2-3-1999

Variance Explained: Why Size Does Not (Always) Matter

Mark Fichman
Carnegie Mellon University, mf4f@andrew.cmu.edu

Follow this and additional works at: http://repository.cmu.edu/tepper
Part of the Economic Policy Commons, and the Industrial Organization Commons

Published In
Research in Organizational Behavior, 21.
Variance Explained: Why size does not (always) matter

Mark Fichman
Graduate School of Industrial Administration
Carnegie-Mellon University
Pittsburgh PA 15213-3890
412-268-3699
e-mail:mf4f@cmu.edu
fax:412-268-7064

February 3, 1999

1This paper is dedicated to Doug Jenkins, who showed me the value of and pleasure in studying things methodological. He was a wonderful guide. I wish I could show him this paper; I hope he would have liked it. Thanks to Fernando Olivera for encouraging me to write this paper. Thanks to Bob Sutton for really pushing me to write this paper. Thanks to Barry Staw and Bob Sutton for their editorial comments and advice, and to participants in the Organizational Behavior and Theory seminar at GSIA.
Abstract

I examine the role of explaining variance in the construction of explanatory theory. Explaining variance can be an insufficient basis for evaluating a theory (Lieberson, 1985). Starting with this insight, I suggest that models that provide explanations of variance do not necessarily provide good explanations of causal mechanisms. I then explore the utility of process models and theories (Mohr, 1982) relative to variance theories. I clarify the role of stochastic processes in such model building and discuss the implications of such processes for evaluating explanatory ‘adequacy’. Under some conditions, explaining variance may be neither a necessary nor a sufficient condition for good explanatory theory. I then identify some implications of this argument for developing and analyzing explanatory theory. These arguments are applied to two examples: (1) meta-analysis and (2) the disposition versus situation debate (a variant on the nature vs. nurture argument) to illustrate the implications of this process theory point of view.
The Variance Explained Criterion

The Pursuit of Variance

Organizational and social researchers frequently try to build variance theories (Mohr, 1982). The goal is to develop a theoretical structure that is consistent with (explains variance in) the data (Pfeffer, 1982) while being parsimonious, interesting and conceptually sound. Given the applied underpinnings of organizational research, we are interested in developing theories that maximize this consistency; explain as much of the variance as possible¹. A common assumption in our field is that the amount of variance explained is an indicator of the quality or explanatory power of a theory. In reading papers, we consider goodness of fit statistics like $R^2$ or $\eta^2$ and use them to evaluate explanatory adequacy. When effect size statistics are lacking, statisticians develop them in order to meet the demand for them (e.g., Bentler & Bonett, 1980). Effect size statistics are an essential requirement in any research report (Rosenthal & Rosnow, 1991), allowing readers to assess more than the significance of the effect, but the strength of the effect itself. However, we often subtly move from assessing strength of effects in a study to using effect sizes to assess the underlying theory, indexing the explanatory power of a theory by strength of effects. In some cases, this becomes a sort of competition or race between theories for which theory or model explains the most variance, and hence is declared the winner in

¹I use explanatory theory or explanation with reference to theory in the sense of providing a causal account (Salmon, 1984). The usage ‘explain variance’ and its variants are a different usage not invoking this sense of a causal accounting. Rather, explaining variance is a common usage meaning a statistical association between an independent variable set and a dependent variable set. The slippage between and confusion of these two senses, a causal account or explanation and a statistical association, explaining variance, is what I am concerned with here.
comparisons between theories. Variance explained indicators are sometimes explicitly used to compare theories (Cooper & Richardson, 1986) or used to gauge the adequacy or power of a theory against implicit standards or expectations.

Consider the study of leadership. The rise and fall of interest in and concern about leadership in organizational studies can be roughly indexed by the amount of variance explained by studies in the area. Pfeffer (1977) reviewed research on leadership, concluding that leaders have small effects, particularly relative to environmental factors. This led him to ask why so much attention is paid to leaders given their small effects on performance. In effect, Pfeffer argued that the theoretical assumption that leaders matter was questionable given that changes in leadership explained relatively little variance in organizational performance. The basis for his argument was that changes in leadership showed small effects on organizational performance relative to other indicators in several studies of leadership change (Lieberson & O’Connor, 1972; Salancik & Pfeffer, 1977(b)). Pfeffer concluded that leaders serve an important symbolic function, giving owners and others a sense of control. Others, looking at the same data, concluded that “leadership seemed to have a substantial effect on some organizational outcomes.” (Bass, 1990: 7). House, Spangler & Woycke (1991) presented an analysis of charismatic leadership effects which placed great emphasis on the amount of variance explained. The gist of their argument was that explaining anywhere from 20% to 66% of the variance was strong evidence of the impact of leadership, particularly if alternative explanations have been considered. In that paper, they clearly argue that such effect sizes call into question Pfeffer’s claims. This is one prototypic example of the use of vari-
ance explained arguments as a metric for the relative explanatory value or power of different theories or theoretical constructs.

Consider the study of job design. Salancik and Pfeffer’s (1977(a),1978) social information processing analysis of job attitudes led to a great interest in whether social information processing was a ‘better’ explanation than job characteristics theory for reactions to job characteristics. For example\(^2\), White and Mitchell (1979) set out to test if “social and informational cues may be better predictors of employees’ reactions to a task than the objective characteristics of the task. That is, job satisfaction and motivation may be more a result of how one’s co-workers react to a task than the task characteristics themselves.” (p. 1-2). They concluded that “the comments of co-workers were a more powerful motivating force than the actual properties of the task.” (p. 8). O’Reilly and Caldwell (1979) noted “the small amount of variance accounted for in the associations between job characteristics and outcome measures” (p. 157), and investigated and compared “the importance of objective task characteristics with the importance of information … from others” on “perceptions of task characteristics and the affective responses to the job.” (p. 158). Frequently, theory comparison involves assessing which of several theories explains more the variance in the dependent variable(s). Usually, the theory that carries the day is the one that accounts for or explains the most variance in the dependent variable(s). For any particular theory, the explanatory value of variables is often evaluated using the same proportion of variance explained criterion. “Researchers often act as if one variable is ‘better’ than another if it can account for more of

\(^2\)This example, from a different perspective, is also discussed by Cooper and Richardson, (1986). Their views are touched on below.
the variance; indeed we even talk sometimes about the ‘explanatory’ power of variables in exactly these terms” (Lieberson, 1985: 90).

Such uses of variance explained or other goodness of fit indices are an everyday part of how we evaluate research. Such uses can be extremely misleading, and their use can have deleterious and often pernicious effects. Social scientists focus too much attention on explaining variance (Lieberson, 1985). This focus on explaining variance can have detrimental consequences for theory development. In this paper, I begin by critically examining Lieberson’s argument. I offer some new and related arguments going beyond Lieberson’s position. In this examination, I focus on the relationship of scientific explanation to explanation of variance, broadening the scope of Lieberson’s criticisms. In particular, I identify and extend the implications of Lieberson’s argument for theory building and theory testing in social and organizational research.

**Variance Theories Defined**

In these examples and in much organizational and social research, explanatory theories have as one goal explaining variation in some dependent variable matrix $Y$ given a matrix of independent variables $X$. Mohr characterizes the variance approach, in contrast to the process approach (that I will turn to below), as having four characteristics.

1. $X$ is a “necessary and sufficient condition” (Mohr, 1982: 38) for $Y$.

   Logically, this means that $X \Rightarrow Y$ and $\sim X \Rightarrow \sim Y$, where $\sim$ is read ‘not’ and ‘$\Rightarrow$’ is read ‘implies.’
2. “A variance theory deals with variables” (Mohr, 1982: 39). That is, elements in $X$ and $Y$ can take on different values, and there is a mapping of variation in variables from $X$ on to variables in $Y$.

3. “A variance theory deals with efficient causes” (Mohr, 1982: 41). This condition allows Mohr to exclude spurious associations masquerading as causes. Others have made a similar observation in trying to relate statistical association to causality (Holland, 1986; Salmon, Jeffrey & Greeno, 1971; Suppes, 1970). In essence, there is a search for what Mohr calls “push-type” causality (Mohr, 1982: 41). Further, the effects of multiple causes, however complex their relationships with each other and the dependent variable(s), are potentially separable and identifiable.

4. “In variance theory, time ordering among the contributory (independent variables) is immaterial to the outcome (Mohr, 1982: 43).” Although there may be intervening or mediating variables which have a temporal order, they are not critical to meeting the necessary and sufficient condition (1) above. That is if

$$x_a \rightarrow x_b \rightarrow y,$$

where $\rightarrow$ denotes causality, and $x_a$ is causally prior to $x_b$, then $x_b$ is not needed to have a necessary and sufficient explanation of $y$. $x_b$ gives us information about the mediating path by which $x_a$ exerts its causal influence on $y$, but I do not need to know $x_b$ to have a complete explanation. Rather knowing $x_b$ provides more detail about the intervening mechanisms; something that is often desirable and
always, in principle, possible.

Variance theories are the primary means social researchers use to try to explain social phenomena. The variance approach informs experimental and quasi-experimental approaches to the design of research. The underlying premise is that an appropriately constructed variance theory will provide a complete explanation of the phenomenon in question, the observations in $Y$. Consequently, the tone of much discussion of theories and comparisons between theories involves assessment of variance explained. This is reflected in the two examples cited at the beginning of this paper.

