Vita Contemplativa : Why I Stopped Trying to Understand the Real World

William H. Starbuck

Organization Studies 2004 25: 1233
DOI: 10.1177/0170840604046361

Vita Contemplativa

Why I Stopped Trying to Understand the Real World

William H. Starbuck

William H. Starbuck
New York University, USA

Abstract

Years ago, I believed that rationality could manufacture understanding. I lived in physical and social environments that were real and I wanted to understand the social realities. I wanted to create a genuine 'behavioral science' based on mathematical models, computer simulation, and systematic experiments. Various experiences over the years have challenged these beliefs. I discovered that rationality can not only be a deceptive tool but a potentially dangerous one, and I learned a few techniques to help me challenge my rational thought. I discovered that research findings have very low reliabilities, that some fields make no discernible progress over many decades, and that societal cultures strongly influence researchers' judgments about what constitutes useful knowledge. I saw that much that passes for research is merely random noise dressed up in pretentious language. Rather than realities, the social systems I was studying proved to be arbitrary categories created by observers or social conventions. I became an advocate for research that actively attempts to change situations rather than merely to observe what happens spontaneously.

Key words: research, methodology, social construction, hypothesis tests, rationality

This article describes an intellectual journey. Since the journey developed gradually, the account is generally chronological. But a completely chronological account would be as confusing for you as it was for me, so I have rearranged time to create some conceptual threads. As one result, the article ends near a temporal middle.

Real Science in a Real World

While I was in graduate school and for several years thereafter, I believed that rationality was a tool one could use to manufacture understanding. Both my physical and social environments were real and I wanted to understand the social realities. What laws govern human behavior? I imagined that these laws, when we researchers discovered them, would be as universal and timeless as the laws of physics or chemistry. That few such laws had been discovered as yet, I thought, was mainly a result of the lack of rigor in psychological and social science methods. I wanted to create a true 'behavioral science' based on mathematical models, computer simulations, and systematic experiments.

Organization Studies
25(7): 1233–1254
ISSN 0170–8406

www.egosnet.org/os

DOI: 10.1177/0170840604046361

These notions came from my environment, of course. I had studied science in college, including graduate-level courses in mathematics and electronic engineering. My science teachers had emphasized the reliability of mathematical formulations and systematic experiments, so it seemed obvious to me that behavioral research had been lacking both in mathematical theories and in systematic experiments. I had developed an interest in computers early in the 1950s, had worked at IBM during summer vacations from college, and had participated in the engineering of IBM's first large computers (the binary 701 and the decimal 705). It seemed to me that computers would allow much greater theoretical complexity than would algebra.

Although when I left college I intended to get a doctorate in mathematics, I became a doctoral student in administrative science instead, after Dick Cyert and Jim March hired me to run some experiments for them (Starbuck 1993a). I was so greatly impressed by Thibault and Kelley's (1954) analysis of laboratory studies of small-group behavior that I chose social psychology as my major field. During my doctoral studies, I took a course in 'mathematical social science' that was taught by Alan Newell, and I heard Herb Simon advocate the value of mathematical modeling. Newell and Simon wrote computer programs that modeled human problem solving, and Cyert and March created simulation programs that imitated the decision processes of managers and companies. I attempted several doctoral dissertations, one of which involved a computer simulation of a large division of the Koppers Company and two of which involved mathematical modeling.

The Graduate School of Industrial Administration (GSIA) at Carnegie was an exceptional and exciting research environment. A major cultural theme there was the idea that administration lacked a needed scientific basis, and many people at GSIA saw themselves as adventurers in the frontier of science applied to administration. Some of the faculty had participated in the founding of the Institute of Management Sciences (a professional association that has merged into INFORMS). Several professors were developing mathematical approaches to decision making, such as production scheduling and inventory control. Franco Modigliani and Merton Miller were creating a foundation for finance based in economic theory. Alan Newell and Herb Simon were seeking to place cognitive psychology on a stronger basis through computer simulation. In addition, Simon (1950, 1952) was an enthusiastic missionary for 'organization theory', which he saw as a broad category that embraced scientific management, industrial engineering, industrial psychology, small-group psychology, human-resources management, and business strategy. Jim March and Herb Simon (1958) published one of the first books that integrated ideas drawn from many fields about organizations. Of the fifteen or so professors at GSIA during the late 1950s, three have received the Nobel Prize in Economic Sciences, five more were nominated repeatedly for that prize, and one received a similarly prestigious award in sociology.

