© Academy of Management Review
2001, Vol. 26, No. 4, 530–547.

# BUILDING BETTER THEORY: TIME AND THE SPECIFICATION OF WHEN THINGS HAPPEN

TERENCE R. MITCHELL
University of Washington

LAWRENCE R. JAMES
University of Tennessee

**In any investigation of a causal relationship between an X and a Y, the time when X and Y are measured is crucial for determining whether X causes Y, as well as the true strength of that relationship. Using past research and a review of current research, we develop a set of X,Y configurations that describe the main ways that causal relationships are represented in theory and tested in research. We discuss the theoretical, methodological, and analytical issues pertaining to *when* we measure X and Y and discuss the implications of this analysis for constructing better organizational theories.**

Our purpose in this article is to discuss a very specific issue concerning time and organizational theory. It is our contention that when a hypothesis involving causal relationships between X and Y is proposed and tested, a key issue for correctly testing the hypothesis and for making correct inferences based on the empirical results of the test is knowing *when* X and Y occur. Theory, method, analysis, and inference must be appropriate. Any inference involving the existence of an X,Y relationship or its strength is dependent upon when X and Y are believed to occur and when they are measured. We believe that the neglect of this issue poses serious problems for the evolution and acceptance of our field.

## BACKGROUND

We should initially point out that there are a number of journal reviews, books, and edited volumes on the topic of time. Books by McGrath and Kelly (1986), McGrath (1988), and Kelly and McGrath (1988) cover social psychological perspectives, and the recent book by Pentland, Harvey, Lawton, and McColl (1999) focuses on how people distribute activities across time. A review by Bluedorn and Denhardt (1988) broadly covers the topic in the organizational research area, and Clark (1985) discusses time in the field of sociology. Other papers focus on narrower reviews. For example, George and Jones (2000) and Zaheer, Albert, and Zaheer (1999) provide excellent analyses of the complex ways time can and should be incorporated in the theory-building process. Okhuysen (1999) looks at time in decision making, Gersick (1988) examines time in groups, and Ancona and Chong (1996) provide an overview of recurring cycles and rhythms in organizations. Chan (1998) has an excellent discussion of how we can analyze relationships involving change over time. We draw on these sources in various parts of this article.

Our interests, however, are very precise and are basically not covered in detail anywhere in the organizational literature. Our domain of interest is both the theoretical statement and empirical testing of causal relationships. Our specific focus is on *when* the variables involved in the relationship occur. At the simplest level, in examining whether an X causes a Y, we need to know when X occurs and when Y occurs. More complex relationships may involve repeated incidents of X or Y or predicted changes in either X or Y, or both. Without theoretical or empirical guides about when to measure X and Y, we run the risk of inappropriate measurement, analysis, and, ultimately, inferences about the strength, order, and direction of causal relationships. Thus, the integration of theory, design, and analysis is critical.

We thank the three anonymous reviewers and special issue editors who helped make the paper better as a result of the revision process. Terry Mitchell especially thanks Joe McGrath for lessons that have lasted a lifetime.

Our treatment of time is embedded in a fairly traditional view of how time is represented in science, an orientation that Gurvitch (1964) and Clark (1985) describe as standard time or clock time. Time is treated as a commodity that can be broken into meaningful segments or blocks. It flows evenly and continuously. It is precise and quantifiable. It has the properties of an ordinal scale. We recognize that there are other philosophical positions (see Faulconer & Williams, 1985, and Gergen, 1973) in which time is viewed more hermeneutically. Time may be experienced psychologically and physically in very different ways (Berger & Luckmann, 1966; Mosakowsi & Earley, 2000). However, the traditional position we examine is by far the most frequently represented in the social sciences (Clark, 1985): "Standard or clock time has become the dominant orientation toward time in the organizational literature" (George & Jones, 2000: 659).

We focus only on research involving the theoretical statements concerning causal relationships that are empirically tested in a manner that allows causal inferences. Thus, nonempirical research is omitted from our analysis. We also omit descriptive and qualitative studies (with or without empirical data) in which no causal relationships are suggested or tested and studies in which associational (noncausal) hypotheses are made and cross-sectional data are gathered and analyzed.

Is the category of research we are examining important? Is it represented in the literature? A decade ago Sackett and Larson (1990) reported that they reviewed every article published in the *Journal of Applied Psychology*, *Organizational Behavior and Human Decision Processes*, and *Personnel Psychology* from the years 1977, 1982, and 1987. They reported that in 286 of 577 studies (about half), researchers used designs and analyses with which they tested a causal hypothesis with empirical data. To provide a current update directly relevant to the field of management, we examined all the studies published in the *Academy of Management Journal* (*AMJ*) and *Administrative Science Quarterly* (*ASQ*) in 1999. We discuss more details of this analysis using the *AMJ* articles (simply for convenience) throughout the paper, but an overview is helpful at this point.

We describe the numbers for *AMJ* and place the *ASQ* numbers in parentheses. *AMJ* published forty-three (twenty-three) studies in 1999. Of these, one (one) has no data. There are two (zero) meta-analyses and two (three) descriptive papers. Six (five) do not involve the testing of causal hypotheses. Thus, thirty-two (fourteen) of forty-three (twenty-three) papers have theoretical statements or hypotheses involving causal hypotheses (about 60 to 70 percent). We should add that some papers include some causal hypotheses, along with associational and descriptive statements, but as long as some causal relationships are part of the study, we included them in these summary numbers.

We discuss later, in more detail, the methodological and conceptual factors that are necessary for testing causal hypotheses and making causal inferences. However, it is important to point out that nine (one) of the thirty-two (fourteen) papers suggesting causal hypotheses have survey designs in which it is impossible to infer causality. All the variables are measured at the same time and, while statistical analytical procedures are available that may be helpful for suggesting causality, to be conservative, we did not include these papers in our category of having causal hypotheses and designs allowing causal inferences.

In summary, twenty-three of forty-three *AMJ* studies and thirteen of twenty-three *ASQ* studies fall in the causal theory, causal design category. We add that there are a few studies in both *AMJ* and *ASQ* in which there is ambiguity about whether the hypotheses or designs address causal inferences. The important point, however, is that our data are similar to the Sackett and Larson (1990) results in that about half the published work we examined reflects this type of scientific activity.

Of course, a second and equally important question is whether the theory or designs used in our research specifically address the issue of *when* the variables involved in the relationships actually occur. That is, does the theory and/or the design focus on when X or Y should be or was measured? Kelly and McGrath (1988) reviewed all of the articles of a leading journal (unspecified) in one year, and although they were looking at a wider variety of issues than we are, they concluded that, frequently, "The duration of the interval between cause and effect is left unspecified in our theoretical formulations and in our interpretation of concrete findings" (1988: 19).

Our summary of the 1999 *AMJ* papers leads us to a similar conclusion. In only four of the twenty-three *AMJ* causal theory, causal design studies is a theoretical rationale specifically mentioned with respect to the issue of the timing of the variables in theory or in the design. Bloom (1999) mentions that the salary distributions for a specific year should influence baseball players' performance in the immediate subsequent year as contrasted with other years. Judiesch and Lyness (1999) look at the effects of leaves of absence on career success and measure the latter after a few years in order to provide "ample time" for the effect to take place. Waller (1999), studying pilot crews, suggests that task prioritization in response to a nonroutine event only influences performance positively if the prioritization occurs soon after the nonroutine event. Waller's (1999) design is sensitive enough (variables assessed in a continuous fashion) that the timing of the variable could be assessed and analyzed. Finally, Welbourne and Cyr (1999), who researched the effect on performance of having a highly placed human relations executive when a company went public, specifically mention that they waited three years because other research indicated this was the appropriate time lag.

Thus, most of our research involves causal hypotheses and designs presumed to support causal inferences. Yet, very few papers specifically address, from a theoretical perspective, the time elements involved in X causing Y. Our research designs also infrequently allow us to make precise inferences about the time involved for X-causes-Y relationships. It is our belief that the field of organizational studies can and should do better: failure to consider the when in our theories and methods places the entire enterprise of causal inference on a tenuous footing.

## TIME AND THEORY

Before proceeding to a detailed analysis of the role of time in making causal X,Y inferences, we need to note that time is important in organization theories in other ways. Some authors focus on individual differences or cultural differences in the perception of time (Bluedorn, Kaufman, & Lane, 1992; Mosakowski & Earley, 2000). Gersick (1988) uses time as a marker of group process transitions. Time cycles in one system (e.g., the academic year) can influence other individual or group cycles in a process labeled *entrainment* (Ancona & Chong, 1996), and these cycles can occur at different organizational levels (Goodman, 2000).