**Comparing Explained Variance Between Theories**

Scholars including Lieberson (1985) and Cooper and Richardson (1986) have raised concerns about the appropriateness or fairness of such comparisons between theories. Spirited discussion about the relative contribution (in variance explained terms) of different variables occurs frequently in the study of problems in all the social sciences. Certainly arguments about nature versus nurture (e.g., disposition vs. situation, personality vs. situation) have taken this form. Since investigators often rely on variance explained to help decide which theory is ‘better,’ the concern with the fairness of comparisons is understandable. Cooper and Richardson’s (1986) analysis suggests that if fairness criteria (e.g., competitive theories were equally well operationalized, variables representing competitive theories both vary sufficiently, variables are measured with equal reliability and so on) are met, then such comparisons between theories are reasonable ways to decide which theory is a better explanation of the phenomenon. Others, like Duncan (1970) have suggested that comparisons of variance explained across models and/or theories are
not important, and are often detrimental to the basic scientific enterprise. Discussing path models, Duncan (1970: p. 46) observed that “In problems where this kind of system (path analysis, or a structural system of equations) is an appropriate model, the calculation of explained variances is often an irrelevant or at best a secondary objective.” A structural model (in the path analytic sense) can be invariant across populations while explaining different portions of the variance across those populations. Consequently, interpretation of the relative magnitudes of $R^2$ is often a discussion of population differences, not theoretical or model differences. For Duncan, $R^2$ helps tell us about the precision of estimates, which is important, but that is all (Duncan, 1975).

King (1986, 1990) has reiterated this point forcefully in critiquing the use of $R^2$ in political science (where the research methods, customs, and uses of $R^2$ are very similar to those in organizations research). King concludes that there is only one use for $R^2$, “comparing two equations with different explanatory variables and identical (emphasis added) dependent variables” (King, 1986; p. 677). With Duncan, King notes that changing variability in the independent variable across settings will change the estimated $R^2$ even if the relationships of the $X$ variables to the $Y$ variables are invariant. Most treatments of $R^2$ treat it as a method to determine which model is better, or closer to the ‘truth.’ King notes that there is a logical fallacy in such an interpretation; namely that one can not logically determine which model is ‘true’ given the data. The best one can do is determine which model is most likely to have generated the data in hand.

Lieberson (1985: 104) suggests that “any judgment about the ‘explanatory power’ of a theory . . . will be affected by the nature of the data set . . .
There are two equally valid and at least partially separate questions involved here. The first refers to whether a theory is correct. The second refers to whether it helps us understand a given phenomenon.” In the case of explaining the motions of leaves on a breezy autumn day, Galileo’s Law is still a correct theory but it does a poor job of explaining variability of movement of leaves to the ground. In a perfect vacuum, Galileo’s Law is correct (as before on a breezy day) but it now does a good job explaining (in a predictive or goodness of fit sense) the movement of leaves to the ground. Thus evaluating the contribution of a variable, or the relative contribution of one variable to the explanation of a phenomenon, is context dependent. Simon (1977) makes a similar argument from a similar example to forcefully argue that statistical tests are of little help in ultimately evaluating a theoretical model when applied to data. Lieberson (p. 106) concludes that

It is impossible to use empirical data in nonexperimental social research to evaluate the relative importance of one theory vs. another (or if you will, one independent variable vs. another.)

Buss (1989) in a discussion of the personality versus situation controversy in personality psychology, where arguments about relative amounts of variance explained by personality and situation abound, concludes that “Asking whether traits or manipulations (environment) control more variance is useless because researchers can plan paradigms that favor one or the other. (Buss, 1989: 1379)” However, his conclusion does not favor the argument proposed by Cooper and Richardson (1986) for ‘fair’ comparisons untainted by biases embedded in designs or introduced by investigators. Rather, Buss suggests that laboratory experiments on personality (which focus on estimat-
ing effects of environmental variation) and longitudinal studies of personality traits ask fundamentally different questions in different types of contexts. The experimentalist is usually focusing on environmental variables and letting traits vary. Consequently, under random assignment in a laboratory experiment, trait variance is treated as residual error variation. Personality trait students focus on a different question; what consistencies across occasions can be identified in individual behavior. Consequently, the focus is on person variance while variation in settings is now treated as residual error variance.

An observation about Buss’s argument should be made here. Buss suggests that how variance is partitioned is in part a function of the investigative context; he does not suggest that a particular theoretical construct operates differently across contexts (in this case research context), but that the answer you get to your question will vary across contexts, even if the underlying processes generating the variation due to traits and situations is constant. Second, Buss’s argument is applied to both experimental and nonexperimental research. Whereas Lieberson (1985) focuses his criticisms of the problems of explained variance on nonexperimental research designs (which is predominant in his discipline, sociology), Buss clearly suggests that the difficulties with the explained variance criterion apply to both experimentally controlled and nonexperimental research settings.

Buss concludes that such different methods of asking research questions are bound to get different answers, and the answers are not readily comparable with each other. He finds that studies focusing on situational or environmental effects are strikingly different from studies focusing on stable trait effects. In his characterization of the research, Buss characterizes the
research strategies favoring each type of outcome.

if one wishes to load the dice in favor of manipulations (i.e., situation effects), research should be conducted in novel, formal, public settings, with detailed instructions to subjects who are allowed little or no choice, in a brief period of study of narrowly defined responses. If one leans toward personality traits, research should be conducted in a familiar, informal, relatively private setting, with few instructions and maximal choice by subjects, whose broadly defined responses are studied over an extensive period of time. Thus, it appears futile to question whether manipulations or traits carry a greater share of the variance, for a researcher can stack the deck in favor of one or the other. (Buss, 1989: 1381)

The critical issue in Buss’s analysis is that since questions about situations and traits are often both conceptually and methodologically different, the comparison of answers to the two questions is misguided. Buss’s position implies that the Cooper and Richardson (1986) argument for fair comparisons is ill-conceived. Generally such comparisons cannot be made in a useful way. Some situations and measurement strategies favor situational effects (experimental strategies) while some research strategies favor person effects (longitudinal studies that examine person’s across a broad range of situations). Interestingly, Buss’s analysis is generally consistent with the results on person vs. situation effects in the social psychological and organizational literatures. Longitudinal designs have favored strong person effects, while short-term laboratory studies have favored situational effects.
These arguments can be pushed farther, and may lead to perverse consequences. For example, competition amongst theories is often judged using variance explained criteria. One possible outcome is that better (in the sense of explaining more variance) theories may be ones which are selecting on better contexts. Gerhart (1987) has made precisely this argument regarding the relative influence of dispositions vs. situations on job reactions. Staw and Ross (1985) found evidence that people’s job reactions were relatively invariant across job environments using the National Longitudinal Survey (NLS; a national sample study using a cohort design). Gerhart argued that Staw and Ross’s results were in part due to the use of the mature cohort of the NLS. He argued that mature workers (in this case in their 40’s and 50’s) had more fully developed their outlook and attitudes, and were less likely to respond to variation in jobs. On the other hand, younger workers who were more reactive to environments ought to show greater responsiveness to situational variation. Gerhart’s results support his argument.

Sears (1986) has made a similar argument in experimental social psychology, suggesting that the propensity of experimental social psychologists to use college students leads psychologists to wrongly conclude that environmental or situational influences swamp differences between individuals (dispositional responses, in Staw and Ross’s terms, or personality trait effects, in Buss’s terms). Sears argues from data that young people in college are precisely the ones who are most likely to exhibit high sensitivity to situational cues. In both these cases (Gerhart’s and Sears’s critiques), the arguments about the relative effects of variables are context dependent (in both the Gerhart and Sears analyses, the context variable is age), and do not have much force when considered apart from the particular context.
This reinforces Buss’s position that personality trait or situational effects will reliably vary by context.

This set of observations leads us to a first clear conclusion:

*Typically, the variance explained criterion, when used to contrast the relative theoretical value of different sets of constructs, is not a useful criterion for evaluating and comparing theories.*

**Evaluating Variance Explained**

In applied practice, explained variance can help the analyst evaluate the utility of a set of variables for prediction or as an indicator of treatment effectiveness in an evaluation context. In fact, much concern for the ‘weakness’ of social theories and interventions has arisen because of ‘low’ correlations in soft psychology (e.g. Meehl, 1978, 1990). Acceptance of Meehl’s argument that variance explained is an appropriate metric for evaluating theories can have detrimental consequences for applied practice. Investigators often are not well calibrated in their evaluation of effect size estimators and their applied value, underestimating the impact of treatments which have small effect sizes. Such miscalibration can mislead a practitioner in choice of appropriate treatment or intervention.

Rosenthal (1990) has convincingly argued that treatment benefits are frequently underestimated by looking at variance explained indicators like correlation coefficients, $\omega^2$ or other effect size estimators. Small amounts of variance explained can still make big practical differences, as Rosenthal

---

3 Meehl’s critiques are far broader than just the issue of small correlations, and far more than focusing on application alone. However, Meehl has clearly made this argument concerning the relatively small amounts of variance explained in ‘soft’ psychology.
Rosenthal and Rubin (Rosenthal & Rosnow, 1991; Rosenthal & Rubin, 1982) developed the BESD (binomial effect size display) so that investigators are not misled by variance explained estimators. The BESD is designed to show the “practical validity” (Rosenthal & Rosnow, 1991: 282) of a correlation coefficient. In the BESD, a correlational result is displayed as a dichotomous outcome, allowing for an intuitively accessible interpretation of an effect size. As an illustration, Rosenthal and Rosnow consider \( r = .32 \), which translates into an effect size estimator \( r^2 = .1024 \) or approximately 10% of the variance in the dependent variable. This certainly falls in Meehl’s neighborhood of soft psychology’s small effects. Yet, when displayed in a BESD as in Table 1, it clearly makes a difference.
Abelson (1985) has shown that one can overestimate the amount of variance explained. In an entertaining demonstration, he analyses the percentage of variance in baseball hitting performance that can be attributed to differences in skill levels of a baseball player for a single time at bat\textsuperscript{4}. Abelson’s informal poll of colleagues at Yale showed a median predicted value of 25\% for the percentage of variance in hitting due to differing skill levels for a single at bat. Abelson’s estimate based on data for 1983 is about .22\%!

