systems. It is also the case that often codings are recoded into higher-order categories. The simplest examples of this has been described already: the translation (read coding) of raw data into incidents is then recoded into theoretically meaningful events. These events can then, in turn, be recoded into higher-order variables.

A limitation of many coding systems is that they reduce rich qualitative data to a single dimension of meaning. For example, a failure to get renewed funding for an innovation may influence the development of the idea behind the project; it may also result in layoffs of innovation personnel; and it may signal a change in the relationship of the innovation to external resource controllers. If we code this incident simply as the termination of an idea, we omit other dimensions of meaning of the event. To avoid this problem, incidents may be coded on several dimensions. For example, we might code the incident on four event dimensions: a negative outcome (resource cut), a change in the core innovation idea, people leaving, and a change in transactions (relations with resource controllers).

One way to organize these multidimensional data into a format to analyze change processes is to array them on multiple tracks that correspond to conceptually meaningful categories. The procedure of coding incidents along several event tracks evolved in Poole's (1983a & b) studies of decision development in small groups. Poole argued that previous models of group decision development—which commonly posited a rational decision process of three to five stages—were too simple. He was interested in testing the hypothesis that decisions did not follow a fixed sequence of phases, but instead could follow several different paths. He also believed that the previous practice of coding only one dimension of group behavior, such as task process, was responsible for previous findings supporting the single sequence models. To examine a richer model of group development, he developed a three-track coding system: one track coded the impact an incident had on the process by which the group does its work (in this case, an incident was a member's statement), a second coded the same incident in terms of its effect on group relationships, and a third track indexed which of several topics the incident referred to. By coding an incident on several conceptually relevant dimensions simultaneously, Poole was able to derive a richer description of group processes than had previous studies.

The CIP researchers, along with other MIRP investigators, also coded each innovation incident according to multiple dimensions or constructs of events. Their coding scheme captured key dimensions of changes in innovation ideas, people, transactions, context, and outcomes in an observed incident. Figure 5.1 summarizes the general MIRP categories and displays them as a set of layered tracks. Within each conceptual track a number of more specific codings are possible, depending on the particular questions being addressed by the researchers. For example, in their MIRP study Ring and Rands (1989) coded incidents in terms of more refined dimensions of transactions in order to examine their model of formal and informal transaction processes, while Garud and Van de Ven (1989) expanded the context track into a number of dimensions to examine their model of industry emergence. Thus, the coding scheme can be tailored to meet the needs and interests of individual MIRP study teams.



People Track: a coding of the people/groups involved in an incident, the roles and activities they perform at a given point in time.

Ideas Track: a coding of the substantive ideas or strategies that innovation group members use to describe the content of their innovation at a given point in time.

Transactions Track: a coding of the formal and informal relationships among innovation group members, other firms, and groups involved in the incident.

Context/Environmental Track: a coding of the exogenous events outside of the innovation unit in the larger organization and industry/community which are perceived by innovation group members to affect the innovation.

Outcomes Track: when incidents provide evidence of results, they are coded as representing either positive (good news or successful accomplishment), negative (bad news or instances of failure or mistakes), and mixed (neutral or ambiguous news, indicating elements of both success and failure).

Figure 5.1 Coding tracks on core MIRP dimensions

Assessing the Reliability and Validity of Coding Systems

A number of steps can be taken to enhance the reliability and validity of coding incidents into indicators of event constructs or events into higher-order constructs. First, operational definitions and coding conventions can be drafted for the coded constructs. Periodic meetings can be conducted with researchers and other colleagues to evaluate the construct validity of these definitions; that is, the extent to which operational definitions appeared to be logical and understandable indicators of the constructs under consideration.

Van de Ven and Ferry (1980) found that a useful way to conduct these meetings is to begin with an overall presentation of the conceptual model being studied. Then participants are handed a paper that defines each construct in the model and the sugggested indicators to be used to measure each construct. Participants can then be asked to "suggest better indicators for measuring this construct as defined previously." Often using a Nominal Group Technique format (see Delbecq, Gustafson, & Van de Ven, 1975), reviewers are provided a brief period to think and respond to the questions in writing. Then a general discussion ensues to obtain group opinions. The qualitative written comments from these review sessions are especially helpful to sharpen the norms of correspondence (Kaplan, 1964) between definitions of constructs and event indicators, and to clarify ambiguities in decision rules for coding event indicators.