However, without reservation we can say that most theory involves fairly simple relationships of the X-causes-Y variety, with X and Y representing substantive variables other than time. Within this set of relationships there are five major ways in which theory informs method with respect to time. First, we need to know the time *lag* between X and Y. How long after X occurs does Y occur? Second, X and Y have *durations*. Not all variables occur instantaneously. Third, X and Y may change over time. We need to know the *rate* of change. Fourth, in some cases we have *dynamic relationships* in which X and Y both change. The rate of change for both variables should be known, as well as how the X,Y relationship changes. Finally, in some cases we have *reciprocal causation*: X causes Y and Y causes X. This situation requires an understanding of two sets of lags, durations, and possibly rates.

We should point out that time can enter into our theories in much more complex ways than those described above. Nonlinear relationships over time are possible, as are cyclical and oscillating ones. Change can be incremental or discontinuous. Cycles can spiral up or down, and the intensity can change. Various relationships can have rhythms or patterns over time. We do not delve into these more complex relationships for one major reason: they are, as yet, not well represented in the published literature. For example, none of the papers published in *AMJ* during 1999 have theoretical propositions or methodological procedures that test for such relationships. However, Kelly and McGrath (1988) make it clear that such relationships are possible and important, and Ancona and Chong (1996) and George and Jones (2000) conclude that our theory is impoverished without an examination of such relationships.

Unfortunately, in most of our research we do little more than say that one event will be followed by another or that more time results in different behaviors or activities than less time or that X may influence Y, which will, in turn, influence X or a third variable, Z:

> Although theories in organizational behavior, more often than not, specify relationships among constructs in causal terms, the duration of effects, the time lag between causes and effects, and differences in rates of change are often left unspecified (George & Jones, 2000: 670).

With impoverished theory about issues such as when events occur, when they change, or how quickly they change, the empirical researcher is in a quandary. Decisions about when to measure and how frequently to measure critical variables are left to intuition, chance, convenience, or tradition. None of these are particularly reliable guides and, as we shall see, investigations using these guides are open to criticism and prone to error.

It appears that roughly 40 to 50 percent of our published research involves theory suggesting causal X,Y relationships. Table 1 presents a summary of different ways X and Y are represented in these theoretical relationships. We cover each situation briefly and provide theoretical examples for each configuration. Although these configurations are simplified and not exhaustive, they are meant to cover the major ways in which time is involved in our theoretical statements suggesting causality. Throughout this discussion we use examples from our review of volume 42 (1999) of *AMJ*. We should initially note that because of our background, many of the studies we cite are micro examples. Macro examples are readily found in Zaheer et al. (1999), George and Jones (2000), Goodman (2000), and Mosakowski and Earley (2000).

*Configuration 1:* X causes Y. This is the simplest, most frequently appearing relationship stated in our theories. The main time implication is that X must precede Y. In addition, theorists should, but infrequently do, state exactly when Y will occur—specify the time lag. In most cases this is simply presumed to be immediate, and Y is not characterized or even discussed as changing over time. For example, Westphal (1999) demonstrates that interactions between CEOs and their boards of directors positively influenced firm performance two years later.

*Configuration 2:* X causes Y, and the relationship is stable over time. In this relationship X is introduced (or naturally appears) in some regular, continual fashion, and the effect on Y is hypothesized to be stable. The Xs precede the Ys. Every time X occurs, Y will occur, and it will do so consistently. A good example is provided by Bloom (1999), who demonstrates that over nine seasons (1985–1993) the pay dispersion (according to pay contracts at the beginning of that season) on major league baseball teams was negatively related to individual and team performance for that year.

*Configuration 3:* X causes Y, and Y changes over time. X precedes Y, and the time lag is important between X and all three occasions when Y occurs. But also important is the type of change in Y between time 2 and time 3 and between time 3 and time 4. In some cases this may be a very predictable mean change. In other cases we may know the direction of a change but have less precision in specifically stating the rate of change. Callister, Kramer, and Turban (1999) provide an example of feedback seeking after a job transfer. The dependent variable was measured right after the transfer, after three months, and after one year. Inquiring

**TABLE 1**
**Theoretical Relationship and Time**

| Configuration | Time 1 | Time 2 | Time 3 | Time 4 |
|---|---|---|---|---|
| 1 | X →→→→→→→ | Y | | |
| 2 | X →→→→→→→ | Y | X →→→→→→→ | Y |
| 3 | X →→→→→→→ | Y →→→→→→→ | YΔ →→→→→→→ | YΔ |
| 4 | X →→→→→→→ | Y | X →→→→→→→ | YΔ |
| 5 | X →→→→→→→ | Y | XΔ→→→→→→→ | YΔ |
| 6 | X →→→→→→→ | Y →→→→→→→ | XΔ→→→→→→→ | YΔ |
| 7 | X →→→→→→→ | Y →→→→→→→ | Z →→→→→→→ | Q |
| | | Y<br>↓ | | |
| 8 | X →→→→→→→ | →→→→→→→ | Z | |

Note: An arrow (→) implies causality, and delta (Δ) implies a change in X or Y.

for feedback from peers and supervisors decreased over time.

*Configuration 4:* X causes Y, but over repeated exposure to X, Y changes. Time lags again are important here but in two ways. First, we need to know the X,Y and the X,YΔ lags, but we also need to know at what point in time the exposure to X produces a change in Y. A classic example is the concept of satiation from learning theory, where the effect of some reinforcer wears off at a specific rate over time. A not so obvious example comes from ingratiation theory (Liden & Mitchell, 1988). The repeated use of some tactics (e.g., flattery by person A) has a predictable diminishing effect on the liking of person B for person A. The same issue is true for other ingratiation tactics, although some may increase in effectiveness while others may decrease. One of the key problems with ingratiation approaches has been the inability to say whether the changes in Y occur simply as a result of the frequency of use of X or whether there are changes in the effect of X as a result of other things that change in a relationship over time—or both. Note that the type of change can be a relatively simple question of rate change or a more complex change involving changes in Y due to X but also change due to other variables that are associated with the passage of time that are not X. We return to this issue in our methods section.

*Configuration 5:* X causes Y, and then a changed X causes a changed Y. In some cases theorists suggest that systematic changes in X over time will lead to systematic changes in Y over time (Mone, 1994). Thus, different levels of X appear or are introduced in a systematic way over time, and Y changes as a result. For example, Carlson and Zmud (1999), who looked at the relationship between experience with e-mail and perceptions of the richness of that communication channel over a thirteen-week period, found that while experience (X) increased (manipulated through specific tasks), the richness perceptions (Y), which were measured after the experience, initially increased and then decreased.

Note that what is happening here is that, over time, there is systematic variance in X. Individual learning and experience are good examples of where the passage of time can produce different levels of these variables. Here we have added to the need to know lags and types of change the need to know the time involved that results in different levels of X. In terms of our diagram, we need to know the time between $X_1$ and $Y_2$, $X_3$ and $Y_4$, $Y_2$ and $X_3$, and $X_1$ and $X_3$. Many evolutionary or developmental theoretical approaches include such relationships.

*Configuration 6:* X causes Y, which causes a changed X, which causes a changed Y. This pattern introduces cyclical recursive causation. Besides the obvious $X,Y_2$ lag, the $Y_2,X_3$ lag is critical. It may be different from the $X,Y_2$ lag. For example, self-efficacy performance spirals, as described by Lindsley, Brass, and Thomas (1995), indicate that people who have moderately high self-efficacy will perform well and that good performance will result in higher self-efficacy, which will result in higher performance and so on. People can spiral up or down.

*Configuration 7:* X causes Y, which causes a different variable Z, which may, in turn, cause a different variable Q. This type of relationship presents itself in two different ways. First and most obvious is the situation suggested by theories involving mediation. The relationship between X and Z is mediated by Y. The time of occurrence and the lags between all the variables in the chain may be important. Thus X,Y; Y,Z; and Z,Q relationships all play roles in supporting theoretical propositions. Amabile and Conti (1999), for example, demonstrate that downsizing (X) presents obstacles at work (Y), which eventually reduce creativity (Z).