My first academic employer, Purdue University, also offered me an exceptional and exciting research environment. A group of us united to develop experimental and mathematical methods of studying human behavior. At first there were about six of us, but the group gradually grew to about a dozen. We met regularly for lunches at which we discussed research; we created a Department of Administrative Sciences that offered an avant-garde doctoral program; we invited renowned researchers to visit Purdue for short periods; and we quickly developed a wide reputation for innovation and scientific excellence.

Thus, my research during the 1960s involved mathematical models, laboratory experiments, and computer simulation. My experiments focused on choices made by individual people, negotiations between two people, and teams managing hypothetical companies that competed in computer-based markets. In promoting experimentation, I collaborated with Vernon Smith, who was pioneering experimental approaches to economic behavior, for which he too received the Nobel Prize in Economic Sciences. Vernon helped me by inviting me to join him in seeking a large research grant, and he and I jointly obtained funding for, designed, and built a laboratory for experimental research (Fromkin 1969). I also collaborated with John Dutton. John and I spent six years trying to simulate the behaviors of a factory scheduler, and we co-edited an anthology of exemplary computer simulations.

Then, my reality started to come apart.

Flaws in Research Methods

My enthusiasm for experiments waned in the mid-1960s after I attended a workshop about experimental studies of economic behavior. Its organizers intended the workshop to reinforce interest in experimentation, but it had quite the opposite effect on me. The experiments we discussed and ran showed me the supreme importance of having 'subjects' who sincerely want to support an experimenter to produce desired behaviors. Subjects who want to undermine an experiment can easily do so.

The organizers of the workshop inadvertently facilitated the most telling example of experimentation gone wrong. During the first session, we participants were told that the organizers had promised on our behalf that we would all be subjects in an experiment that would be run by one of us. There was an immediate bristling reaction as people realized that they were being required to participate involuntarily. We were then divided into groups of three competitors; each person was supposed to pretend to be making bids in a three-person market. We were then handed written instructions and told to read them. When people attempted to ask questions about the instructions, the experimenter gave terse responses or dismissed the questions as trivial. I recall someone, possibly myself, pointing out that the instructions said 'There will be no collusion detectable by the experimenter' and asking 'Does this mean that you do not want us to collude?' The experimenter replied, 'Just follow the instructions on the paper.' As soon as we had the opportunity, my two 'competitors' and I met to discuss how we could collude in a way that the experimenter would be unable to detect. We decided that the bid prices in our market would rise monotonically, and that the exact increases and the identities of the winning bidders would be determined by random numbers, such as the number of words spoken by the experimenter when asking us to submit bids. None of us aspired to perform well according to the experiment's performance criteria; all of us enjoyed proving that we could screw up the experiment. We later discovered that several other groups of three competitors had also met to discuss how they could collude and they had each invented an artificial behavior pattern that they expected to violate the experimenter's expectations.

Of course, this was an extreme instance. Most subjects try to follow an experimenter's instructions insofar as they can understand what the experimenter wants. But as I saw it, this observation only underscores the most central point — experimenters can control their subjects' behavior to high degrees. If experimenters give complete and precise instructions, nearly all their subjects will make every effort to carry out these instructions. If experimenters leave something ambiguous, the subjects do whatever they please, and different subjects are likely to do different things. For example, if an experimenter tells subjects to 'Try to earn as much money as you can,' the subjects will act as if money is their primary goal, and nearly all will do this even if the amounts of money seem trivial to them. Experimenters can take various steps to assure that their subjects understand their instructions and to motivate them to follow the instructions carefully. For instance, experimenters can offer larger monetary payments or they can make actual cash payments immediately.

I also came to doubt that I could produce in a laboratory, with students as subjects and the budgets available to me, conditions that resembled in a meaningful way those outside a laboratory. It was obvious that subjects' behavior is strongly affected by the salience of rewards and that it changes over time. I could not assign tasks that would require specialized skills or weeks of training. In actual work organizations, people come to know each other over months and years and their rewards may involve significant wage changes, promotions, and social statuses that persist for years.