\textsuperscript{4}Please pause for a moment and generate your own estimate of the variance attributable to skill for a single time at bat before continuing.
No, that is not a typographic error; estimates were off by approximately two orders of magnitude! Even taken over 1000 at bats, only 1.3% of the variance is attributable to skill differences when contrasting a poor (.220 hitter) and good (.320) hitter. This result clearly violates the intuitions of many baseball fans and managers. It suggests that our intuitions about variance explained and the value of a particular independent variable are not well calibrated with other estimators of the effect of an independent variable on a dependent variable. For baseball, Abelson argues that what is critical is that the value of hitting skill *cumulates* across time and over people. Hence a small advantage at one at bat can make a tremendous contribution over the course of a series of games and a season. In the context of predicting behavior from personality traits, Epstein (1979) makes precisely the same observation. Prediction from a single trait to a single behavior leads to low predictive power, but by aggregating behaviors across occasions, the predictability of behavior from personality traits increases. Given the low reliability of individual behaviors, this effect of aggregation is well understood and derivable from the Spearman-Brown prophecy formula in psychometrics. In the Abelson case, aggregating across people (batters) and occasions (games) gives increased predictability.

The Rosenthal and Rubin BESD example in Table 1 suggests that effects are ‘stronger’ than our intuitions about corresponding correlation coefficient estimators. The Abelson baseball example shows occasions where people overestimate effect sizes when expressed using proportions (batting averages) to infer effect size estimates. These two results are actually complementary in the sense that the BESD shows the translation from effect size estimates to hit rates, while the Abelson result illustrates the translation from hit
rates to effect size estimators. Each result is a transformation of the other. These illustrations by Rosenthal and his colleagues, and by Abelson suggest people do not do a good job using variance accounted to evaluate the applied utility of a particular variable or model.

These several observations lead to my second conclusion.

If the variance explained criterion is an appropriate evaluation criterion for model evaluation (as it may sometimes be in the evaluation of predictions and treatments in applied settings), it can be misapplied such that the conclusions about model or treatment quality are inconsistent across different effect size estimators.

There are several other relevant characteristics of variance explained estimators. First, it is not necessarily the case that explaining a great deal of variance is testimony to good theory. One can not blindly compare theories by such a criterion. A theory may be inelegant and unparsimonious, so one may prefer less variance explained if the price of explained variance is a huge host of variables. King, Keohane & Verba (1994) describe this criterion as a preference for leverage, explaining or accounting for a ‘lot’ of a phenomenon with a ‘little’ set of variables. Second, one can ‘overfit’ a set of data. Often, adjustments must be made for overfitting of equations in regression, for example. Techniques such as cross-validation on a hold out sample, bootstrap procedures for cross-validation (Efron, 1983) and reduction in $R^2$ formulae are employed to allow for chance variation explained. For these reasons as well, variance explained estimates and their interpretation should be approached with caution.
My arguments about the role of variance theories and their effects on research design and interpretation go deeper than misestimating effect sizes or setting up unfair comparisons of variables within theories or between theories. These issues are of importance if one accepts that explaining variance is a useful criterion to apply when building explanatory theory. Lieberson (1985) suggests explaining variance often is not useful, and frequently is inappropriate. Since his argument concerning variance theories is central to understanding the problems the variance approach can create for social science research, I present it in some detail and then extend it.

**Statistical Analysis and Variance Explained**

The primary tool kit of quantitative social scientists is statistics, and central to the statistical enterprise is the analysis of variation. Variation is the necessary condition for much statistical analysis. Without variation, there is no need for a statistical characterization of an empirical phenomenon. Variation is the defining feature of any statistic. Singer and Marini (1987) note that statistics largely grew out of genetics, where the analysis of variation is a central substantive question. In social research, accounting for variation is not necessarily the central substantive question (Lieberson, 1985). Rather, I may be interested in determining if a particular theoretical effect can be observed as a critical test of the difference between two theories. In such cases, the explanation of variance is irrelevant. The existence of an effect, not its magnitude, is what is at issue (Berkowitz & Donnerstein, 1982; Mook, 1983; Prentice & Miller, 1992).

Lieberson suggests that under certain circumstances, the pursuit of variance obscures critical theoretical constructs, and the pursuit of variance per
is an inappropriate research strategy. His example is striking (pun intended), so it is offered here for illustration. The study is focused on falling objects in a natural setting.

...suppose we visualize a study in which a variety of objects is dropped without the benefit of such a strong control as a vacuum – just as would occur in nonexperimental social research. If social researchers find that the objects differ in the time that they take to reach the ground, typically they will want to know what characteristics determine these differences. Characteristics of the objects such as their density and shape will affect speed of the fall when there is not a vacuum. If the social researcher is fortunate, such factors together will fully account for all of the differences among the objects in the velocity of their fall. If so, the social researcher will be very happy because all of the variation between objects will be accounted for. The investigator, applying standard social research thinking, will conclude that there is a complete understanding of the phenomenon because all differences among the objects under study have been accounted for.

Something must be wrong if social research methods would lead the investigator to believe he had a full grasp of falling objects without ever invoking gravity.5

---

5Gravity is a very interesting phenomenon for illustrating statistical issues in social science. I note Simon and Lieberson’s related use of the example. Wahlsten (1990) uses gravity to illustrate the insensitivity of ANOVA to interactions. Gravity is a useful example for serious issues.
Clearly the problem here, as Lieberson rightly notes, is that the researcher is analyzing “difference(s) in velocity” (Lieberson: 103), not velocity and gravity, which causes the movement itself. The question is misspecified theoretically, but the tools of variance analysis lead to an explanation of the variance but not of the phenomenon! When the causal force is constant (e.g., gravity), variance partitioning methods are ill-suited to the problem.

What is going on here is that the conditions under which free falling bodies are studied diverge from ideal conditions. Simon’s (1977) discussion of theory building and judgments of plausibility makes the point forcefully. Simon does not want to confuse deviations from a simple lawful generalization with the critical theoretical insight; in this case gravity. “When a physicist finds that the ‘facts’ summarized by a simple, powerful generalization do not fit the data exactly, his first reaction is not to throw away the generalization, or even to complicate it by incorporating additional terms. When the data depart from $s = \frac{1}{2}gt^2$, the physicist is not usually tempted to add a cubic term to the equation... What the physicist must learn through his explorations is that as he decreases the air pressure on the falling body, the deviations from the law decrease in magnitude.” (Simon, 1977: 27-28) Simon explicitly rejects the use of statistical testing in such situations, since significant variations will be observed in any but ideal circumstances (a perfect vacuum), given sufficiently precise estimates and numbers of observations. Simon’s physicist wants to prune away such deviations (ignoring variance) as irrelevant, while Lieberson’s social researcher embraces and seeks out precisely those deviations (seeking variance to explain). This example clearly illustrates one critical aspect of the variance seeking problem.

Lieberson’s and Simon’s discussion’s of gravity identify two related prob-
lems. One is the inappropriateness of variance oriented thinking and analysis when the fundamental process does not imply such variation. Gravity is a physical constant in this example, and only the deviations due to air pressure and other departures from a true vacuum lead to variations\textsuperscript{6}. Such variations are not random; they have systematic components. To understand gravity, I do need to determine the impact of such deviations in order to control them. However, I am left with an invariant process; a process that variation seeking tools and habits of mind will not detect. Second, the pursuit of variance will lead to a focus on the deviations and the statistical modeling of the variations, rather than the fundamental process. Finally, if the deviations are large, the fundamental process will be judged to be a weak theory. That is, Galileo’s Law will not provide much help in explaining variation in the behavior of falling bodies, and will be judged a poor theory. The appropriate characterization is Simon’s; that Galileo’s Law is a good generalization, since its performance approaches the ideal case as the deviations from the ideal case approach zero. However, a metric of variance explained would be wholly inappropriate for the evaluation of Galileo’s Law.

Lieberson’s gravity example forcefully suggests that investigators may focus on variation in observations while neglecting to understand the process itself. Thus, if I use a variance approach to study variation in status attainment in society, I may think I understand status attainment, but my analysis does not provide insight into status attainment, but into differences in status attainment. This is directly analogous to the variance analysis of gravity, where the variation in gravity induced velocity is understood, but

\textsuperscript{6}I ignore the mass of the falling object and the earth and the inverse square law of gravitation because its effects here are incidental to the central issue in the argument.
Given that

1. problems are posed using variance theories,
2. variation is necessary for many types of statistical analysis,
3. statistical analysis is often necessary to conduct quantitative social research,

problems will be framed and addressed in terms of variance and explaining variance. This is in fact the case. Lieberson suggests that social scientists have lost contact with problems by focusing on them through a statistical lens. This statistical lens causes social scientists to select for variance-framed theories and this lens filters out all but variance theories. This is a version of Kaplan’s (1964) boy with a hammer, who, once given a hammer, hammers everything in sight. One could also argue that variance-oriented tools will cause us to build theories which reflect variational concepts, as Gigerenzer (1991) shows in the case of the development of causal attributional reasoning which treats the perceiver as calculating an intuitive analysis of variance.

To this point, I have focused on difficulties with the the variance explained criterion such as its limited utility, many real problems in effectively applying it, and that it is often incidental to the goal of explanation and theory building. Investigators frequently can not evaluate variance explained estimators well, can be misled by the effects of contexts on variance explained estimators, and by the vagaries of estimation procedures which can be Procrustean, overfitting models to data.

However, I have tacitly accepted the argument that given a context spe-
cific focus, sound design, and statistical analysis, one might be able to make limited statements about variance explained and the quality of explanation. To do this, I have assumed that the model(s) to be tested are *complete* explanations of the phenomenon to be explained. If so, appropriately specified models of such complete explanations will explain all the variation in the phenomenon of interest. I now turn to this assumption.