This type of configuration also may be used to describe relationships that emerge from theories of interaction among dyads. For example, Andersson and Pearson (1999) recently introduced the idea of incivility spirals, where a benign behavior from person A may be perceived as uncivil, leading to incivility on the part of person B that leads to incivility on the part of A, which may lead to aggression from B. Again, the lags are important, as is what else happens with the passage of time. But note that since X and Y may be different variables from Z and Q, the way that time influences these relationships also may be different.

*Configuration 8:* X causes Z, but the strength of the relationship varies as a function of the level of Y. This is the classic example of a moderated relationship. The X,Z issues pertaining to time are the same as for configuration 1. The start, stop, and lag times are important. But also important and almost never discussed are the X,Y time lags or the Y,Z lags. In some cases X and Y may occur simultaneously, or Y may follow X in time, or Y may also represent a variable that differs across individuals or groups but is presumed to be relatively stable across time.

Gibson (1999) presents an example of both types of moderation. In her laboratory study she assessed the degree to which group efficacy (measured before the task was completed) interacted with task uncertainty (manipulated) to affect group effectiveness. She also measured group collectivism as a contextual moderator—presumably, a variable (Y) that was relatively constant and present at the time X was assessed. The efficacy/collectivism combination did not affect group effectiveness.

Configuration 8 is the last one we present. Obviously, more complex configurations can occur in any given theory or any given paper. We can have longer chains combining configurations. There is mediated moderation (James & Brett, 1985) and theories of entrainment where, for example, certain X,Y relationships become associated with other X,Y relationships in groups over time (Ancona & Chong, 1996; McGrath & Kelly, 1986). Theories exist in which curvilinear or other configural relationships are predicted (George & Jones, 2000), and there are theories including rhythms, cycles, and pacing and descriptions of how these processes can occur at different conceptual levels (Goodman, 2000). Thus, much more complexity exists than we have described—our examples are not exhaustive.

However, we should point out that the above configurations capture a substantial portion of what appears in our literature. For example, in the 1999 *AMJ*, the twenty-three empirical papers that tested for causal relationships all have hypotheses that fit in the above configurations.[1] About one-half are represented by configuration 1, but as the examples illustrate, most of the other configurations are also represented. There are no examples of configurations 4 or 6.

A review of these studies overall shows that time can be involved in theoretical statements in many different ways. In turn, the appropriate timing of measurement influences whether we can unambiguously conclude that a variable is a cause, an effect, a mediator, and a moderator. It is our task as theorists to be precise about these questions if we are to both test theories and be better able to reject or accept their propositions.

## TIME AND METHODS

Methods (research designs and analytical tools) are used to test theory. In many cases, especially in the field, when we measure or how frequently we measure is only partially under our control; the realities of organizational constraints and data collection dictate our actions. But, in other cases, our methods are selected because they are easily used, familiar, or popular. One purpose of this section is to discuss the ways that theory drives method and, more specifically, how the specific and clear theoretical treatment of time and appropriate research design can help confirm or disconfirm our theoretical propositions. Including time in our theories raises issues that need to be addressed in our methods and measures. In addition, it is important to point out that method informs theory. By using the proper designs and analyses, we can help to confirm, revise, or disconfirm theoretical propositions involving time.

Remember that our focus is on theoretical propositions involving causal relationships. In order to demonstrate causality, we usually employ three standards to evaluate our research designs (Popper, 1959). First, X must precede Y in time. In the following section, using the 1999 *AMJ* examples, we discuss various ways in which this design requirement is met. In many cases the time ordering of X and Y is obvious, and we shall see that such ordering is not necessarily dependent upon when X and Y are measured. In addition, just because X precedes Y (and therefore meets one causal standard), this does not mean that Y cannot, in turn, cause X. Second, variance in X must be associated with variance in Y. Traditionally, simple significance tests are used to make this judgment. However, of even greater importance is how much vari-

[1] Providing an exact breakdown for the twenty-three studies is not easy and would require more space than is warranted. Some papers have multiple hypotheses, most (but not all) of which are of the same configuration, and we classified them in terms of the dominating configuration. Also, some papers have multiple studies (e.g., lab and field). In those cases, if one study had causal hypotheses and non-cross-sectional data, we classified it as one of the twenty-three and assigned the configuration to one of the eight presented.

ance in Y is accounted for by variance in X. That is, the strength of this relationship is important, and good research designs are better able to accurately assess this strength. Finally, we have to be sure that there is not a third variable that accounts for the X,Y relationship. Both design and statistical controls are frequently used to help us evaluate this standard.

### Research Designs

> Most study designs in social and behavioral science research imply that the delivered increment of X will yield a fixed increment of Y—and that variations in the duration of the X-Y interval will neither add to nor diminish that increment. According to this view, Y exists at a certain level. The Treatment, X, is applied, Y moves to a higher level, and stays there (Kelly & McGrath, 1988: 79).

This representation is indeed what emerges from our analysis of the 1999 *AMJ* papers. In thirteen of the twenty-three papers that we analyzed, a simple X,Y causal relationship is proposed. Four others include a moderator, and one has a mediator. But the types of designs used to test these propositions vary dramatically, and since they represent the designs used in the studies suggesting other theoretical configurations, we mention them here. We also add that this article is not meant to be a tutorial on research design. Many competent sources are available on this topic (e.g., Cook & Campbell, 1979). We focus on the issues related to the timing of measurement, since this is an issue about which Campbell and Stanley (1963) and Cook and Campbell (1979) are essentially mute.

The most frequent design used to test X,Y relationships in the 1999 volume of *AMJ* is a simple XO design—used in eight studies. These researchers employed surveys and/or archival data, and it was fairly clear that X preceded Y (e.g., the publication success of academic job candidates as indicated on their vitae was associated with their subsequent salary when they were hired; Cable & Murray, 1999). There are also a number of cross-sectional time-series designs of the XOXOXO variety. Combs and Ketchen (1999), for example, using archival data, demonstrated that the scarcity of capital predicts subsequent franchising activity over time. A third design that appears twice is a longitudinal field study (e.g., an $O_1XO_2O_3O_4$ design). Amabile and Conti (1999) looked at creativity before and after downsizing. Finally, in two studies posttest-only control group designs with random assignment are used. Jung and Aviolo (1999), for example, randomly assigned leaders with two different leadership styles to work in groups and observed performance.

Obviously, this last example is one of Campbell and Stanley's (1963) "good" designs. With random assignment and a control group, confounds associated with time, such as history, maturation, mortality, testing (reactivity), and instrument change, are usually controlled. But these are not confounds uniquely associated with the particular time and frequency of measurement. The other designs, without comparison or control groups, do pose problems for our analysis. A number of issues arise. First, *when* X and Y are measured does not determine their causal order. In many cases X and Y come from archival records, and although they are measured at the same time, it is clear that X precedes Y. However, in some cases people are required to recall or recollect a previous X, and in these cases the time reflected in those reports may vary substantially in validity and across subjects.

Whether and by how long X precedes Y may also be vague. Part of the problem is that archival data often include a window of time to measure X and Y. For example, if one studies the acquisition of firms over a three-year period (e.g., 1992–1995) and the performance of these firms in the year 2000, the X,Y time lag may differ for members of this sample. A third problem is that by not manipulating X, some of the variance in Y accounted for by X may really be accounted for by an earlier Y. This situation suggests that configuration 6 (with reverse causality) may be present, which requires attention to the Y,X measurement lag as well as the X,Y lag.

Finally, although in most of these studies the researchers use statistical procedures to control for possible third variable explanations, none of them address how a third variable might interact with the X,Y time lag such that the strength of the relationship is incorrectly inferred. Thus, these XO designs, which are frequently used because they allow the researcher to capture some "real-world" phenomenon, not only have the problems associated with Campbell and Stanley's (1963) list of confounds but also have problems associated with when X and Y are measured.

### The Timing of Measurement

"Time orders and intervals play a central role in virtually all of the major research strategies (laboratory experiments, field studies, surveys, and the like), and a known time relation between variables of interest is essential to the interpretive logic of all of our study designs" (McGrath, 1988: 265). Yet, as Kenny informs us, "Normally the lag between measurement is chosen because of convenience, not theory, since theory rarely specifies the exact length of the causal lag" (1975: 894).

One question posed by a causal hypothesis is determining the initial measurement time of X. In laboratory studies this may be obvious, but in field studies it constitutes the "start" time (Ettlie, 1977). If relationships are cyclical (e.g., seasonal) or if some sort of precondition "base rate" is needed, then *when* one starts to measure X is important. Also, if one is predicting a pattern of events that occur in some order, a theory must inform the researcher about start time.