In addition to incorporating these protections into the construction of categories, it is also important to assess reliability and validity of coding systems based on empirical performance. Reliability refers to the consistency of coding classifications across raters. Two types of reliability can be distinguished (Folger, Hewes, & Poole, 1984): Unitizing reliability refers to consistency in dividing the stream of activity into units. Classificatory reliability refers to consistency in assigning units to categories. The two sorts of reliability must be assessed separately. One widely used index of unitizing reliability is Guetzkow's U (Folger et al., 1984, pp. 119–120). Folger et al. describe a more sensitive method of assessing unitizing reliability based on comparison of lengths of units, which Guetzkow's U takes into account only indirectly. The best measure of classification reliability is Cohen's kappa, available in most nonparametric statistics programs (Popping, 1988). In addition to reliability across the entire category system, it is also important to compute reliability for each individual category. This helps identify problem categories.

While consistency of classification (reliability) is an important criterion, equally important is accuracy of interpretation (validity). It is not common to assess the validity of coding systems, but it is just as important to do so as it would be to assess the validity of a scale. What evidence is necessary to establish the validity of a coding system depends on the type of meaning it is designed to capture. If the system is designed to capture only observer-privileged meanings then the same types of evidence used to assess face and construct validity for quantitative measures can be obtained to evaluate the validity of the category system. To the extent that the constructs in the coding system relate to other constructs in ways that would be expected based on the theory of the construct, then the coding system has a measure of validity (see Folger et al., 1984).

Things are more complicated when the researcher aspires to capture subject-privileged meanings with a classification system. In such cases, face and construct validity must be established, but they are not sufficient to ensure validity of the system as a whole. In addition, the researcher must show that the interpretations yielded by the coding system correspond somehow to those of the participants or insiders. This involves mapping participant or insider intepretations and comparing them to the interpretations of the system. Folger et al. (1984) discuss several ways of doing this.

Chronicles: Event Sequence Data Files

Moving from raw data to incidents results in a stream of data consisting of a chronological listing of all incidents observed in the development of an organizational change process. Each incident represents a datum that is entered as a unique record into a qualitative data file for each innovation. Table 5.5 shows an example of a few incidents in such a data file. A variety of database software programs can be used to organize and manage the qualitative data files. MIRP researchers used R:Base©, but Access™ or any other relational database program can be used equally well. Weitzman and Miles (1995) provide a compilation of qualitative data analysis programs that could also be used for recording such data.

When events and higher-order constructs representing types of events are identified from the sequence of incidents, they can be recorded in the database in several ways. The appendix to this chapter describes the steps involved in creating a qualitative chronicle similar to that used in the MIRP research. This example utilizes the MIRP categories discussed above and presented in detail in the appendix to chapter 4.

While these examples from MIRP are useful for illustrative purposes, it should be noted that many other types of developmental processes can be represented, such as decision-making discussions and other deliberations (in this case the actual verbiage and the data are almost identical, save what is left out by transcription rules), critical events in organizational careers, occurrences during the implementation of information technologies, and the history of the addition and adaptation of features of a organizational structure in a firm.

The chronicle, or event sequence, presents the basic data to be analyzed with the methods outlined in the rest of this book. To render the data suitable for analysis, some transformations are required.