**Variance Explained and Determinism**

In the discussion to this point, I have taken as given the assumption of determinacy. The statistical analysis that seeks to account for variance in elements of the $Y$ matrix of dependent variables rests on an assumption that variation in $Y$ is completely determined by some set of independent variables $X$. Otherwise, the calculation of proportion of variance explained is undefined. The traditional index of variance explained in ordinary least squares (OLS) regression is $R^2$, with a maximum of 1.0. Consequently, unexplained variance is expressed as $1 - R^2$. The assumption is that a complete model would account for all the variation in the $Y$ matrix of outcome variable(s) to be explained. As Hamblin, Jacobsen & Miller (1973) suggest

> The rule of determinacy prescribes the assumptions that natural phenomena are completely determined or very nearly so, and that it is always possible to find a correct answer – an equation that accurately describes or explains almost all the variance in the data relationships under investigation. It implies that scientists should push their investigations until they find an equation or set of equations serving to explain nearly 100 percent of the variance for the phenomena in question. (quoted in Lieberson,
The origin of such arguments is with LaPlace (1951), who suggested the following argument.

Given for one instance an intelligence which could comprehend all of the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit all these data to analysis—it would embrace in the same formula the movement of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes. (quoted in Salmon, 1984: 15)

Subscribing to the application of the variance explained criterion is equivalent to assuming LaPlacian determinism. Is it reasonable to accept the assumption of LaPlacian determinism? That is, assume

1. that all the technical and algorithmic problems (overfitting of linear models, correlated measurement error, and so on) which make us uncertain that we have explained all the variance in a set of observations are solved, and

2. agreement on how to tradeoff parsimony against explained variance to find an optimal number of variables such that additional variables are of no interest although they explain more variance.

For a given problem, I can now evaluate whether all, or nearly all, the variance is explained without recourse to extratheoretical arguments and
assumptions. Is the assumption that all the variance is determined an appropriate one? There are many situations when it is “reasonable to expect a perfectly complete theory to account for less than all of the observed variation between units in the dependent variable of interest” (Lieberson, 1985: 93). Put differently, what if we take seriously the idea that there is variability that is not due to errors of measurement and other errors associated with the error term $\epsilon$ in OLS regression or its equivalent in other contexts. Rather, the random variability is intrinsic to the mechanism (Salmon, 1984) or process which generates the social phenomenon of interest. Such theories are termed stochastic process models, or process theories, in Mohr’s (1982) terminology.

**Process Models**

Stochastic process models are widely used in the sciences, including the social sciences. Darwinian evolution, diffusion of innovation, population and labor mobility, models of epidemics, birth and death processes in people and organizations, decision models, wars, learning and so on have all been modeled using stochastic process models (Bartholomew, 1982; Lave & March, 1975; Mohr, 1982). Neyman (1960) noted the rise of such stochastic process theories, and distinguished them from work in the Pearson and Fisher statistical traditions. Neyman suggested that Pearsonian and Fisherian statistics, which form the intellectual basis for much current statistical practice in the social sciences, are characterized by “static indeterminism.” Such “statistical methodology . . . was not designed for the study of chance mechanisms behind the phenomena developing in time and space” (Neyman, 1960: 629).
He describes stochastic process models as instances of “dynamic indeterminism.”

Mohr defines process theories in contrast with variance theories. “A process theory is one that tells a little story about how something comes about, but in order to qualify as a theoretical explanation of recurrent behavior, the manner of the storytelling must conform to narrow specifications” (Mohr, 1982: 44). My reading of his definition of a process theory is as follows. The essential specifications are these:

1. X is a necessary condition for Y, but not a sufficient condition

2. X will cause Y stochastically. That is, whether X causes Y depends on a draw from some probabilistic process. For example, Mohr describes a malaria epidemic process model. The presence of a malarial parasite in an anopheles mosquito, acquired by biting someone with the parasite, is the necessary precondition. However, whether one gets malaria or not depends on whether one is bitten or not by an infected mosquito, which reflects the confluence of 2 probabilistic processes (the mosquito having bitten someone with malaria, and then that infected mosquito biting someone else). Therefore, the operation of X requires a lawful probabilistic process as a necessary condition. It is non-deterministic because there is no necessity that a mosquito bite a particular person or that such a bite will necessarily lead to malaria.

Thus, a process theory stands in stark contrast to a variance theory with respect to variance explained. By definition, the malaria process theory above can not provide a deterministic explanation of malaria. This theory can not account with certainty for whether a particular individual
will become infected. The random process that drives the process theory ‘story’ by moving people stochastically from one discrete state (health) to another state (malaria) insures that I cannot know with certainty who will be sick and who will be well. I can then say with some precision, where precision is jointly determined by our sample size, measurement procedures, research design, and the quality of our theoretical model, what will be the relative frequency of sick and healthy people observed. The lawful quality of the probabilistic processes that drive the model allow us to do that, but that is all. We cannot determine the outcome $y_i$, given $x_{ij}, x \in X$ for any given individual $i$ in some state $j$ with certainty. As a result, process theories provide a mechanism that insures, by construction, that all the variance in $Y$ will not be explained.

A partisan of variance theories might ask, why ceteris paribus prefer a process theory that leaves residual uncertainty by construction to a variance theory that, when successful, provides a ‘complete’ explanation. There are two kinds of answers to this question. In order to make the issue concrete, consider an interesting process model identified by Lieberson.

**A Stochastic Model of Racial Disturbances**

Spilerman (1970) proposed a stochastic process model of the outbreak of racial disturbances in cities in the United States in the 1960’s. He distinguished two kinds of variables; “underlying causes and immediate precipitants of racial disturbances” (Spilerman, 1970: 628) in his model. Spilerman suggested that underlying causes influence differences in disorder-proneness (a necessary but not sufficient condition for a disorder) and that immediate precipitants were random processes. Therefore, two cities with equal values
on disorder proneness would not necessarily both have disorders, and there is no value of disorder proneness that is both necessary and sufficient for a disorder to occur. This fits the definition of a process model proposed by Mohr.

Spilerman (1970: 627) identifies “what general assumptions must be met by any satisfactory explanation of the distribution of (racial) disorders.” He then evaluates proposed solutions against his criteria. He first assumes there are no differences in underlying causes, and that racial disturbances are purely random, specified as a Poisson process. This implies no differences between cities in disorder-proneness, no change in the rate of disorders over time, and that disorders are independent events. Clearly, these are very restrictive assumptions. The observed data fail to fit the expected Poisson distribution. Spilerman then relaxes different assumptions, and finds that the best set of assumptions are that cities vary in disorder proneness (heterogeneity of disorder proneness), and that this heterogeneity is constant over time. Further, he finds little evidence of geographic contagion (where disturbances in one city increase the likelihood of disturbance in geographically nearby cities) or reinforcement effects (having a disturbance increases (positive reinforcement) or decreases (negative reinforcement)) the likelihood of a disorder in the future. He concludes that disorder proneness is a stable characteristic of a city. He then identifies location (North vs. South) and population size (number of blacks in the city) as explaining variance in disorder-proneness. Beyond these two variables, other proposed explanations offered in the literature explain little additional variance. 46% of the variance is explained by location and population size, and 54% of the variance is unexplained. Spilerman’s framework does not account for all the
variability in the dependent variable, occurrence of a racial disorder.

A LaPlacian view is that this process model is an analytic convenience that uses random processes as an alias for ignorance. If more were known, Spilerman could reduce the magnitude of $1 - R^2$, and drop this stochastic framework. Spilerman (1971) asks if the Spilerman (1970) conclusions are a complete explanation of riots. Does the process model specify a plausible process where social life has ‘true’ stochastic components, or does this process model ‘finesse’ ignorance? Spilerman (1971: 433) uses an analogy with quantum physics.

a process may be conceptualized as inherently stochastic, and this type of explanatory model is consistent with there being a complete explanation (in the sense that all relevant knowledge has been incorporated) in which the $R^2$ value is less than one.

Spilerman notes that in modern physics (see Salmon (1984) for a philosophical treatment of this issue) radioactive decay is fundamentally a stochastic process. The rate of decay of an element is known with a high degree of certainty, but which atoms will exhibit decay or the exact number of atoms decaying in a specified unit of time is not known. Thus, “if, after the fact, the number of disintegrations is regressed against mass (the observations being different sized pieces of the same radioactive substance), the resulting value of $R^2$ will be less than one” (Spilerman, 1971: 433).

Spilerman (1971) takes a disorder proneness measure, size of black population, and simulates the data by region as if the only explanatory variable was size of the black population. Then one can see how much variance there is to be explained in such a simulation and compare it to the actual vari-
ance explained. If the amount of variance to be explained in the simulation is approximately equal to the actual explained variance in the data, then there is good reason to believe that the stochastic model provides a full and complete explanation of the data, even though $R^2 < 1$. Spilerman finds that for the North, nearly 98% of the variance that is theoretically specified can be explained, given the simulation results. For the South, only 72% of the variance to be explained is accounted for by the stochastic model. Thus, Spilerman concluded that his model is a full explanation of the variance in disorders for the North, while there is variation to be explained for disorders in the South. The result for the North is the important one for our purposes. Spilerman has shown that a full explanation for the distribution of the racial disturbances in the North is provided by a model that accounts for only 73% of the variance in the dependent variable\(^7\). This is in stark contrast to the usual reaction to $R^2 < 1$; to wit, more research is needed. Spilerman’s analysis is of interest because he explicitly considers completeness of explanation in the context of a process model.