Another "start time" issue is related to the stability or steady state of X. Some variables may have predictable times when they change before they reach a steady state. For example, newcomers on the job change their feedback-seeking behavior and their relationship with their boss substantially over their first six months on the job. A steady state tends to occur after that, and any research testing a general hypothesis treating any one of these variables as an independent variable (X) would need to be concerned with when X was measured. In addition, some variables, such as effort, may have predictable cycles over the course of a day or a week such that when one measures it is important. What we are suggesting is that X may take some amount of time to unfold, and the shape of that unfolding (e.g., linear, cyclical) may vary. Thus, getting the start time correct is important. X is not always discrete and instantaneous, nor is Y.

But probably the biggest problem is the issue of when Y should be measured. Does Y appear at a specified time after X? This timing is critical in a number of ways. First, by measuring too soon or too late, one does not provide a good test of the theory and is left with two explanations as to why there was no support: bad theory or bad design. Second, where reciprocal relationships are involved, one needs to be precise about the $X_1,Y_2$ time lag and the $Y_2,X_3$ time lag (configuration 6). Similarly, for mediation and moderation, the $X_1,Y_2$ and $Y_2,X_3$ (configuration 7) or $Y_2,Z_3$ (configuration 8) time lags are important. *When* one measures Y and the subsequent measurement of X are important. Schmitz and Skinner (1993), for example, suggest that efficacy, performance lags may be different from performance, efficacy lags.

The real issue here is that the lag represents a window of time when various things can happen. It obviously includes the X,Y causal process, but it also includes the operations (on X,Y) of the surrounding environment, as well as a window for the entry of confounds. For example, even with a control group and the prediction of a stable Y, by increasing the lag, we increase the chances of confounds (e.g., like measurement decay), which can influence the strength of the relationship. Thus, if a lag is too big, X wears off or other variables may come into play. If it is too small, the effect may not be complete or reactivity may occur (responses to a "treatment" independent of its content). Zaheer et al. (1999) provide a good discussion of this issue and related issues, which they call "time scales."

Both theory and past research can be useful in helping the researcher make predictions and gather data at times that should be helpful. However, as we mentioned earlier, in very few studies is the issue of lag actually addressed (only four from the 1999 *AMJ* papers). As Chan says, "The bad news is that we almost never have a good approximation of what the true causal interval is" (1998: 476). Zaheer et al. (1999) make a similar point.

### The Frequency of Measurement

Related to the question of *when* to measure is the important question of how often to measure. In our simplest case, where there is some ambiguity about when Y occurs, multiple measures can help to determine the lag. Notice that the investigator has to select the number of occasions, as well as the lags between them.

This question becomes even more important when specific changes are predicted in either X or Y (configurations 3, 4, 5, and 6). Y may be predicted to change over time in some systematic but not particularly complex way. For example, certain things, like the accuracy of a performance appraisal (Heneman & Wexley, 1983) or the effects of inequity (Cosier & Dalton, 1983), wear off over time. The obvious implication is, at the minimum, to measure Y right after X occurs and at other specific points in time.

Of perhaps greater importance is the situation in which some specific rate of change is predicted. To assess rate of change, we need multiple assessments, and many time-series designs are available that describe multiple measurement procedures (Cook & Campbell, 1979). However, knowing *when* and *how frequently* to measure requires the theoretician to address issues of intraindividual change, interindividual change, and contextual change (Willett & Sayer, 1994). For example, if we were testing a theory about the effects of socialization on adjustment over time, we might want to know how individual attributes (like feedback seeking) were important (and when) over time. When, for example, is there no longer an effect?

What is advocated in most cases, since theorists are also often mute on this issue, is to measure as frequently as possible, without causing other problems, such as reactivity or subject discontent (Cook & Campbell, 1979). For various analysis issues, more people and more occasions are usually better. They help us to ascertain various sources of error and, thus, provide us with more confidence in our inferences. Creative strategies like diaries, passive technological monitoring, or beepers are sometimes used (Robinson, 1999), and these provide more or less control over the when and frequency of measurement question.

We should add that the use of multiple measures over time not only helps to determine the lag for a particular X,Y relationship but is the *only* way we can test for more complex relationships. As Kelly and McGrath (1988) point out, we need at least three assessments to look at a curvilinear relationship; four for oscillation; and perhaps more for rhythms, spirals, and cycles. Given the recent George and Jones (2000) recommendation that we construct and test for more elaborate theories involving these types of changes over time, we need to be especially sensitive to the frequency of observation and the timing of these assessments.

### Stability

The frequency of measurement question also raises the issue of measure stability (reliability). Whenever X or Y changes over time (whether predicted or not), the question we must ask is why. Changes in the assessment of a variable over time can be due to random error, systematic sources of error, or systematic change. Golembiewski, Billingsley, and Yeager (1976) suggest the terms *alpha*, *beta*, and *gamma* for systematic changes, where alpha change represents real substantive change, gamma change occurs when there is a change in the meaning of the construct, and beta change occurs when subjects alter their subjective metric or scale, resulting in a recalibration of the metric or scale. Thus, a systematic change in scores does not always reflect a "real" change in the construct. When the measurement takes place and how often are important.

Test-retest assessments (if variables are assessed during a steady-state or equilibrium period) can provide information on the stability (reliability) of X and Y, and time-series designs, as we mentioned earlier, also assess sources of error over time. More complex designs may involve multiple occasions when X occurs or the addition of comparison groups and patch-up designs (Algina & Swaminathan, 1979). Although, again, most authors do not address *when* these measures should take place, the proper use of these designs can help to uncover other sources of inferential error that may be important for this when question. Most important, the multiple assessments can help to capture "surrounding conditions" or other variables, besides X and Y, that vary over time (e.g., fatigue, satiation, wear, and even longer-term variables associated with seasonal cycles). "Attention is almost never given to the misspecification due to a serious unmeasured variables problem" (James & Brett, 1984: 317).

Time-series designs often include numerous observations on a relatively small number of persons, objects, or organizations. In some cases they are used with only one individual or organization. Panel designs, however, usually involve the use of large samples of subjects but only a few occasions of measurement (Willett & Sayer, 1994). Obviously, multiple measures on large samples would be ideal for assessing various factors influencing interindividual and intraindividual change. The type of question asked (relationship hypothesized) helps to determine both the appropriate method and analysis.

## THE CAUSAL CYCLE AND THE MODERATION BY CAUSAL CYCLE CURVE

Without theoretical guidance as to when to measure X and Y (in any of our configurations), we may gather data at inappropriate times. By "appropriate" we mean that the timing of measurement provides data that can be used to obtain unbiased estimates of a causal parameter (or parameters) linking X to Y (cf. Bryk & Raudenbush, 1992; Heise, 1975; James, Mulaik, & Brett, 1982). In this section we introduce two concepts to assist in this examination of when to measure: the *causal cycle* and the *moderation by causal cycle curve* (MCC curve). We use a simple example drawn from configuration 1 to acquaint readers with these concepts.

A bank wishes to see if increased monitoring (X) increases employee aggression (Y; e.g., malingering). Management takes two branches that are essentially alike and uses a camera and review of e-mails in branch I but not branch II. Aggression is assessed several weeks prior to monitoring, and the question becomes when to obtain postintervention aggression measures.

The causal cycle is designed to assist with this question. Let Y represent frequency of aggressive acts per week in the bank. The causal cycle for Y in branch I is

equilibration period →

equilibrium condition → entropic period

The equilibration period is the causal interval—the time period that it takes X to affect Y and for the values on Y to equilibrate, which means to change, and for the changes to reach a temporary state of constancy (Heise, 1975; James et al., 1982). Presumably, it may take some months for the weekly frequency of acts of aggression to increase and then to stabilize. We personally do not know what the equilibrium period is, but for didactic purposes, let us assume for now that it is three months.

When the scores on Y have equilibrated or stabilized, Y is said to have entered an equilibrium-type condition (cf. James et al., 1982). The effects of X on Y have worked their way through the system and have reached a temporary state of constancy. The scores on Y (and X) may change, but to maintain an equilibrium-type condition, the change must be small and very rapid so that a temporary equilibrium is quickly reestablished (Simonton, 1977).

Let us assume that following the three-month equilibration period, the frequencies of aggressive acts in branches I and II remain in an equilibrium-type condition for another four months. This is when Y should be measured. The major transition, transformation, or alteration of Y, caused by X, has transpired. The changes in the true scores on Y have equilibrated and will remain stable for a period. This period—the equilibrium-type condition—may be long or short. If short, it should be of sufficient length to justify statistical inference and generalizability (James et al., 1982).