Thus, I began to view laboratory experiments as exercises in the writing of instructions and the motivation of subjects, who would perform tasks having little significance outside the laboratory. I could elicit the behaviors I wanted if I wrote instructions that were clear enough and complete enough and I made sure the subjects understood and wanted to follow them. But was this a useful goal? To demonstrate that I could write instructions and persuade subjects to follow them? Certainly, the results I got from experiments strongly reflected my own goals and my own beliefs about what behaviors I wanted to observe, and my experiments were producing evidence about characteristics of myself rather than universal laws of human behavior. Was this not a silly game to be playing?

In fact, laboratory experimentation seemed to bear a strong resemblance to computer simulation, although simulation also raises some different issues (Starbuck and Dutton 1971; Starbuck 1983). When simulating, researchers try to write programs that correctly express their assumptions. Researchers face no motivational challenge: computers follow instructions precisely insofar as they can do so. The computers' actions trace out implications of the researchers' programs, and if the programs accurately represent the researchers' assumptions, the computers' actions demonstrate the logical implications of the researchers' assumptions. Thus, computer simulation is very similar to mathematical analysis. When one creates a mathematical model, one states a set of assumptions and then uses algebra to extract some implications of these assumptions. One can experiment with different assumptions until the model exhibits the properties one desires. Likewise, when one creates a computer simulation, one states a set of assumptions and the computer generates some implications of these assumptions. Since computers do nothing on their own initiative, simulation can only reveal the logical implications of what researchers believed before they created the simulations or what they assumed during the process of creating their models. These implications may surprise the model builders, and when they do, the model builders have to decide whether to change their assumptions. Because assumptions are always somewhat arbitrary, model builders can experiment with different assumptions until the computer generates the kinds of outputs they desire.

One does computer simulation because one does not know how to model one's theory mathematically. This might occur because one has little knowledge of mathematics, but it can also occur because mathematics is not capable of providing answers. Generally speaking, mathematical formulations are difficult to manipulate unless one limits the mathematics to linear functions. Linear formulations remain solvable even if they include many, many equations. But, nearly all nonlinear functions pose insurmountable challenges, especially when the formulations involve several equations. The nonlinear functions include ones that change in different directions for different values of the variables and ones that involve abrupt branching. Computers allow researchers to develop flawlessly the logical implications of their nonlinear assumptions, and computers impose very weak restrictions on the complexity of simulation models. In principle, simulation can disclose the consequences of a multitude of nonlinear, discontinuous, interacting assumptions.

However, in practice, simulation has traps for the unwary. A multitude of nonlinear, discontinuous, interacting assumptions has the potential to generate outputs that appear mysterious, even magical. Because simulations are process oriented, researchers have to specify activity sequences even when they lack information about them. Large, complex simulation models are virtually impossible to validate in detail. Computers generate outputs without explaining their reasoning. Researchers can add instructions to their programs that record calculation sequences but simulation programs typically incorporate so many microscopic steps that the explanations themselves pose serious data-analysis challenges. As a result, researchers are likely to end up with simulated behaviors that they cannot understand. In this fashion, simulation confronts Bonini's Paradox. As I phrase it, Bonini's Paradox is: 'As a model grows more realistic, it also becomes just as difficult to understand as the real-world processes it represents.' A researcher builds a model to gain or demonstrate understanding of a causal process, and the researcher states this model as a simulation to allow complex and realistic assumptions. The resulting program generates outputs that resemble those observed in the modeled situation. But the model itself is very complex, and the interdependences between subroutines are obscure, so the model is no easier to understand than the original causal process.

I call this proposition Bonini's Paradox because I first encountered it in the context of Charles Bonini's doctoral dissertation. Chuck was a doctoral student with me at Carnegie Institute of Technology, and both of us attempted to create computer simulations of business firms. Chuck was much more successful than I, for he completed a simulation model and a dissertation, whereas I abandoned my modeling effort and I never completed a proper dissertation. Chuck's model represented a hypothetical firm's detailed decision making as it decided how much to produce, what prices to charge, and so forth. In a short time, he could generate many years of decision making, and he could vary elements of both the decision processes and the environment of the firm. But in his dissertation, Chuck (Bonini 1963: 136) wrote: 'We cannot explain completely the reasons why the firm behaves in a specific fashion. Our model of the firm is highly complex, and it is not possible to trace out the behavior pattern throughout the firm ... Therefore, we cannot pinpoint the explicit causal mechanism in the model.'