**A Stochastic Model of Individual Behavior: Absence and Attendance**

Fichman (1988) developed a similar conceptual and statistical argument in a study of the time dynamics of individual attendance and absence in an organization. Fichman argued that absence reflects the dynamic operation of a set of time-varying motives. To explain the timing of absence and attendance, one must consider the changing strength of motives to attend

\(^7\)In Spilerman (1970), 48% of the variance was explained across both the North and South. The 73% reflects variation in the North, an estimate that was not discussed in the presentation of the Spilerman (1970) results above.
work and motives to engage in activities which require absence from work. Following Atkinson and Birch (1970), he argued that unfulfilled motives increase in strength with time, and that this changing motive strength can be modeled as a set of differential equations. Given these assumptions, if all motives were internal, there were no external constraints on time allocation, and a person could act on their motives instantly without cost, then one could construct a deterministic model of time allocation and fully explain the timing and duration of activities. However, no realistic model of time allocation can take as given the assumptions of no external constraints. Some external constraints are stochastic, such as work stoppages, accidents and illness. The stochastic external constraints on time allocation must be considered in a realistic model of time allocation. Such stochastic shocks will cause the observed pattern of time allocation to deviate from the theoretically specified pattern, even if the theory is correct. Thus, these stochastic elements assure us that there will be ‘residual’ variation in time allocation to be explained.

Consider the following example. With no external constraints, Person A want st o go to work on Wednesday. A’s motive to work will decline such that by Friday, they will want to take a day off. If I measure motive strengths and the rates of change of these strengths, apply the Atkinson and Birch model, and this model provides a full explanation of time allocation, then I will fully account for time allocation. However, suppose the plant is shut down due to an unforeseen failure to receive sufficient supplies to operate. Then, A will be forced to engage in a less preferred activity, and the strengths of all the various motives will change. The result will be that the previously predicted pattern of absence will now not obtain, and an alternative time allocation
pattern will evolve due to the new pattern of motive strengths. To test such a model, Fichman (1988) used a strategy similar to Spilerman’s.

Absences were classified as voluntarily controlled or involuntarily controlled (where involuntarily controlled means the absence was taken due to a random exogenous event such as an illness or accident). Then, the attendance event histories of 60 people were simulated, using the allocation of voluntarily controlled absences fixed at the points in time where they were actually allocated. Then, involuntary, random exogenous absences were randomly interpolated into the attendance event history. The simulations suggested that (1) such a random interpolation of absences generated simulated patterns very similar to those found in the actual data but (2) several of the parameter estimates were different from actual estimates from the sample, suggesting “that the random interpolation of involuntary absences (as random events) does not fully capture the complexities of the attendance process.” (Fichman, 1988: 131). Fichman’s (1988) results and theory suggest that while there are systematic causes (voluntary motives and their dynamics which can be expressed deterministically) analogous to Spilerman’s underlying causes, there are also random events which interact with systematic causes such that the systematic causes can not fully determine the pattern of time allocation. Note that the simulation results might have suggested that the observed patterns and the simulated patterns do not differ, and the stochastic model is a complete explanation of the pattern of time allocation.
Stochastic Events and Explanation

Bandura (1982, 1986) made a similar, more general observation about stochastic events. He suggests “that chance encounters play a prominent role in shaping the course of human lives. In a chance encounter the separate chains of events have their own causal determinants, but their intersection occurs fortuitously rather than through deliberate plan.8 ” (Bandura, 1982: 747) Even for theories which provide complete explanations of the determination of some domain of behaviors, stochastic processes unrelated to the explanatory variables can play a significant role in determining events. Consequently, in a set of data, anomalous outcomes which diverge from predictions will be observed. These anomalous outcomes are not due to poor measurement or errors which, through proper control and rigorous methods, would be reduced to zero. Rather these anomalous outcomes are fundamentally stochastic. Even if the consequences of a random event can be fully explained, the random event stands outside the explanatory domain of the theory and consequently cannot be considered in any way part of the theory. This means that a good process theory will inevitably have some residual variation unaccounted for in the set of dependent variables the theory is intended to explain.

Similarly, Spilerman’s theory of racial disorders, by admitting stochas-

---

8 This is similar to but distinct from the idea of path dependence in non-linear or dynamic systems. In such non-linear systems, initial conditions determine the path of some process. This is sometimes referred to as the ‘butterfly effect.’ The argument is similar in that variation in initial conditions will lead to a different set of states. The argument here differs from a path dependent non-linear systems argument since such initial conditions are determined stochastically in my argument, while in some types of path dependent non-linear dynamic systems, there is no stochastic component, but the non-linearity may lead to the appearance of randomness.
tic events, does not suggest that one could not explain how a particular stochastic event contributed to a racial disorder. It is just that the stochastic event is probabilistic, and its occurrence cannot be determined. Thus I might obtain records of a racial disorder and find that a police officer shot a young man in broad daylight on a crowded street on a hot, humid summer day. Such a shooting may have precipitated the racial disorder. Further, we might even agree with the counterfactual statement that if the shooting had not occurred and it had rained that day, the racial disorder would not have occurred. In principle, I may find as I try to explain an individual instance of rioting, that there is some point at which cases in all ways identical may still not consistently show one pattern of behavior. This is the essential limitation of indeterminate systems, which was first identified in modern physics in quantum mechanics. In quantum mechanics, some electrons are observed to tunnel through a surface, while other electrons are reflected back. Quantum mechanics provides an explanation of the relative probabilities of tunneling and reflection, but does not explain why one particular electron tunneled rather than was reflected back (Kitcher, 1989). This is irreducibly indeterminate. In either the quantum tunneling or the racial disorder case, stochastic indeterminacy is a deep and difficult problem for causal explanation. Assume that there are two cases which are identical in all causally relevant respects and no other causally relevant respects on which they can be differentiated exist. Yet one case can show a different outcome than the other on a dependent measure for which the causal explanation is intended to account. In the physics realm, it may be that two identical electrons are observed; one tunnels through a barrier and the other does not. In the sociological realm, two cities may have identical disorder
proneness scores, but one has a disorder and the other does not.

There are at least two different senses in which this indeterminacy can limit explanation. In the quantum mechanics case, there is irreducible indeterminacy; the clearest contrast with LaPlacian determinism. In the riot proneness case, I may be able to ‘more’ fully explain a riot ex post by looking at the particular occurrences which occurred in that city that day or week. However, I cannot provide this fuller explanation ex ante because the precipitating event is stochastic. Thus, in the quantum mechanics case, I will not be able to provide a fuller explanation after the tunneling of an electron. I may be able to provide a fuller explanation in the rioting case. These are two different senses in which indeterminacy limits explanation. Note however that in both cases there is an informative ex ante explanation, even if the informative full explanation does not reduce indeterminancy to 0.

Scriven (1959) discusses an instance where one can fully and satisfactorily explain ex post what one could not predict ex ante; evolutionary theory. Evolutionary theory explicitly considers the role chance events play in the changing distribution of life forms over time under changing environmental conditions. As Scriven puts it, “one cannot regard explanations as unsatisfactory when they do not contain laws, or when they are not such as to enable the event in question to have been predicted” (Scriven, 1959: 477). I can take Scriven’s definition of lawful prediction to be the case of deterministic, variance theory. The expectation is that with sufficiently good information and theory, all the variability in \( Y \) can be explained (and consequently predict the value of each entry in \( Y \)). It is precisely the absence of lawful prediction, and its replacement with ex post probabilistic explanation with no necessary predictive ability, that was (and continues to be) the
revolutionary feature of Darwin’s theory.

The revolutionary quality of Darwin’s thinking is made manifest in the attacks upon Darwin for not having followed the simple prescribed model that was supposed to be the accepted way of doing science. In order to see the extent to which Darwinian evolution by natural selection represented a departure from traditional norms of scientific thought, . . . one has only to take account of the fact that Darwinian evolution is nonpredictive, but nevertheless causal. (Cohen, 1985: 294)

Not coincidentally, Neyman’s characterization of the first models exhibiting ‘dynamic indeterminism’ are precisely models which are evolutionary, such as Darwinian evolution and Mendelian genetics. 9

Evolutionary theory is concerned with the distribution of life forms over time, which are differentially selected by changing environmental circumstances over time. Further, these environmental circumstances are exerting selection pressures on life forms whose characteristics are randomly mutating. 10 Clearly, random mutations are by definition unpredictable. With random variation in life forms and environmental pressures, life forms are under “statistical pressure” from the environment (Scriven, 1959: 479). Clearly, evolutionary theory is a process theory. As Scriven (1959: 479) suggests by analogy, “We can explain the unlikely outcomes of partially random pro-

---

9Interestingly, it is precisely because Mendel posited a stochastic process model that Fisher was able to demonstrate that Mendel fudged his data. Fisher showed that, given the stochastic process model implied by Mendelian genetics, Mendel’s reported results were ‘too good’; very unlikely given the stochastic process generating the frequency distributions Mendel reported (Freedman, Pisani, Purves & Adhikari (1991: 417).

10There is some controversy about whether one can have directed mutation, but recent work suggests that such directed mutation is unlikely (Beardsley, 1997).
cesses, though we cannot predict them. We are not hard put to explain that a man’s death was due to his being struck by an automobile, even when we could not have predicted the event. Now this kind of case does admit of hypothetical probability prediction.”

Scriven makes an even stronger point, that there will be occasions when an event can be explained even when it can not be predicted at all! That is, there are some events can be explained \textit{a posteriori} that we would not predict \textit{a priori}. Salmon (1984) raises a related point; low probability events can require explanation. Scriven (Scriven, 1959: 480) provides the following example (also discussed by Salmon (Salmon, 1984:31-32)).

Here, we can explain but not predict, whenever we have a proposition of the form “The only cause of \(X\) is \(A\)” (I) – for example, “The only cause of paresis is syphilis.” Notice that this is perfectly compatible with the statement that \(A\) is often not followed by \(X\)–in fact, very few syphilitics develop paresis. Hence, when \(A\) is observed, we can predict that \(X\) is more likely to occur than without \(A\), but still extremely unlikely. So we must, on the evidence, still predict that it will not occur. But if it does, we can appeal to (I) to provide and guarantee our explanation.