At the end of the four-month equilibrium-type condition, the true and observed scores on Y begin to change. The anger could wear off or be spent, or some third variable could start to influence Y. The basic point is that (1) Y is no longer stable and begins to change, and (2) the conditional probability linking X to the prediction of Y tends toward uncertainty (Kelly & McGrath, 1988). These events herald the beginning of the entropic period. The entropic period is essentially infinite, culminating when changes in Y are completely uncertain with respect to a given set of measurements on X. The entropic period is the final state in the causal cycle. Note that the entropic period applies only for values on X associated with a specific causal cycle. New measurements on X start a new causal cycle.

We indicated above that measurements on Y were taken during the equilibrium-type condition. Several avenues are now available to test the hypothesis that monitoring engenders aggression. Suppose for analysis purposes that we select ANCOVA, conducted as a form of hierarchical regression (cf. Cohen & Cohen, 1983). The key analytic equation takes the form

$$Y = A + B_1 PRE + B_2 X + e \tag{1}$$

where A is the intercept, $B_1$ is the unstandardized regression weight for the scores on the pretest, PRE is the pretest (i.e., frequency of aggressive acts prior to introduction of monitoring), $B_2$ is a regression weight that indicates the effects of monitoring, X is the treatment variable that indexes whether monitoring did or did not occur (i.e., $X = 0$ or $X = 1$), and e is the error or disturbance term.

Not shown is a term carrying the interaction between the pretest and X (i.e., $B_3PRE \times X$). We assume that a hierarchical analysis has shown $B_3$ to be nonsignificant.

If monitoring causes a net increase in the frequency of aggressive acts, then $B_2$ in Equation 1 will be significant. Given that Y is appropriately measured during the equilibrium-type condition, let us assume that $B_2$ is significant, perhaps at an impressive level (e.g., $p < .00001$). The accompanying multiple correlation is a respectful value, such as .60.

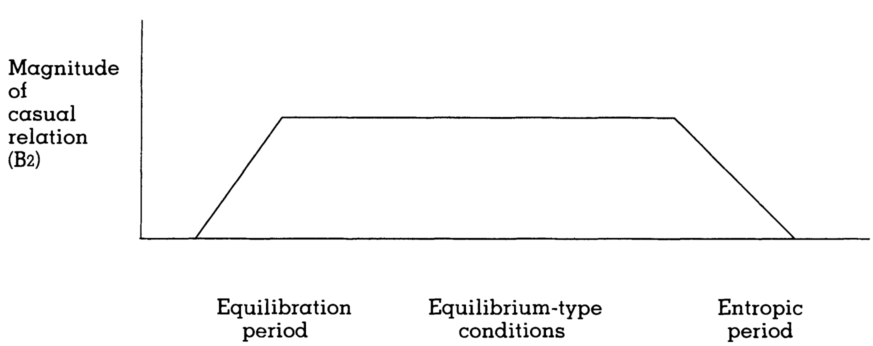
Now let us consider the ramifications of misspecifying the when of measurement. If we measured Y during the equilibration period, before the true scores on Y stabilized, we would expect $B_2$ to be underestimated in relation to the $B_2$ we estimated when Y was measured during the equilibrium-type condition (James et al., 1982). Instability in the true scores on Y acts much like random measurement error; the scores introduce an aspect of randomness or unpredictability to regression of Y on X, thus attenuating the absolute magnitude of $B_2$. The more unstable Y is, the greater the underestimation. Thus, we can infer that (1) the earlier Y is measured in the equilibration period, the greater the underestimation of $B_2$, whereas (2) the later Y is measured in the equilibration period, the smaller the underestimation of $B_2$. The mirror image of this process can be expected for measuring Y during the entropic period. When Y begins to change, the ability of X to predict Y becomes increasingly uncertain. Thus, (1) the earlier Y is measured in the entropic period (i.e., the shorter the time since the equilibrium-type condition elapsed), the smaller the underestimation of $B_2$, whereas (2) the later Y is measured in the entropic period, the greater the underestimation of $B_2$.

Figure 1 summarizes the arguments presented above regarding accuracy of estimation in relation to correctly versus incorrectly specifying when to measure Y. On the ordinate we have the magnitude of the causal relation, which is represented by $B_2$ in the current example. (The magnitude of causal relation could be estimated by many statistics, including path coefficients, structural parameters, Rs, variance components, and so forth.) On the abscissa we have the causal cycle. The causal relation reaches its maximum, and accurate, value when Y is correctly measured during the equilibrium-type condition. The causal relation is increasingly underestimated as the measurement of Y departs from the optimum equilibrium-type condition.

Figure 1 suggests that the *magnitude of the causal relation* (not the value of Y) linking X to Y is moderated by the stage of the causal cycle in which Y is measured and, thus, $B_2$ is estimated. This pattern is the MCC curve. As long as the X to Y relationship is linear, and often when it is not, the MCC curve will have the convex shape shown in Figure 1. The height of the curve, the degrees of underestimation, and the steepness of the slopes all depend on the data at hand. However, the common denominator for all MCC curves is that as long as X causes Y, then misspecification of when—that is, measuring Y during the equilibration period or the entropic period—will engender underestimation of the strength of the causal relationship.

We also note that the slopes in the equilibration and entropic periods need not be symmetric. Y might stabilize quickly, but the effects of X might take a comparatively longer time to wear off. Or Y may take a long time to equilibrate, but the effects of X on Y might wear off very quickly.

**FIGURE 1**
**The MCC Curve**



Relationships among all of the time periods in the causal cycle are essentially uncharted. This makes it all the more crucial that our theory help us in determining when and how frequently we should assess X and Y.

### Additional Illustrations

We can extend the discussion of causal cycles and MCC curves to more complex configurations relating causes to effects. In the first of two illustrations, we consider how changes in X are related to changes in Y (configuration 4). The hypothesis tested is that higher average annual raises (given in January) increase the average incidence of organizational citizenship behaviors (OCBs) during that subsequent year. Data are collected for 100 banks; the average raise and average incidence of OCBs are collected for each bank for each of twenty years. A time series is computed for each bank, in which yearly average OCBs are regressed on yearly average raises—that is, OCB(tj) = A(j) + B(j)AVG RAISE (tj) = e(tj), where t refers to year (t takes on values of 1 to 20), and j refers to bank (j takes on values of 1 to 100). The resulting 100 estimates of B(j) are treated statistically as level 1 estimates in a hierarchical linear model (HLM; cf. Bryk & Raudenbush, 1992; Hofman, Griffin & Gavin, 2000).

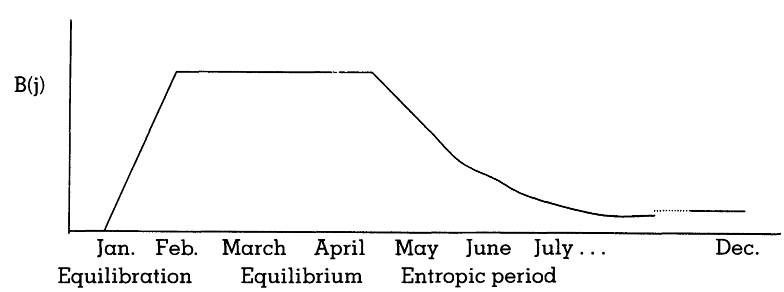
A level 2 analysis is conducted to determine if (1) the B(j) vary over banks, which would suggest that the within-bank OCBs on AVG RAISE regression slopes vary over banks, and (2) variation in mean OCBs is related more strongly to differences in raises between banks or to pooled variation in raises within banks.

Every one of the preceding analyses will misspecify the X,Y relationship if the B(j) are underestimated because of misspecification of the causal interval. To avoid such misspecification, the causal cycle relating the timing of raises to the timing of measurement of OCBs is crucial. Suppose the causal cycle is very rapid. Raises given in January quickly affect OCBs (i.e., a short equilibrium-type condition lasts for three months [February through April], after which the effect of raises on OCBs dampens quickly [the entropic period starts in May and lasts through December, when a new cycle starts]). The MCC curve produced by this causal cycle is shown in Figure 2a. Of key importance is that any measurement of OCBs taken after April of each year will underestimate the effects of average raises on average OCBs. The entire HLM analysis would be based on biased estimates.