Insofar as one is creating a theory to enhance human understanding, one needs to respect the limitations of human cognitive abilities. People need theories to simplify their worlds as well as to represent them. McClain (1981) observed that when policy recommendations are supported by simulation models, policy-makers only believe the models that have few equations and transparent relations between inputs and outputs. Several studies have shown that being able to understand computer programs requires the ability to draw both forward inferences and backward inferences — to explain what specific assumptions cause the programs to do what they do as well as to explain what consequences follow from various assumptions (Fitter and Sime 1980; Green 1982; Green et al. 1980). So theorists need to be wary of the freedom that simulation appears to offer.

Science Struggles Against Human Rationality

My efforts to analyze the results of an experiment showed me how deceptive rationality can be as a tool for understanding. In 1966, I was trying to write a paper about the behaviors of teams that had played a business game, but my efforts were going nowhere because the outputs from the statistical analyses differed greatly from the hypotheses I had held when designing the study. I tried to introduce various correction factors, but they helped not at all. So I decided to figure out what the data were telling me. I constructed diagrams that represented high correlations by thick lines and low correlations by thin lines, and then I began to play the game of 'why X correlates with Y but not with Z'. After a couple of weeks I had constructed a complete and logically integrated explanation for the relations among variables.

But something nagged. It was a good theory, but it was quite at odds with the one I had held when designing the study. I decided to trace back through all of the statistical analyses. The statistics had been produced by my research assistant, and although I trusted his work, I had not had first-hand involvement with the raw data or the calculations. I thought that a close look at the data and the analysis process might help me to comprehend the differences between my initial expectations and the findings. To my surprise, I discovered that very early in the analytic process, my assistant had made a data-entry error. The experiment involved four treatments, so some statistical analyses required adding correction factors that would make the treatments comparable. When correcting for one of the four treatments, my assistant had omitted two minus signs and so he had added instead of subtracted, which had displaced the data from this treatment even farther from the other three treatments. Hence I had just spent weeks trying to make sense of data that contained large systematic errors. Moreover, I had been quite successful. In effect, I had constructed a logically satisfying theory based on random noise!

Hayek (1975: 92) observed: 'It may indeed prove to be far the most difficult and not the least important task for human reason rationally to comprehend its own limitations.' The realization did not come easily or quickly to me, but I gradually began to view rationality as a potentially dangerous scientific tool. Rationality arises from human physiology; our minds feel comfortable when we perceive relations as being logical. Our shared rationality helps you to understand what I am saying. But rationality also constrains our ability to understand because our judgments about whether we do understand involve rational assessments of our explanations: when our minds say we understand, we stop seeking for further understanding. Rationality also warps our perceptions, and it leads us to oversimplify (Faust 1984). Such distortions are probably consequences of the physiology of human nervous systems.

Scientific rationality is an extreme ideal type that has been constructed through centuries of discussion by philosophers and scientists and implanted in researchers through education and socialization. One can observe the participants in academic seminars shifting into a ritualistic mode of rationality. This mode reduces all conditions to binary states — good or evil, true or false, and consistent or inconsistent. It rejects loose ends and fosters ludicrous extrapolations. But this scientific rationality generates logical contradictions, distorts our observations, and extrapolates incomplete knowledge to ridiculous extremes. We seek scientific rationality because it pleases our minds, but what gives our minds pleasure may not give us insight or useful knowledge.

Despite my growing skepticism about the trustworthiness of rational thought, I continued to publish articles that incorporated mathematical reasoning until 1973. That 1973 article characterized the conditions needed for a social system to undergo a rapid dramatic revolution. The mathematics suggest that it is not meaningful to try to explain why such revolutions begin; one can say when a revolution began but not why it began at that time. I find such analyses revealing, but the limitations of mathematical analysis do lead model builders to pursue the implications of analytically tractable assumptions even though the assumptions are far from plausible as descriptions of the modeled situations. As a result, nearly all research seminars that include mathematical models become ceremonial displays of speakers' mathematical skills.