Scriven is arguing here that, having accepted the idea of probabilistic explanation, instances will occur where even the probabilistic explanation can not be readily indexed by the ability to accurately predict or postdict (after the fact) the fate of individual cases. Consequently, an analog to \(\sigma^2\) such as predictive or postdictive accuracy to indicate explanatory adequacy.
will not do. Salmon (1984) explores this issue in greater depth, and deals with criticisms of Scriven’s argument.

One criticism of the Scriven position is that such low probability events are so characterized because the explanations offered are inadequate and incomplete. That is, the low probability of paresis given \( A \) just suggests the need for more and better terms such that consideration of these additional terms with yield high accuracy of explanation of paresis. Salmon suggests this is in principle not acceptable. Some events are low probability, and will not be rescued by the addition of other explanatory terms. Salmon (1984) provides the following example.

Suppose, for example, that two individuals, Sally Smith and John Jones, both commit suicide. Using our best psychological theories, and summoning all available relevant information about both persons... we find that there is a low probability that Sally Smith would commit suicide, and a high probability that John Jones would commit do so. This does not mean that the explanation of Jones’s suicide is better than that of Smith’s, for exactly the same theories and relevant factors have to be taken into account in both (Salmon, 1984: 88)

A similar argument can be made in a more general form, following Salmon. Assume a probabilistic process \( \Theta \) that generates a set of outcomes \( \alpha, \beta \) such that \( p(\alpha) = .7 \) and \( p(\beta) = 1 - p(\alpha) \); \( p(\beta) = .3 \). Any explanation of \( \alpha \) will invoke \( \Theta \) as an explanation. By the criterion of predictive or postdictive accuracy, \( \Theta \) is a good explanation of \( \alpha \). However, if \( \Theta \) explains \( \alpha \), it must also be explaining \( \beta \). Therefore, Salmon (1984: 88) concludes
“that the degree of probability assigned to an occurrence in virtue of the explanatory facts is not the primary index of the value of the explanation.”

This analysis of explanation and process theories suggests several conclusions. In the domain of process theories, the notion of complete explanation is quite different than in the domain of deterministic theories. By construction, process theories will present unexplained residual variability. The amount of residual variability may not provide a good indication of the process theory quality. Strategies and criteria such as Spilerman’s (1971) and Simon’s (1977) will improve the assessment of the quality of explanation. However, as Scriven and Salmon show, a statistical criterion such as predictability may not be useful for assessing the quality of an explanation. There will be stochastic causal processes which account for infrequent events. Nevertheless, a good explanation of a rare event may be complete, although the event was not predictable \textit{ex ante}.

\textbf{Implications for Empirical Social Research and Theory Building}

Several interrelated arguments have been proposed.

1. Goodness of fit criteria such as proportion of variance explained are problematic. People do not know how to assess and evaluate such criteria well. Such criteria are context dependent and do not provide a useful guide for appraising theory quality.

2. If a particular social phenomenon has intrinsically stochastic elements such that it is more appropriately treated using stochastic process modeling, then goodness of fit criteria such as variance explained can
mislead and frustrate investigators. Investigators will be misled because the variance explained criterion, when applied to a stochastic process phenomenon, will persistently signal to investigators that their explanation is incomplete since there is unexplained variance. The investigator may undertake new studies, which will prove fruitless if the underlying process is stochastic. As Spilerman (1970, 1971) and Fichman (1988) have suggested, a stochastic process explanation can be examined to see if it is providing a complete explanation of the phenomenon. Thus, in the case of racial disorders, Spilerman’s results suggest that his model essentially is a complete explanation of racial disorders in the North. Such a conclusion is quite encouraging, when contrasted with the frequent pessimistic analyses of the quality of social research (Meehl, 1978, 1990). Meehl’s pessimistic analysis may require reappraisal given the arguments and results discussed above.

These arguments, when applied to some research areas and programs, raise interesting questions. Some examples are provided in order to illustrate the generality of the issues and applications.

**Meta-Analysis**

Meta-analysis is a set of statistical techniques developed to integrate research findings across studies. Meta-analysis is now a widely used method for integrating studies in areas as diverse as organizational behavior, education, psychology, medicine, and epidemiology. These techniques have had a substantial impact in many fields (particularly behavioral research and medical research) and provided insights into often diverse and confusing literatures. Meta-analyses are now sufficiently common in behavioral research
that a meta-analytic review of 302 meta-analyses has been published to assess the overall impact of psychological and behavioral treatments (Lipsey & Wilson, 1993)\textsuperscript{11}. One premise underlying all meta-analysis is that the process generating the observed distribution of outcomes across studies is deterministic. Meta-analysts try to identify sources of error and restriction of variation across studies. The meta-analyst then tries to estimate the proportion of variation in outcomes attributable to such factors as transcription errors, unreliability, range restriction and so on. This is done so the analyst can determine if there is any residual variability to be explained after accounting for such sources of error and variance restriction. This is the motivation behind procedures developed by Hunter and Schmidt (1990). Other procedures, while having different objectives (Bangert-Drowns, 1986), also share similar assumptions. The premise in meta-analysis is precisely that the underlying process generating the distribution of observed results is deterministic. For example, a meta-analysis of studies of riot proneness in the 1960’s, assuming Spilerman’s results are correct, could yield a population correlation estimate much less than 1.00, suggesting the need for more research to uncover additional explanatory variables.

Suppose this assumption were not true, and the process generating observed outcomes were stochastic. Suppose the Hunter and Schmidt procedures are applied to a process such as that described by Spilerman, where less than half the variation is determined by the identified predictors (if modeled as a deterministic process), then we would probably conclude that there are other explanatory variables which have not been identified. The

\textsuperscript{11}The meta-analysis result is that these behavioral, psychological and educational interventions have a “strong, dramatic pattern of positive overall effects.(Lipsey & Wilson, p. 1181)”
meta-analyst will reach a conclusion that is similar to the conclusion reached by analysts of single studies where significant portions of variation are unexplained; that there is a need to consider additional predictors. However, such a search in a stochastic system may be fruitless, and lead to one of two possible outcomes. The first possibility is continued frustration as additional studies fail to provide additional insight into the problem. The second possible outcome is that a study, through capitalization on chance variation, overfitting of a model to data, or the particular context and design of the study, accounts for more variance. This study, having won the variance explained race, may identify additional variable(s) which become candidates for investigation. In effect, such a scenario can lead to the triumph (temporary, one would hope) of an incorrect model over a more correct model.

One additional observation about meta-analysis should be made. Meta-analysis has evolved in part as a response to the neglect of statistical power and attention to effect sizes (Bangert-Drowns, 1986; Cohen, 1962; Glass, McGaw, & Smith, 1981; Hunter & Schmidt, 1990). These criticisms of low power and inattention to effect sizes were and are warranted. When evaluating the effects of class size on educational achievement (Glass, Cahen, Smith & Filby, 1982) or other social interventions (Lipsey & Wilson, 1993), effect size is an essential issue. However, the argument raised above suggests that if the social processes generating the outcome of interest are stochastic, then the estimate of unexplained variance may be biased upwards. Further, small effect sizes may lead investigators to pursue new theory, when there is far less systematic variation out there to be explained than is expected assuming deterministic social processes.

Clearly, in such cases, techniques such as those developed by Spilerman
(1971) can give investigators better insight as to whether a relatively complete explanation has in fact been put forward (assuming a stochastic specification is sound). This ought to lead to more effective resource allocation and less pursuit of false positive findings.

**Debating Nature vs. Nurture**

In organizational behavior, great energy is periodically devoted to debating variants of the nature vs. nurture argument. Most recently, the debate has been over whether individual affective responses to organizational life (attitudes) are largely dispositional (a function of the individual’s characteristics and invariant across settings) or situational (a function of variation in organizational characteristics) (Davis-Blake & Pfeffer, 1989; Gerhart, 1987; Staw & Ross, 1985). Studies of identical and fraternal twins raised apart have added much interesting information and fueled this debate (Arvey, Bouchard, Segal, & Abraham, 1989). Similar issues are raised in other social sciences. In development and life course dynamics, this issue is actively studied, such as whether intelligence is environmentally or genetically determined (e.g., Devlin et al., 1997). Controversy has often emerged in psychology concerning the relative impact of personality and the situation; one visible outbreak of controversy was prompted by Mischel’s (1968) critique of the impact of personality on behavior (see Ross & Nisbett (1991) for a review of the issues).

These variants of nature vs. nurture or person vs. environment arguments share certain fundamental premises. First, the proportion of variability explained can be considered independent of context (however, more recent work in the effects of personality vs. the situation shows far greater
sensitivity to this issue (e.g., Buss (1989)). It has been argued here and by others (e.g., Buss, 1989; Duncan, 1975; Lieberson, 1985) that the proportion of variance attributable to person effects (or situation effects, the issue is the same) is not invariant but depends on characteristics of the measurement, the research design as well as characteristics of the sample. Therefore, questions about relative proportions of variance explained in these types of debates are probably not useful.

Davis-Blake and Pfeffer (1989), in their critique of the organizational behavior variant of this work, identify important problems in this type of research. Much of their specific criticisms of efforts to partition variance between dispositions and situations are sound. For example, if there are omitted variables, this can lead to misattribution of variation to other identified variables. Davis-Blake and Pfeffer (1989) suggest that one cannot accurately estimate the variation attributable to dispositions and situations without a fully specified model. However, they do not recognize the deeper problem that even a fully specified model, invariant across populations, can have variation in $R^2$ across populations. Even with a fully specified model, such estimates will not be invariant. Consequently, I would disagree with Davis-Blake and Pfeffer’s premise that such attempts at variance partitioning are useful, even if done properly.