Now let us turn to a bit more complex problem. Our hypothesis here is that time on task (variable X) leads to self-set goals of greater difficulty (variable Y). The rationale is that time on task serves as a surrogate for experience and learning, and as people learn and develop expertise and a sense of self-efficacy (Gist & Mitchell, 1992), they set progressively more difficult goals. The regression of (self-set) goal difficulty on time (experience, learning) is called a "growth curve." We further hypothesize that (1) different individuals have different growth curves, (2) intelligence serves to explain these differences such that the growth curves are steeper initially for more intelligent folks, and (3) irrespective of initial steepness and final goal difficulty, all growth curves flatten out.

To test this model we fit a growth curve using HLM to each of the N subjects for whom we have data on goal difficulty. These data are collected on a continuous basis (e.g., collected on a weekly basis for twelve months). The within-subject times-series equation fit for each subject is

**FIGURE 2a**
**MCC Curves for Additional Illustration**



$$Y(ij) = A(k) + B^2(i)X(ij) + e(ij) \tag{2}$$

where i = 1 to N subjects and j = 1 to 52 weeks.

The parameter estimate for the slope is in quadratic form to capture the predicted curvilinear relationship within subjects between goal difficulty and time (i.e., difficulty levels off). N such equations are estimated in the level 1 (within-subject) HLM analysis. The level 2 (between-subject) analysis consists of testing the $B^2(i)$ on measures of individual intelligence. If intelligence accounts for variation in the $B^2(i)$, then individual differences in the steepness of the within-subject growth curves are assumed to be attributable to learning potentials.

When X and Y are in essentially continuous form and X exerts constant influence on Y, we tend to think of the causal cycle in terms of a time interval that is bounded by a logical start point and a logical end point (Kelly & McGrath, 1988). In the present case, a reasonable interval appears to start when the subject begins a new activity and to end when the subject has completely learned all tasks and is no longer changing his or her self-set goals. This would generally be where goal difficulty reaches an equilibrium-type condition. The equilibration period encompasses the period prior to the flattening out of the growth curve. There may or may not be an entropic period. That is, an individual may continue to set the same goals for the remaining time that she or he is in the activity.

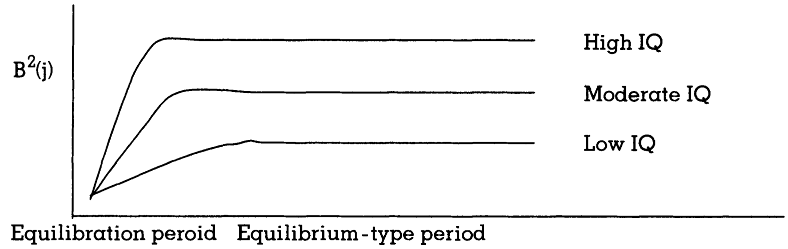
Irrespective of whether an entropic period is involved, this model generates a series of MCC curves, as illustrated in Figure 2b. Because level of intelligence moderates the length of the equilibration period (i.e., the steepness of the growth curve), we must have different MCC curves for each key segment of values on the moderator (one might think of the moderator as a blocking variable with multiple values, technically infinite). The key messages conveyed by Figure 2b are that (1) a sufficient length of time has to elapse in the collection of data (i.e., goal difficulty has to stabilize) before an attempt is made to fit the level 1 time series, and (2) those periods vary over individuals (although there is also a lengthy period when all the growth curves have equilibrated). The primary threat to parameter estimation is to estimate too early. It is possible, for example, to fit a linear model what is truly a curvilinear function.

## Analytical Tools

One of the reasons that we can ask the questions posed here and have any hope of resolving them is that serious advances have been made recently in the methods of studying time-embedded data. Of particular note are HLM (Bryk & Raudenbush, 1992; Deadrick, Bennett, & Russell, 1997; Hofmann, 1997; Hofmann et al., 2000; Vancouver & Putka, 2000), latent growth modeling (LGM; Chan, 1998; Duncan, Duncan, & Hops, 1996; Eid & Langeheine, 1999; McArdle & Woodcock, 1997; Muthen & Curran, 1997; Willett & Sayer, 1994), and pooled cross-sectional time series (PCSTS; Farkas & Tetrick, 1989; Hom & Griffeth, 1991; Sayrs, 1989; Simonton, 1977).

While space limits preclude a thorough description of all the ways these procedures can help, especially in the assessment of change (see Chan, 1998), there are some points that should be mentioned. First, these techniques can tell us if and when a variable changes over time and how it changes. With just repeated measures on a single variable for a single person, we can assess the type of trajectory or growth curve and the rate of change. With multiple assessments of multiple subjects and vari-

**FIGURE 2b**
**B(j) for Growth Curves Moderated by Intelligence**



ables, we can tell whether these trajectories are homogeneous or heterogeneous (and what variables cause that heterogeneity) across subjects. We can tell whether the change is linear or not. Finally, tests for some types of correlated residuals (causing autocorrelated errors) are part of the evolving LGM models (Chan, 1998). Thus, powerful statistical tools are available to help us *inform* theory by discovering how and when variables change or to *test* more precise hypotheses that predict when and how change will occur.

### Theoretical Thinking

What we indicate in this section is that theoretical propositions need to display a specific awareness of time and context. Such issues demand both complex and creative thinking on the part of the theorist. We suggest, as a first step, that all the relationships be diagrammed, as we did in Table 1. This exercise gives one an immediate sense of what time issues are relevant, like lags, change, and reciprocal causality. It would then be important to try to state what the lags would be and why. What is associated with the lags? When do particular events occur? Zaheer et al. (1999) are helpful in suggesting different ways to think about time lags. Second, if the configuration suggests that change in X or Y is predicted, the theorist must consider the amount and rate of change. In addition, one must speculate about other variables that may be associated with and/or cause the change (Chan, 1998). We suggest, as a third step, that the theorist construct an MCC curve. This activity would require specification of the equilibration, equilibrium, and entropic periods for all the variables involved. Fourth, based upon the first three steps, the theorist should select a design that specifies the timing and frequency of measurement. Fifth, the theorist should become acquainted with the appropriate statistical procedures available to help discover or test causal relationships involving changes over time. Data and theory interact using both induction and deduction to inform the investigator about time.

Of course, the process of thinking about time and theory is much more than a series of steps involving specific questions and structural aids. The theorist also needs to "think out of the box." While little has been written about this more creative aspect of theorizing about time, there is some precedence for it. First, if possible, the theorist should go to the place of the actors involved and should talk to these people, asking them about the timing of events as they unfold. Second, the theorist should know the literature concerning self-reports involving time. For example, Buehler, Griffin, and Ross (1994) have demonstrated that people often underestimate the time it will take them to do something while overestimating for others. Also, people's reconstruction of events in the past may be different from their report while the experience is unfolding (Mitchell, Thompson, Peterson, & Cronk, 1997). Third, the theorist should try to envision a larger cycle of events than simply his or her focus (Zaheer et al., 1999) and should concentrate on antecedents and consequences that may be more distal. Fourth, historians also suggest the use of analogies (May, 1986). The theorist should make explicit comparisons of similarities and differences between analogies, with a focus on time. May (1986) also suggests the construction of a story, including the textual details suggested by Whetten (1989) of who, what, where, why, and, of course, when. Such story building helps in the understanding of patterns of events and the unfolding of a timeline. Finally, Weick (1989) suggests that we engage in "thought trials" to enrich our theoretical understanding. One can imagine how events might be different, given specific changes. Such trials can include thinking about when effects occur and at what rate. In addition, speculation about "other" variables that cause error or could be systematically related to theoretical elements should be examined. Thoughtfully engaging in such activities can substantially aid the theorist's understanding of a phenomenon.

## DISCUSSION

Karl Popper reminded us years ago that "the work of the scientist consists in putting forward and testing theories" (1959: 31). Popper goes on to elaborate on the idea that only by introducing more complex theory and subjecting it to empirical verification and falsification can a science grow and develop. Our purpose here is to suggest that one way to enrich our theories *and* subject them to falsification is to be more precise theoretically, and methodologically, about *when* events occur.

### Why Is This Important?

The discipline of management has been criticized on many issues. The research we do is described as trivial, insignificant, and nothing but common sense. Bedeian goes so far as to say that "much of what passes as management research is arguably sterile, simple-minded, and, consequently, increasingly irrelevant to management practice" (1996: 317). Kilduff and Mehra summarize one view of organizational research as "oversimplified, narrow, and disappointingly irrelevant" (1997: 454). These are serious indictments.