Table 5.5 Partial Bitmap of Incidents in CIP Study

INCIDENT NUM	DATE	DAYS	IC	IR	PE	TR	CI	CE	OP	ON	OM
1	01/01/68	1	1	0	0	0	1	0	0	0	0
2	01/01/68	1	0	0	0	1	0	0	0	0	0
3	01/01/68	1	0	0	0	0	0	1	0	0	0
4	01/01/68	1	0	0	0	0	0	1	0	0	0
5	01/01/68	1	1	0	0	1	0	0	0	0	0
6	01/01/68	1	1	0	0	1	0	0	0	1	0
7	01/01/74	2193	1	0	1	0	0	0	0	0	0
8	01/01/74	2193	0	0	1	0	0	0	0	0	0
9	12/01/78	3988	0	0	1	0	0	0	0	0	0
10	10/01/79	4292	1	0	1	1	0	0	0	0	0
11	10/07/79	4298	1	0	0	1	0	0	0	0	0
12	01/01/80	4384	0	0	0	0	0	1	0	0	0
13	01/04/80	4387	1	0	0	0	0	0	0	0	0
14	01/08/80	4391	0	0	0	1	0	0	0	1	0
15	01/12/80	4395	1	0	0	0	0	0	1	0	0
16	04/01/80	4475	1	0	0	0	0	0	0	0	0
17	04/07/80	4481	0	0	0	1	0	0	1	0	0
18	05/01/80	4505	0	0	0	1	0	0	1	0	0
19	10/01/80	4658	1	0	0	0	0	0	0	1	0
20	10/07/90	4664	0	0	0	1	0	0	1	0	0
21	11/01/80	4689	0	0	1	1	0	0	0	1	0
22	11/07/80	4695	1	0	0	0	0	0	0	0	0
23	11/15/80	4703	0	0	0	1	1	0	0	1	0
24	11/21/80	4709	0	0	0	1	0	0	0	1	0
25	12/01/80	4719	1	0	1	1	0	0	0	0	0
26	12/11/80	4729	0	0	0	1	0	0	0	1	0

Variables (number in [] represents the frequency of 1s):

num: Incident Number
date: Incident Date
days: Number of Days from 01/01/68 (thr first incident)
ic: Idea-core [68]
ir: Idea-related [4]
pe: People [49]
tr: Transaction [165]
ci: Context-internal [8]
ce: Context-external [22]
op: Outcome positive [58]
on: Outcome-negative [59]
om: Outcome-mixed [7]

TRANSFORMING CODED DATA INTO FORMS SUITABLE FOR ANALYSIS

We will go into greater detail on transformations appropriate for each of the data analysis methods in subsequent chapters. At this point, we will discuss in general terms several transformations of the qualitative data that have proven useful.

Summary Data

One transformation involves converting the sequence data into summary statistics, such as the total number of events in various categories in the entire sequence or in segments of it; or the total number of phases in the process (see below for definition of phase). This data can then be used to test developmental models with variance analysis. For example, if a three-phase life cycle model holds, then the monthly count of events characteristic of phase 1 should be highest early in the process and decline, the monthly count of events characteristic of phase 2 should peak after phase 1 events start to decline and then decline thereafter, and the count of events characteristic of phase 3 should peak near the end of the process, as shown in Figure 5.2. This transformation, one of the most common in developmental research, essentially collapses the data over time, removing any temporal information. The remaining three transformations preserve temporal information.