In my discussion of this version of nature vs. nurture, I have accepted the assumption made in this literature that the dependent variable is embedded in a deterministic system of relations, that when fully specified, will lead to a full accounting of variation due to nature or nurture.
A Stochastic Approach to Nature vs. Nurture

I suggest that it is worth considering how a process theory approach, where variation in attitudes (the focus in organizational behavior) or behavior (the focus in personality and social psychology) is not fully determined by independent variable(s) like personality and situation would work. There may be residual variation not accounted for by the independent variables that is essentially indeterminate. This follows Bandura’s suggestion “that chance encounters play a prominent role in shaping the course of human lives.” (Bandura, 1982: 747).

I will elaborate this argument, using the process theory approach of systematically analyzing the implications of Bandura’s little story of chance encounters. For example, I could model Bandura’s notion using a Darwinian selection model. This is similar to what Lave and March (1975) call an adaptation model. Suppose there are situations $S$ which arrive stochastically that select for (e.g., reward) or against (e.g., punish) some particular disposition $d$. By select for, I mean that $s_i, s_i \in S$, increases the probability of the predisposition being enacted behaviorally. By select against, I mean that $s_i, s_i \in S$ decreases the probability of the disposition $d$ being enacted behaviorally. This is a plausible model. Gazzaniga (1992) has suggested just such a process for a broad range of cognitive development. In particular, he suggests that rather than viewing the development of cognitive capabilities as learning, they may be better viewed as the selection by the environment for potential capabilities that are part of our genetic endowment. In an extreme version of his views, people do not learn French, but the environment selects for that capacity. This view is modeled on the mod-

44
ern biological conception of immunity of Jerne, which is an environmental selection model.

As a consequence of encountering $s_i$, the future probability of $d$ affecting attitudes and behavior changes. One could make assumptions about the likelihood of encountering such situations in $S$, and the distribution of the strength of the selection effects of $s_i$ on the subsequent probability of $d$ manifesting itself. Plomin (1991) has suggested that in some cases, selection of environments is itself in part genetically determined. There may be a genetic basis for environment selection such that one might be genetically disposed to select certain environments. If that is the case, Plomin argues, estimates of environmental effects on behavior may be overestimates. Ross and Nisbett (1991), in their analysis of person and situation effects, make a related observation. They note that the fundamental attribution error, the tendency to overestimate person effects relative to situation effects, is partly due to the selection of environments by actors. In their argument, they note that observers fail to realize that actors often select the environments they are in such that stable individual behavior patterns may be due to an individual systematically selecting certain environments. For example, an academic may be viewed as consistently quiet and shy and that trait is attributed to their personality. What may be happening is the person systematically selects settings like libraries which drive all actors in those settings towards behaving in a quiet and shy manner. Clearly such notions have broad applications.

Using such a framework, one could model the effects of dispositions on job attitudes and see if the model fits the distribution of the predicted outcome(s). Such a story is basically an evolutionary selection formulation.
applied to the effects of dispositions on reactions to some $s_i$ such as a job (drawing on the organizational example discussed by Davis-Blake and Pfef-
fer (1989)). Lave and March (1975: 302-304) discuss just such learning or adaptation in the context of personality development. It would apply equally well in the context of job reactions. The person is in some $s_i$ with a set of dispositions $d_j$ such that each $d_j$ has a probability of being enacted given $s_i$. Suppose the reward consequences for acting on any of $d_j$ are positive. Further, suppose the person’s positive response to the outcome of doing $d_j$ leads to positive outcomes from $s_i$. This can lead to repeated enactment of $d_j$. However, choosing any other option from the set of actions enabled by $d_j$ may also lead to positive outcomes. However, those positive outcomes will be unobserved given the initial pattern established by the choice of a particular behavioral enactment of $d_j$. Lave and March (1975) characterize this as superstitious learning, since other behavioral choices may lead to equally good outcomes, but the person does not search the space of possible choices given the initial positive outcomes. The critical point in this example is that several persons, with similar dispositions, given the arrival of different $s_i$ and the stochastic choice of different behavioral enactments of $d_j \mid s_i$ can show great variation despite similar starting configurations of dispositions. Further, one would expect such dispersion to occur early and be negatively accelerated with respect to time as particular behaviors are overlearned and less liable to extinction. This probability learning example can be applied quite plausibly to personality development or to learning patterns of behavior on the job. Presumably such patterns stabilize as persons seek out $s_i$ that are most congenial to the repertoire of behaviors they have developed.

In fact, such a viewpoint is embedded in the interactionist approach set
out by Kenrick and Keefe (1991), in their analysis of biological predispositions towards certain behaviors and affective states. They propose that “a biological predisposition is one factor in a multiplicative equation that also includes environmental events.... For instance, depression may be predisposed by certain genetically based biochemical thresholds, but it will not be activated without inputs from the environment (e.g. repeated failures, loss of resources, disease, certain weather conditions)....” If one assumes that such environmental events are acting as selection devices and are stochastic, then Kenrick and Keefe’s framework operates precisely in the stochastic fashion described above. That is, some disposition is selected for by environmental conditions which are arriving stochastically. If such processes are at least partially independent of the predispositions of the individual (i.e., individuals do not reliably sort themselves into environments which are conducive to the elicitation of behaviors to which the person is predisposed), then this is a partially determined process much like that suggested by Bandura. Campbell (1988) makes a similar observation about attitudes as acquired behavioral dispositions. He notes that a behavioral disposition can be acquired through multiple mechanisms. By implication, without those mechanisms, the disposition will not be acquired. He also points out that the expression of such a disposition is contingent on the availability of modes for its expression.

This selection by the environment argument suggests that dispositions influence responses to environments depending on the effects of past occurrences of $s_i$ which have selected for or against a particular $d_j$ that influences job reactions. Depending on the history of such events, the person may be more or less strongly influenced by their initial disposition. One could
further elaborate the Bandura argument by arguing (plausibly) that such chance encounters exert greater influence when they occur early in the life of the person. This is a plausible assumption, consistent with some observations of variability in maturation and growth (Kenny, 1975). Let me illustrate this point with two very different examples which illustrate the generality of this perspective.

The first example is drawn from the population dynamics of organizations. Hannan and Freeman (1989) review much research supporting the concept of the liability of newness. The liability of newness hypothesis suggest that events and conditions early in the life of a firm are imprinted on the firm, exerting an influence throughout the life of the firm (Stinchcombe, 1965). Hannan and Carroll (1992) have shown an effect called density delay as one way in which such conditions are imprinted on the firm. Density delay occurs when the population density in a firm’s niche at birth affects available resources, leading to a permanent positive effect (if a firm is born into a niche when there are few other competitors) or negative effect (if a firm is born into a niche when there are relatively many other competitors). For my purposes, one can view this process as follows. The premise is that these initial conditions $s_i$ select for a certain set of behaviors from $d_j$. Those who survive are, by definition, rewarded for enacting the chosen behaviors from $d_j$ and will persist in those choices. Behaviors learned in difficult times (many competitors in a niche) will be very different from those learned in relatively benign times (few competitors in a niche). These are in fact the results shown by Hannan and Carroll (1992). These population ecology arguments applied to organizations suggest that $s_i$ are more likely to occur earlier in life course of the organization, and exert a continuing influence
throughout the life course of the organization.

These types of arguments, as suggested by the Lave and March (1975) model discussed above, can be made at the individual level as well. Besides learning, such influences can grow out of social processes such as individual level searching and sorting mechanisms that tend to match certain dispositions with situations. Thus, in the realm of jobs, we would expect that dispositions which have been selected for in the past will now influence current choices an individual makes about which jobs and social situations to select. That is, individuals will search for jobs where they can obtain valued outcomes. Simultaneously, employers who are making choices about the individual will also be influenced by observations of behavioral indicators of underlying dispositions (albeit only weak indicators) of the person. Such mechanisms can be applied to a range of social matching processes (Mortensen, 1988; Fichman & Levinthal, 1991). Consequently, both employers and job seekers will seek out information to facilitate matches of employers and job seekers which will lead to positive outcomes for both parties. Given the stochastic quality of the strength and arrival times of \( s_i \), inevitably there will be residual variability unattributable to either situations or dispositions. The size of such residual variability can vary with both the homogeneity/heterogeneity of the distribution of some disposition in the population, and with the heterogeneity and distribution of elements of \( S \). Note that the residual variability is not ‘error variance.’
Further Implications of Stochastic Selection Models for the Nature and Nurture Debate

Crawford and Anderson (1989) present several interesting implications of the sociobiological perspective on behavior which are quite consistent with the selection model just presented. Sociobiologists are interested in studying the manner in which behaviors are determined by both genetic heritage (nature) and environmental forces. Crawford and Anderson’s (1989) observations suggest several difficulties for analysis of dispositional effects consistent with the arguments presented above. Sociobiological viewpoints also are based on the Darwinian notion of selection. Not surprisingly, the role of environments is prominent in such models (see Kenrick, 1987 for a social psychological example of the application of the sociobiological viewpoint where environment plays a prominent role.) In these models, the behavioral dispositions of individuals interact with the characteristics of the environment. Thus some environments increase the probability of certain behaviors, while other environments decrease the probability of the same behavior. The probability of behaviors can not be assessed independently of the environment in which the behaving person is located. In fact, Crawford and Anderson (1989) provide two important theoretical results of interest for stochastic modeling of dispositional versus situational effects.

Genetically Influenced Behavior Is Not Invariant Behavior

The standard reading of the nature side of the nature-nurture argument is that if behaviors are genetically predisposed, behavior patterns should be predetermined and invariant. Thus, if twins reared apart both have a genetic predisposition towards positive affect, this should be manifested in
correlated levels of job satisfaction in identical or fraternal twins raised apart (Arvey, Bouchard, Segal, & Abraham, 1989). However, an environmental interactionist perspective suggests that a given predisposition can lead not to invariant behavior, but to contingent behavior, where the behavior enacted is determined in part by the environment in which the behaving organism is currently located. Crawford and Anderson suggest such contingency can be concurrent or developmental. By concurrent they mean that the choice of behavior to emit (where choice is genetically constrained to some set of behaviors) is determined by the current environmental conditions. Thus, for concurrently contingent behavior, the current environment selects for the genetically predisposed behavior. By developmentally contingent, Crawford and Anderson mean that an environment during the organism’s development was the selection device for a particular behavior. In either case, the behavior is selected for by the environment. Note here the very close parallel between developmentally contingent behavior and the notion of imprinting and density delay in the organizational ecology results discussed above. Given their common intellectual ancestry, this is not surprising.