One of my colleagues at the International Institute of Management in Berlin in the early 1970s, Bo Hedberg, announced that he had received a research grant from the Swedish government to study stagnating industries. He said that he wanted to find out why some industries stagnate and drive firms out of business and put people out of work.

Because my background and orientation are social-psychological, I tended to place the responsibility upon the firms rather than their environments. I said something tactful like: 'Bo, your thinking is all screwed up! The interesting question is not why an industry stagnates. Technologies are always evolving, populations are always migrating, prices are always shifting — it's inevitable that things will change, and some of these changes will make some industries obsolete. The interesting question is: why do people remain in an industry that they recognize is stagnating? Why don't firms move into more promising industries when their current ones start to stagnate?'

With equal tact, Bo responded, approximately: 'You're spinning an academic fantasy. A firm can't just pick up their product line and their engineers and plunge into another industry. Their specialized skills and business connections make them captives of their environment. The firms in an industry have to evolve together. It's a societal problem to create incentives that keep industries vital, that keep them evolving in line with social needs and economic and technological opportunities.'

Obviously, we disagreed: I was saying industrial stagnation posed problems for the managers of individual firms, whereas Bo was saying industrial stagnation posed problems for the setters of governmental policies. We decided to resolve our argument by doing research together (Starbuck et al. 1978)

Several years later, we resolved our argument by concluding that both of us had been partly right. Of course, such an outcome was both necessary to preserving our friendship and a result of our friendship, which grew much stronger through years of cooperation.

We decided that business crises are indeed produced by firms' environments, although not exactly in the way that Bo had conceived initially. He had rightly seen that environments change so as to obsolete some markets, products, and technologies, but he had not seen that environments make it hard for firms to adapt to these changes. Environments propound ideas about how to organize, and then the environments behave in ways that make these ideas unrealistic. Business crises are also produced by firms themselves, but somewhat differently than I had supposed at first. Although firms do make mistakes, I had underestimated the pervasiveness of successful firms' efforts to stabilize their environments. A successful firm knows how to succeed in its current environments, and it is not sure it would be equally successful in altered environments, so it wants its environments to stay as they are. Neither of us had foreseen our most interesting finding, which was that serious crises arise from exactly the same processes that enable firms to become successful.

More generally, our joint research showed me how two viewpoints that appear to be completely antithetical can synthesize into a broader integrated understanding. Collaborative research can foster progress by framing issues as conflicts and then encouraging the collaborators to discover gradually that the conflicts did not actually exist. Debates help researchers to clarify concepts, and dialectical reasoning helps researchers to break out of the mental prisons they build with rationality.

Science Struggles Against Social Institutions

I discovered the ambiguity surrounding human judgments about research findings when I became the editor of *Administrative Science Quarterly* in 1968. My predecessor bequeathed me a thigh-high stack of manuscripts that needed review. I was embarrassed that many authors had been waiting months for feedback, so I weeded out the topics that obviously did not suit the journal and then mailed manuscripts to hundreds of reviewers. After a few months, I had received more than 500 pairs of reviews, and I was amazed by the discrepancies among the reviews: only a small fraction of the reviewers agreed with each other as to whether a manuscript should be accepted for publication, returned to the author for revision, or rejected. Counting an 'accept' as 1, a 'revise' as 0, and a 'reject' as –1, I calculated a correlation of 0.12 between the recommendations of pairs of reviews. This correlation was so low that knowing what one reviewer had said about a manuscript would reveal almost nothing about what a second reviewer had said or would say. More generally, the reviewers exhibited almost no agreement about what constitutes good research, what findings are credible, what topics are interesting, or what methods are appropriate.

The low correlation I calculated may have been attributable in part to *Administrative Science Quarterly*'s broad interest during its early years in all aspects of management. The manuscripts were diverse and I sent them to a quite heterogeneous cross-section of reviewers. Subsequently, studies have reported correlations between two reviewers that range from 0.08 to 0.38 (Starbuck 2003b). It is unclear whether higher correlations occur where manuscripts and reviewers are more homogeneous: although more homogeneous reviewers should tend to agree with each other, more homogeneous manuscripts bring smaller differences among reviewers to the forefront. Gottfredson (1978), Gottfredson and Gottfredson (1982), and Wolff (1970) have reported that reviewers for psychological journals agree strongly about the properties they want manuscripts to exhibit, but they agree much less strongly about whether specific manuscripts exhibit these properties.