Have we reached a crisis point? Kuhn argues that "crisis is a necessary precondition for the emergence of novel theories" (1970: 77). We are not sure if that point has been reached in the field of management. However, we are concerned that few if any of our theories are ever disconfirmed. Ignored, yes; rejected, infrequently. Our field has been slow to change.

We would suggest that part of the problem revolves around the quality of the questions we ask. The typical hypothesis suggests simple relationships like X is associated with, or a cause of, Y. With precise measures and large samples, these types of hypotheses approach a .5 probability of support (Meehl, 1967); almost any correlation between two variables or mean difference between two groups is statistically significant. One remedy for this problem is to concentrate on substantive significance rather than statistical significance; look at variance accounted for. But a different solution requires more complex theory. We should push ourselves to be more precise. When does the change occur, at what rate does it occur, and with what in the environment is the change associated? How, exactly, does time influence X or Y, or both, or their interrelationship? Does time moderate the X,Y relationship? Does Y cause X, as well as X cause Y? What errors are associated with time?

It is theory that informs method and analysis. By being more sensitive and precise about how time is involved in our theories, we will develop theories that are richer and more subject to disconfirmation (Bacharach, 1989; Platt, 1964). Ultimately, these processes will result in better theory and better practice. "A deliberate and thoughtful consideration of the role of time in different organizational phenomena would represent a significant advance in the study of temporal concerns in the organizational literature" (Okhuysen, 1999: 23).

### What Is New

Obviously, one reason we in the field should consider time with more precision is because of these criticisms. But another reason is that the field of management has changed dramatically over the last fifteen years with respect to theory, design, and analysis. From a theoretical perspective, such authors as McGrath (1988), Ancona and Chong (1996), and George and Jones (2000) are suggesting we consider time in all of its theoretical richness. Developing theory and testing ideas including notions of cycles, rhythms, spirals, and oscillation will require thinking about time and method in more elaborate and precise ways.

Our methods have changed as well. The growth and infusion of more "macro" orientations, such as business policy, have changed the methodological landscape, as has an increased disdain (or at least a more critical eye) toward laboratory research. Much more research is being conducted in the field. In fact, only two laboratory studies were published in *AMJ* in 1999 (one in *ASQ*). Many more designs without control groups and random assignment are being used, and such research is especially vulnerable to a variety of confounds that are due to time. Of particular concern to us are the problems that arise with an unspecified third variable that may interact with the lags, change, and rate of change for variables in the study. "Management research may be moving even further away from rigor" (Scandura & Williams, 2000: 1259).

Two statistical issues also should be mentioned. One of the more serious problems of any systematic misspecification of when events occur is that an X,Y relationship is over- or underestimated. With the infusion of meta-analyses in our literature, such errors can lead to faulty conclusions about both theory and practice. In addition, many of our confirmatory techniques using structural equation modeling (Kline, 1998) may end up indicating incorrect causal paths and causal effects and may also suggest the incorrect ordering of events. We can get the existence of an effect wrong, as well as its strength and direction. "Without good theoretical work, it is obvious that little benefit is derived from the use of causal models" (Williams & Podsakoff, 1989: 285).

## Conclusion

Paradigms usually present themselves as an integrated whole. As Kuhn (1970) points out, paradigms integrate theory, method, and standards. We believe that the management discipline needs to seriously consider issues of time, especially *when* events occur, in both theory and method. Such reconsideration may change our standards for what is acceptable in our journals and help our discipline to progress through the process of disconfirmation. In addition, power in universities (money, resources) flows to disciplines that have successful paradigms (Salancik, 1987). Kelly and McGrath concluded that there was "a vicious cycle of neglect of temporal effects in substantive, conceptual and methodological domains" (1988: 56). We can and should do better.

## REFERENCES

Algina, J., & Swaminathan, H. 1979. Alternatives to Simonton's analyses of the interrupted and multiple-group time-series designs. ***Psychological Bulletin,*** 86: 919–926.

Amabile, T. M., & Conti, R. 1999. Changes in the work environment for creativity during downsizing. ***Academy of Management Journal,*** 42: 630–640.

Ancona, D., & Chong, C.-L. 1996. Entrainment: Pace, cycle, and rhythm in organizational behavior. ***Research in Organizational Behavior,*** 18: 251–284.

Andersson, L. M., & Peterson, C. M. 1999. Tit for tat? The spiraling effect of incivility in the workplace. ***Academy of Management Review,*** 24: 452–471.

Bacharach, S. B. 1989. Organization theories: Some criteria for evaluation. ***Academy of Management Review,*** 14: 496–515.

Bedeian, A. G. 1996. Thoughts on the making and remaking of the management discipline. ***Journal of Management Inquiry,*** 5: 311–318.

Berger, P. L., & Luckmann, T. 1966. ***The social construction of reality.*** London: Penguin.

Bloom, J. D. 1999. The performance effects of pay dispersion on individuals and organizations. ***Academy of Management Journal,*** 42: 25–40.

Bluedorn, A. C., & Denhardt, R. B. 1988. Time and organizations. ***Journal of Management,*** 14: 299–320.

Bluedorn, A. C., Kaufman, C. F., & Lane, P. M. 1992. How many things do you like to do at once? An introduction to monochronic and polychronic time. ***Academy of Management Executive,*** 6(4): 17–26.

Bryk, A. S., & Raudenbush, S. W. 1992. ***Hierarchical linear models.*** Newbury Park, CA: Sage.

Buehler, R., Griffin, D., & Ross, M. 1994. Exploring the planning fallacy: Why people underestimate their task completion times. ***Journal of Personality and Social Psychology,*** 67: 366–381.

Cable, D. M., & Murray, B. 1999. Tournaments versus sponsored mobility as determinants of job search success. ***Academy of Management Journal,*** 42: 439–449.

Callister, R. R., Kramer, N. W., & Turban, D. B. 1999. Feedback seeking following career transitions. ***Academy of Management Journal,*** 42: 429–438.

Campbell, D. T., & Stanley, J. C. 1963. ***Experimental and quasi-experimental designs for research.*** Chicago: Rand McNally.

Carlson, J. R., & Zmud, R. W. 1999. Channel expansion theory and the experiential nature of media richness perceptions. ***Academy of Management Journal,*** 42: 153–170.

Chan, D. 1998. The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). ***Organizational Research Methods,*** 1: 421–483.

Clark, P. 1985. A review of the theories of time and structure for organizational sociology. ***Research in the Sociology of Organizations,*** 4: 35–79.

Cohen, L. E., & Cohen, P. 1983. ***Applied multiple repression/correlation analysis for the behavioral sciences*** (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Combs, J. G., & Ketchen, D. J., Jr. 1999. Can capital scarcity help agency theory explain franchising? Revisiting the capital scarcity hypothesis. ***Academy of Management Journal,*** 42: 196–207.

Cook, T. D., & Campbell, D. T. 1979. ***Quasi-experimentation: Design and analysis issues for field settings.*** Chicago: Rand McNally.

Cosier, R. A., & Dalton, D. R. 1983. Equity theory and time: A reformulation. ***Academy of Management Review,*** 8: 311–319.

Deadrick, D., Bennett, N., & Russell, C. 1997. Using hierarchical linear modeling to examine dynamic performance criteria over time. ***Journal of Management,*** 23: 745–757.

Duncan, S. C., Duncan, T. E., & Hops, H. 1996. Analysis of longitudinal data within accelerated longitudinal designs. ***Psychological Methods,*** 1: 236–248.

Eid, M., & Langeheine, R. 1999. The measurement of consistency and occasion specificity with latent class models: A new model and its application to the measurement of affect. ***Psychological Methods,*** 4: 100–116.

Ettlie, J. E. 1977. Real-time studies in organizational research. ***Academy of Management Review,*** 2: 298–302.

Farkas, A. J., & Tetrick, L. E. 1989. A three-wave longitudinal analysis of the causal ordering of satisfaction and commitment on turnover decisions. ***Journal of Applied Psychology,*** 74: 855–868.

Faulconer, J. E., & Williams, R. N. 1985. Temporality in human action: An alternative to positivism and historicism. ***American Psychologist,*** 11: 1179–1188.

George, J. M., & Jones, G. R. 2000. The role of time in theory and theory building. ***Journal of Management,*** 26: 657–684.

Gergen, K. J. 1973. Social psychology as history. ***Journal of Personality and Social Psychology,*** 26: 309–320.

Gersick, C. J. G. 1988. Time and transition on work teams: Toward a new model of group development. ***Academy of Management Journal,*** 31: 9–41.