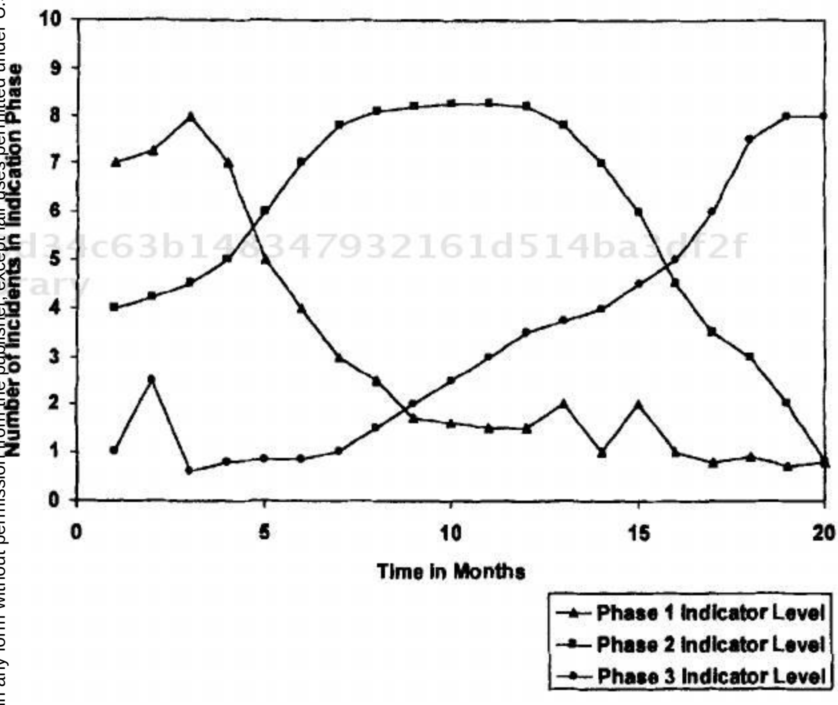


Figure 5.2 Patterns of indicators expected if a three-phase life cycle holds for a process

Bitmaps

Bitmaps (Frey, 1986) can be used to represent nominal level constructs, such as the occurrence or nonoccurrence of a certain type of event. One such bitmap is shown in Table 5.5. In a bitmap a column in the database is created for each event category. If a given type of event, such as an idea event, can be classified into events affecting the core idea or events affecting related ideas, two different columns are created (columns ic and ir in Table 5.5). When an event occurs that affects the core idea, a "1" is entered into column ic; otherwise, the value for a given incident is coded as "0." Any data which can be exhaustively described by a set of nominal categories can be represented as a bitmap. This bitmap can then be subjected to analysis with various statistical methods to examine time-dependent patterns of relations among the event constructs. In a bitmap such as the one shown in Table 5.5, the chronological listing of qualitative events is time-dependent, meaning that the sequential order of the rows is crucial and should be taken into account when information is to be extracted, although the columns are interchangeable. A method that returns the same results when the rows of an event sequence bitmap are interchanged is not appropriate for identifying dynamic patterns because the information contained in the temporal order of the incidents is not used.

Phase Maps

Phasic analysis presumes that individual events are indicators of underlying phases. It transforms the sequential event stream into a series of phases of various lengths, using procedures to be described in more detail in chapter 7. Phase mapping adopts the third strategy for event identification from indicators: it uses a summary measure to define events or higher-order constructs. For example, one algorithm defines a phase based on three consecutive occurrences of an indicator. Fewer than three is considered an error and does not define a phase. Phase maps assume and try to correct for some unreliability in the indicators of events or higher-order constructs.

Continuous Variables

Some methods require continuous data. In these cases the sequence data, bitmaps, or phase maps may be transformed into continuous data by (a) dividing the sequence into segments of uniform length containing more than one event and (b) calculating the number of events of various types in the segments; the number of events represents a continuous measure. For example, Van de Ven and Polley (1992) coded whether innovation units persisted in their current actions or changed. They subdivided their event sequence into one-month intervals and counted how many of each category of behavior (persist or change) occurred in each month. This served as continuous data for a time series analysis of action persistence.

Comments

These transformations generate data that can be used in a wide variety of analyses, ranging from qualitative interpretation of the data stream to multivariate statistical analysis. Specific applications require quite a bit of tailoring of the original event data, as subsequent chapters illustrate. The important thing is to be ever cognizant of the possibilities for testing models and hypotheses and flexible with respect to methods of analysis.

Conclusion

This chapter introduced methods for designing longitudinal research for studying how and why change unfolds in organizational entities. The chapter outlined a host of choices available to researchers, all of which entail difficult decisions and trade-offs. As always, research design requires the exercise of what Aristotle called "practical wisdom." There is no definitive best design for a given project, and any design requires giving up some data in order to get others.

One objection that might be registered is that the methods proposed here may "overquantify" analysis. This conclusion may be the inadvertent result of our objective in this chapter, which was to introduce some systematic methods to overcome the tendency in much research of relying exclusively on subjective "eyeballing" and anecdotal information in qualitative data. However, in practice, our objective is to combine the special information that quantitative and qualitative approaches provide to understand organizational change processes. After all, by themselves quantitative data provide only a skeletal configuration of structural regularities, often devoid of life, flesh, and soul. Qualitative data, by themselves, are like an amoeba, rich with life but squishy, soft, and absent of apparent structure. Only by combining quantitative and qualitative data in a balanced way do we come to understand the richness of life in its varied regularities.