Other studies have shown that manuscript reviews reflect biases. For example, Mahoney (1977, 1979) found that reviewers are much more likely to recommend acceptance or minor revision for manuscripts that agree with the reviewers' previous publications. Peters and Ceci (1982) showed that manuscripts that had been accepted for publication when their authors came from prestigious psychology departments were almost certain to be rejected when the authors came from obscure institutions. Nylenna et al. (1994) observed that Scandinavian reviewers give higher ratings to manuscripts written in English than to the same manuscripts written in the authors' native Scandinavian languages.

In 1970, I moved to Europe for a short stay that turned into four years. Before I left the USA, I had assumed that idealizations about research methods in social science were much the same the world over, so I was startled by some German professors' perception of American research as being mindless empiricism. One of them likened American social scientists to hamsters that run endlessly on their exercise wheels, but go nowhere. Not initially, of course, but eventually I came to see that much empirical research imitates prior research and adds nothing of value, except more lines on résumés. A few years later, after I had come to admire the insights of several European scholars, I returned to the USA and encountered the other side of this coin — Americans who were exceedingly proud of the superiority of American empiricism and disdainful of the less empirical European social science.

I have now lived and been employed for periods from several months to several years in seven countries, I have briefly visited academic institutions in eight other countries, and I have participated in academic meetings in another nine countries. These experiences have impressed me with the wisdom and diverse insights of people from different societal and academic traditions. Many societies have long traditions of excellent scholarship by very intelligent and perceptive people, and I believe academic researchers from every country can benefit from trying to understand these alternative traditions. At the same time, I think I may have seen loose correlations between academics' efforts to contribute to their societies, the quality of academic research, and the willingness of their societies to support academic research. I have surmised that where citizens perceive their universities to be contributing to their economic and social welfare, academic wages tend to be higher and research funds more available and researchers seem to be generally more committed to doing research. By contrast, where citizens perceive their universities as arcane enclaves, academic wages tend to be lower and research funds scarcer, and researchers seem to be generally less committed to doing research. If my surmise is right, it is not only academics that are passing judgment on the value of academic research.

Around 1980, I read a paper by Peter Grinyer and David Norburn (1975) that examined the relationship of profitability to the use of strategic planning. They discovered that profitable firms are nearly as likely to do no formal strategizing as to do it, and the same is true of unprofitable firms. As I had mindlessly assumed that strategizing was useful, this intrigued me, so I dug out as many studies as I could find of the relationships between profitability and the use of strategic planning. The oldest study, by Thune and House (1970), had reported a rather high, positive correlation. This discovery of a strong relationship stimulated others to make additional studies, partly because these other researchers thought they could improve on the study by Thune and House. Over time, the reported correlations between profitability and the use of strategic planning decreased toward zero. Eventually, in some studies that measured profitability with stock prices, the correlations varied around zero.

I wondered if I had possibly observed a widespread phenomenon. A social scientist reports finding a fairly strong relationship of some sort. This relationship might be quite general and robust, but it might, instead, result from methodological deficiencies or it might be a peculiarity of a specific source of data. The strong finding draws the attention of other researchers, who see deficiencies in the original study and who have access to different data. They too find relationships, but weaker ones. Still more researchers appear, who try slightly different analytic methods and different sources of data. These new findings indicate still weaker relationships. Eventually, as methods evolve, the reported relationships hover around zero and researchers lose interest.

Invited to write a chapter about theory building in organizational behavior, I took the opportunity to explore this conjecture about research evolution (Webster and Starbuck 1988). Jane Webster dug up the histories of nine relationships that had traditionally been important in industrial psychology. Since these relationships had remained important for many years, we expected that studies of them might show constant or increasing strength over time — increasing strength as researchers developed better measures and obtained more appropriate data. But five of the relationships we examined had trended toward zero over time: the correlations of job satisfaction with absenteeism, the correlations of turnover with realistic job previews, the correlations of turnover with job enrichment, the correlations of performance improvements with behavior modification, and the correlations of observed results with Fiedler's contingency theory of leadership. Three relationships had remained approximately constant for many years: the correlations of job satisfaction with job performance, performance improvements with goal setting, and the correlations of subordinates' perceptions of leaders with the leaders' intelligence. Measures of only one relationship had increased, but this increase was entirely due to the very oldest study, which had reported a very weak relationship. Furthermore, this relationship was trivial: some of the people who have stated in private that they intend to quit their jobs actually do quit them.