Gibson, C. B. 1999. Do they do what they believe they can? Group efficacy and group effectiveness across tasks and cultures. ***Academy of Management Journal,*** 42: 138–152.

Gist, M., & Mitchell, T. R. 1992. Self-efficacy: A theoretical analysis of its determinants and malleability. ***Academy of Management Review,*** 17: 183–211.

Golembiewski, R. T., Billingsley, K., & Yaeger, S. 1976. Measuring change and persistence in human affairs: Types of change generated by OD designs. ***Journal of Applied Behavioral Science,*** 12: 133–157.

Goodman, P. 2000. ***Missing organizational linkages: Tools for cross-level research.*** Thousand Oaks, CA: Sage.

Gurvitch, G. 1964. ***The spectrum of social time.*** Dordrecht, Netherlands: Reidel.

Heise, D. R. 1975. ***Causal analysis.*** New York: Wiley.

Heneman, R. L., & Wexley, K. N. 1983. The effects of time delay in rating and amount of information observed on performance rating accuracy. ***Academy of Management Journal,*** 26: 677–686.

Hofmann, D. A. 1997. An overview of the logic and rationale of hierarchical linear models. ***Journal of Management,*** 23: 723–744.

Hofmann, D. A., Griffin, M. A., & Gavin, M. B. 2000. The application of hierarchical linear modeling to organizational research. In K. J. Klein & S. W. J. Kozlowski (Eds.), ***Multilevel theory, research, and methods in organizations*** (SIOP Frontier Series): 467–511. San Francisco: Jossey-Bass.

Hom, P. W., & Griffeth, R. W. 1991. Structural equations modeling test of a turnover theory: Cross-sectional and longitudinal analyses. ***Journal of Applied Psychology,*** 76: 350–366.

James, L. R., & Brett, J. M. 1984. Mediators, moderators, and tests for mediation. ***Journal of Applied Psychology,*** 69: 307–321.

James, L. R., Mulaik, S. A., & Brett, J. M. 1982. ***Causal analysis assumptions, models, and data.*** Beverly Hills, CA: Sage.

Judiesch, M. K., & Lyness, K. S. 1999. Left behind? The impact of leaves of absence on managers' career success. ***Academy of Management Journal,*** 42: 641–651.

Jung, D. I., & Avolio, B. J. 1999. Effects of leadership style and followers' cultural orientation on performance in group and individual task conditions. ***Academy of Management Journal,*** 42: 208–218.

Kelly, J. R., & McGrath, J. E. 1988. ***On time and method.*** Newbury Park, CA: Sage.

Kenny, D. A. 1975. Cross-lagged panel correlation: A test for spuriousness. ***Psychological Bulletin,*** 82: 887–903.

Kilduff, M., & Mehra, A. 1997. Postmodernism and organizational research. ***Academy of Management Review,*** 22: 453–481.

Kline, R. B. 1998. ***Structural equation modeling.*** New York: Guilford.

Kuhn, T. S. 1970. ***The structure of scientific revolutions.*** Chicago: University of Chicago Press.

Liden, R. F., & Mitchell, T. R. 1988. Ingratiation behaviors in organizational settings. ***Academy of Management Review,*** 13: 572–587.

Lindsley, D. H., Brass, D. J., & Thomas, J. B. 1995. Efficacy-performance spirals: A multi-level perspective. ***Academy of Management Review,*** 20: 645–678.

May, N. 1986. ***Thinking in time.*** New York: Free Press.

McArdle, J. J., & Woodcock, R. W. 1997. Expanding test-retest designs to include developmental time-lag components. ***Psychological Methods,*** 2: 403–435.

McGrath, J. E. 1988. ***The social psychology of time.*** Newbury Park, CA: Sage.

McGrath, J. E., & Kelly, J. R. 1986. ***Time in human interaction: Toward a social psychology of time.*** New York: Guilford.

Meehl, P. E. 1967. Theory testing in psychology and physics: A methodological paradox. ***Philosophy of Science,*** 34: 278–295.

Mitchell, T. R., Thompson, L., Peterson, E., & Cronk, R. 1997. Temporal adjustments in the evaluation of events: The "rosy view." ***Journal of Experimental Social Psychology,*** 33: 421–448.

Mone, M. 1994. Comparative validity of two measures of self-efficacy in predicting academic goals and performance. ***Educational and Psychological Measurement,*** 54: 516–529.

Mosakowski, E., & Earley, P. C. 2000. A selective review of time assumptions in strategy research. ***Academy of Management Review,*** 25: 796–812.

Muthen, B. O., & Curran, P. J. 1997. General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. ***Psychological Methods,*** 2: 371–402.

Okhuysen, G. A. 1999. ***The many faces of time: Temporal considerations in the study of organizational decision making.*** Paper presented at the annual meeting of the Academy of Management, Chicago.

Pentland, W. E., Harvey, A. S., Lawton, M. P., & McColl, M. A. 1999. ***Time use research in the social sciences.*** New York: Kluwer Academic.

Platt, J. R. 1964. Strong inference. ***Science,*** 146: 347–353.

Popper, K. 1959. ***The logic of scientific discovery.*** New York: Basic Books.

Robinson, J. P. 1999. The time-diary method: Structure and uses. In W. E. Pentland, A. S. Harvey, M. P. Lawton, & M. A. McColl (Eds.), ***Time use research in the social sciences.*** New York: Kluwer Academic.

Sackett, P. R., & Larson, J. R., Jr. 1990. Research strategies and tactics in industrial and organizational psychology. In M. Dunnette (Ed.), ***Handbook of industrial and organizational psychology,*** vol. 1: 419–428. Palo Alto, CA: Consulting Psychologists Press.

Salancik, G. R. 1987. Power and politics in academic departments. In M. D. Zanna & J. M. Darley (Eds.), ***The compleat academic:*** 61–84. New York: Random House.

Sayrs, L. W. 1989. ***Pooled time series analysis.*** Newbury Park, CA: Sage.

Scandura, T. A., & Williams, E. A. 2000. Research methodology in management: Current practices, trends and implications for future research. ***Academy of Management Journal,*** 43: 1248–1264.

Schmitz, B., & Skinner, E. 1993. Perceived control, effort and academic performance: Interindividual, intraindividual, and multivariate time series analyses. ***Journal of Personality and Social Psychology,*** 64: 1010–1028.

Simonton, D. K. 1977. Cross-sectional time-series experiments: Some suggested statistical analyses. ***Psychological Bulletin,*** 84: 489–502.

Vancouver, J. B., & Putka, D. J. 2000. Analyzing goal-striving processes and a test of the generalizability of perceptual control theory. ***Organizational Behavior and Human Decision Processes,*** 82: 334–362.

Waller, M. J. 1999. The timing of adaptive group responses to non-routine events. ***Academy of Management Journal,*** 42: 127–137.

Weick, K. E. 1989. Theory construction as disciplined imagination. ***Academy of Management Review,*** 14: 516–531.

Welbourne, T. M., & Cyr, L. A. 1999. The human resource executive effect on initial public offering firms. ***Academy of Management Journal,*** 42: 616–629.

Westphal, J. D. 1999. Collaboration in the boardroom: Behavioral and performance consequences of CEO-board social ties. ***Academy of Management Journal,*** 42: 7–24.

Whetten, D. H. 1989. What constitutes a theoretical contribution? ***Academy of Management Review,*** 14: 490–495.

Willett, J. B., & Sayer, A. G. 1994. Using covariance structure analysis to detect correlates and predictors of individual change over time. ***Psychological Bulletin,*** 116: 363–381.

Williams, L. J., & Podsakoff, P. M. 1989. Longitudinal field methods for studying reciprocal relationships in organizational behavior research: Toward improving causal analysis. ***Research in Organizational Behavior,*** 11: 247–292.

Zaheer, S., Albert, S., & Zaheer, A. 1999. Time scales and organizational theory. ***Academy of Management Review,*** 24: 725–741.

**Terence R. Mitchell** is the Edward E. Carlson Professor of Business Administration and professor of psychology at the University of Washington. He received his Ph.D. in social psychology from the University of Illinois. His research interests are decision making, leadership, motivation, and retention.

**Lawrence R. James** is the Pilot Oil Chair of Excellence Professor and director of the Center for Organizational Research in the Industrial/Organizational Psychology Program at the University of Tennessee. He received his Ph.D. in industrial psychology from the University of Utah. His current research interests include confirmatory analysis, measurement of personality, and organizational and psychological climate.