Sociobiologists and social psychologists applying such models to social behavior (e.g., Kenrick (1987), Kenrick & Keefe, (1991)) are arguing that the interaction of genetic behavioral dispositions with the environment determines some observed behaviors. This interaction idea is very similar to environments acting as selection mechanisms for particular behaviors. If such contingent mechanisms are operating, then it may be the case that genetically identical individuals exposed to different environments will show no phenotypic similarity with respect to behavior, when in fact there is genetic influence over behavior, but it acts as only a partial constraint. Crawford
and Anderson (1989) conclude for the example of identical twins reared apart that “alternative behaviors . . . may have zero heritability, may be completely dependent on conditions in the environment, and may be under genetic control. Moreover, the environmental contingencies may be either concurrent or developmental. A finding of zero heritability must not be taken as evidence that genes are not involved in the development of differences in behavior!” (emphasis added) (p. 1453).

**Genetically Influenced Behavior May Be Environmentally Determined**

If the argument for developmentally and concurrently contingent behaviors is correct, then it suggests that such contingent behavioral dispositions can be uncorrelated across environments (which is precisely the null hypothesis of most studies which look at dispositional effects!). Crawford and Anderson suggest that asking for behavioral invariance across environments is the wrong question. The right question, they believe, is “not the heritability of a trait but its correlation with inclusive fitness in the environment in which it occurs or, if the environment has changed, the environment in which it evolved (p. 1453).” A more useful approach is to assess how different environments select for different genetically predisposed behaviors. In the case of twin studies, they suggest “The focus . . . is not heritability or on the relative importance of genetic and environmental variation in accounting for variance in observed behavior. It is on using the twin study to control genetic variation so that we can see how environmental conditions are involved in the production of differences in behavior.” (p. 1456).

This conclusion is very similar in form to Buss's (1989) conclusion that
comparisons of the strength of situation versus person effects in determining the relative effects of personality on behavior was “useless” (p. 1378). A more appropriate question was to ask about the conditions (both substantive and methodological) under which traits (or dispositions) will be dominating and the conditions under which situations (or environments) become dominating. In fact, consistent with the Crawford and Anderson (1989) and Kenrick (Kenrick, 1987; Kenrick & Keefe, 1991) positions, Buss suggests that situations and environments interact such that certain trait effects are observable under some situations and not under others. Such a conclusion is consonant with the interaction model outlined above or the stochastic environmental selection model.

**Leadership as an Occasion for Stochastic Modeling**

One problem that lends itself to stochastic modeling and captures the person vs. situation tension is leadership. Leadership theorizing has run the gamut from trait based work to situational determinism to various amalgams of interactional or contingency models which require the matching of a trait to a situation. Outcomes of leader behavior may often be weakly linked to leader actions (Pfeffer, 1977; Lieberson & O’Connor, 1972). Such conditions may be conducive to this environmental selection argument. One example with such a stochastic flavor is Simonton’s (1985) analysis of the effects of intelligence on personal influence in small groups. His results suggest that intelligence acts in a non-linear fashion in influencing the likelihood that a person will be influential in a small group. He found that children who were more intelligent, but *not too much more intelligent*, were most likely to be influential in small groups. That is children who were very much brighter
than their fellow group members often could not communicate (because they could not be understood) with others in the group. Children who were no more intelligent than the others in their group had relatively little distinctive to contribute. Those who were sufficiently brighter so that they could both make a distinctive contribution and be relatively easily understood were most likely to be influential and take positions of leadership and influence.

This result illustrates how the group (the environment) selects for certain attributes, and that such attributes cannot be viewed independently of the environment. Similar results for intelligence and leadership emergence have been reported in adults (Bass, 1990; Simonton, 1995). Bass cites a number of studies in managers and adults showing that “large discrepancies between the intelligence of potential leaders and their followers militated against the exercise of leadership.” (Bass, 1990: 83).

Simonton does not model influence and intelligence as stochastic processes. One can develop such a model. Group composition, entry and exit are can be modeled as stochastic. In fact, the modeling of labor and manpower flows in economics, sociology and operations has been almost exclusively stochastic (Bartholomew, 1982). If such flows into and out of groups are treated as stochastic, the Simonton theory of leadership emergence would then be developed as a process model. How might this look.

I would suggest considering two factors. First, since the effect of intelligence in this model is contingent on group composition, we need to ask about how group the group is composed. If one treats entry and exit into the group as a stochastic process, the composition of the group would evolve as a stochastic dynamic process. Entry might be viewed as determined by random arrivals of individuals who enter conditional on the availability of
places for the individual. Such vacancies would also arise stochastically as people exit. Thus, the intelligence distribution in the group which sets a context for the emergence of leadership would be evolving probabilistically. A second concept one could develop is occasions for leadership. It is reasonable to suggest that groups and organizations vary in their need for leadership behavior. A number of leadership theories have developed this concept. If problems arrive probabilistically, these problems represent occasions for leadership. An aphorism that captures the flavor of this argument might be ‘For great leaders to emerge, they need extraordinary times.’ This argument has some of the flavor of the work of March and his colleagues (e.g., Cohen, March & Olsen, 1972). This is not surprising, as much of March’s work is in the process theory tradition and he is an advocate of such modeling efforts (Lave & March, 1975). The critical argument here is that the context is best viewed as a dynamic stochastic process, and then occasions for leadership and leadership emergence may be treated best as process theories.

The conclusion from this discussion of several lines of research and theory related to the disposition vs. situation controversy is that a stochastic framework with environment(s) $s_i$ acting as a selection device favoring one disposition versus another is a plausible and reasonable approach to the disposition vs. situation controversy. Recall that I proposed this model to suggest that process models may be quite congenial to this issue, and begin to help address the difficulties which attend the application of deterministic, LaPlacian models to nature vs. nurture controversies.

A second encouraging feature of such an environmental selection model is its broad generality. The Spilerman (1970, 1971) model of race riots can be interpreted as such an environmental selection model. His conclusion
was that cities did differ in their predisposition to engage in riots, but that the predisposition was not a sufficient condition to cause a riot. Rather, some stochastic, non-deterministic process generated an environmental circumstance which kindled a riot. Thus, different cities with identical dispositions had different outcomes, where the environmental stochastic variation selected for the disposition to riot in some cases and not in others. Conceptually, this selection model is similar to one that examines the dynamics of organizational populations, where different environments select for or against particular dispositions of firms (firm strategies) (Hannan & Freeman, 1989). I think such models can be constructed for a host of interesting problems (in fact many already have been constructed) and represent a very fruitful line of inquiry in modeling many social processes (Lave & March, 1975). Further, such models do not impose the deterministic LaPlacian criteria of full explanation that I have argued are so often inappropriate and unrealistic.

**Conclusion**

The question addressed here, explanation of variance and its attendant conceptual, statistical and methodological difficulties has been counterposed against stochastic explanations of social processes. I have suggested a number of reasons why such process models can be more fruitful. The position I am taking is that some frequently encountered questions, like the many variants of the nature vs. nurture question, have embedded in them such assumptions about determinism and indeterminism. This view raises some difficult philosophical problems. Some readers probably can sense in
their own discomfort the emotional difficulties attendant upon accepting a stochastic view. Workers in physics, where quantum mechanics has come to dominate, and biology, where the Darwinian revolution is firmly entrenched, have struggled with this question, and have appreciated the difficulties of this issue.

Jacques Monod, a vigorous advocate of the role of chance processes and selection mechanisms in biology, noted the implications and difficulties of embracing a non-LaPlacian view of the world.

When I say that living beings as a class are not predictable on the basis of fundamental principles, I do not mean to suggest at all that they cannot be explained by these principles, that they are in some way beyond these principles, and that other principles applicable only to them have to be devised. In my view the biosphere is just as unpredictable as the specific configuration of the atoms making up the pebble in my hand. No one will raise as an objection to a universal theory that this theory does not establish and predict the existence of this specific atomic configuration. We will be satisfied if this real and unique object before us is compatible with the theory. According to the theory, this object does not have to exist, but it may. This satisfies us if we are talking about pebbles but not if we are talking about ourselves. We want to know that we are necessary, that our existence is inevitable and was determined since the beginning of time. All religions, almost all philosophies, and even, to some extent, the natural sciences attest to the indefatigable and heroic
efforts of humanity to deny its own accidental origins. (Jacques Monod, quoted in Eigen & Winkler (1981), p. 162)

By contrast, Einstein was always deeply troubled by the quantum mechanical view of the world and the critical role of stochastic processes in such a world. He observed, in a famous letter to Max Born that

Quantum mechanics is certainly imposing. But an inner voice tells me that it is not yet the real thing. The theory says a lot, but does not really bring us any closer to the secret of the ‘old one.’ I, at an rate, am convinced that He is not playing at dice (Einstein and Born, 1971: 91).

Einstein consistently felt that the role of chance processes was a troublesome and deeply dissatisfying feature of quantum physics (Clark, 1971), struggling with the very human beliefs observed by Monod.

Organizational scientists, like all social scientists, must struggle with this same issue. Indeterminism presents many formidable analytic problems. Probably more troubling than the technical and conceptual issues are the deeply troubling emotions scientists have encountered when confronting the notion of indeterminism and the difficulties for causal explanation it raises.
References


59