Jane proposed that industrial psychology might be producing poor research results because it lacked paradigm consensus. Kuhn (1970) had argued that scientific progress alternates between brief spurts of rapid change and long periods of consensus building. Had industrial psychology been fallow during a long period of consensus building? Did industrial psychology exhibit consensus? We looked at three measures of paradigm consensus: citing half-lives, percentages of references to the same journal, and numbers of references per article. We found that according to these measures, industrial psychology looks much like management, sociology, and other areas of psychology, and it does not look very different from chemistry or physics.

However, our measures of consensus made no distinction between substantive consensus and methodological consensus. On the one hand, Garvey et al. (1970) had inferred that editorial practices in the social sciences place more emphasis on methodology than do those in the physical sciences. On the other hand, when Campbell et al. (1982) had asked American industrial psychologists to recommend 'the major research needs' of their field, 105 psychologists had offered 146 suggestions, of which 106 were unique. Campbell et al. (1982: 71) inferred, 'The field does not have very well worked out ideas about what it wants to do. There was relatively little consensus about the relative importance of substantive issues.' Therefore, Jane and I speculated that industrial psychologists might disagree about the relative importance of substantive issues, but agree about proper research methodology. We advocated that, as a starting point toward clearer development, psychologists should establish consensus around a few substantive propositions (also see Pfeffer 1993).

I wondered if the lack of progress might be attributable to low standards for what constitutes a 'significant' relationship. Two of my colleagues at Purdue, Ed Ames and Stan Reiter (1961), had published an article pointing out how easy it is for macroeconomists to discover statistically significant correlations that have no substantive significance. Macroeconomists try to discern significant economic relationships by looking at the correlations between various time series (series of observations over time). But economic time series generally have high autocorrelations and these autocorrelations mean that many time series correlate with each other. An economist who starts with one time series and searches for a second one that correlates with it does not have to search long. Choosing time series entirely at random, an economist would need only three trials on average to discover a correlation greater than 0.71. Even if the economist removed linear trends from series before correlating them, the economist would require only five trials on average to find a correlation greater than 0.71.

I speculated that a similar phenomenon might occur with cross-sectional data. First, a few broad characteristics of people and social systems pervade psychological data — sex, age, intelligence, social class, income, education, or organization size. Such variables correlate with many behaviors and with each other. Second, researchers' decisions about how to treat data can create correlations between variables. Third, so-called 'samples' are frequently not random, and many of them are complete subpopulations even though study after study has turned up evidence that people who live close together, who work together, or who socialize together tend to have more attitudes, beliefs, and behaviors in common than do people who are far apart physically and socially. Fourth, some studies obtain data from respondents at one time and through one method. By including items in a single questionnaire or interview, researchers suggest to respondents that they ought to see relationships among these items. Lastly, researchers are intelligent, observant people who have considerable life experience and who are living successful lives, so they are likely to have sound intuitive understanding of people and of social systems. They are many times more likely to formulate hypotheses that are consistent with their intuitive understanding than ones that violate it; they are quite likely to investigate correlations and differences that deviate from zero; and they are less likely than chance would imply to observe correlations and differences near zero.

Thus, social science researchers should not expect correlations to center around zero, and statistical tests with a null hypothesis of no correlation are biased toward statistical significance. Jane culled through *Administrative Science Quarterly*, the *Academy of Management Journal*, and the *Journal of*

Applied Psychology seeking matrices of correlations. That is, we took account of all the correlations among all variables observed in a study, not merely the correlations relating to hypotheses. She turned up more than 13,000 correlations. In all three journals, the mean correlation was close to +0.09 and the distributions of correlations were very similar. Finding significant correlations is ludicrously easy in this population of variables. Choosing variables utterly at random, a researcher has 2:1 odds of finding a significant correlation on the first try, and 24:1 odds of finding a significant correlation within three tries (also see Hubbard and Armstrong 1992). Furthermore, the odds are better than 2:1 that an observed correlation will be positive, and positive correlations are more likely than negative ones to be statistically significant.

Reports by social scientists routinely overstate the generality of their observations. In particular, researchers often conceal the ambiguity in their observations by focusing on averages and using hypothesis tests about averages to convert ambiguities into apparently clear conclusions. Thus, instead of characterizing statistical findings by stating percentages such as '70 percent of adult men have brown hair,' researchers state, test, and do not reject the hypothesis: 'Men have brown hair.' Then they describe such findings by saying 'Men have brown hair' as if the description describes everyone or every situation. The distribution of hair colors becomes a generalization. Much of the time, such generalizations have no bases beyond computed averages, that is, 'An average man had brown hair.' Since social phenomena often have overlapping frequency distributions, comparisons between averages may say nothing about specific instances. For example, the average height of a man exceeds the average height of a woman, but the heights of men and women have frequency distributions that overlap nearly 100 percent. What is the probability that Robert is taller than Roberta?

Researchers use other language conventions as well to fabricate generality. One especially pernicious convention is the use of definite articles to describe representative instances, where indefinite articles would be accurate. In principle, a definite article ('the') denotes a specific, nameable instance. 'The environment' means one specific environment, such as the British steel industry or Indianapolis, and 'the organization' means one specific, nameable organization, such as Control Data or the University of Wyoming. An indefinite article ('a' or 'an') designates a typical, nonspecific instance. Thus, 'an environment' means one typical environment and 'an organization' means one typical organization. Confusion of definite and indefinite articles causes serious substantive problems throughout the social sciences. For example, their penchant for saying 'the organization' has allowed organization theorists to gloss over the differences between organizations and to speak as if all organizations act the same. Similarly, by saying 'the environment', organization theorists have understated the degrees to which environments are ambiguous, diverse, and selected. Indeed, a proposition that raises no alarms when phrased in terms of 'the organization and its environment' may seem implausible when phrased as 'all organizations and all of their environments'.

The main inference I draw is that the social sciences are being inundated with statistically significant, but meaningless noise — supposed 'findings' that say nothing of lasting value, but enable researchers to publish a multitude of articles. Reports frequently use misleading language to exaggerate the generality of research findings. Social and psychological phenomena may be difficult to study and to understand, but social scientists and psychologists have control over some important factors that make progress difficult. The rituals of academic research, including statistical significance tests and editorial decision making, are smothering potentially useful research. For the social sciences to make real progress, social scientists and psychologists will have to decide that they actually want progress to occur. They will have to decide that successful careers and the maintenance of status hierarchies should take second place to revealing research designs and careful assessments of research contributions.

Organizations Become Less Real

Around 1970, Marvin Dunnette asked me to write a handbook chapter about the relation of organizations to their environments (Starbuck 1976). I started thinking about the boundary between organization and environment, and I began to see that this boundary is not at all discrete. In fact, it seemed that the boundary between organization and environment is, to no small degree, an invention of the observer. Although some activities might be classified as clearly internal to a specific organization, and some activities might be classified as clearly external to that organization, many activities involve interactions in which both organization and environment participate. Thus, there is no clear point at which internal ceases and external begins. Then I began to think about how to measure the degree to which someone or some activity occupies a position near the center of an organization versus the periphery. I found that several possible measures are quite at odds with each other. According to some measures, a company president would be central to the company, but according to other measures, the president is far out on the periphery. Not only can the boundary between an organization and environment depend on what aspects of activities an observer considers, but the boundary can vary from time to time depending on the activities that people are performing. In addition, each organization interacts with several different kinds of environments — legal, financial, social, transportation, technological, and so on.

Thus, organizations began to look less and less like distinct social systems and more and more like arbitrary categories created by observers or social conventions. The previously real phenomena I had been trying to study were vaporizing into mental and social constructions.

When Dunnette's *Handbook of Industrial and Organizational Psychology* appeared, my chapter was adjacent to one in which Roy Payne and Derek Pugh (1976) reviewed scores of studies in which researchers had asked organizations' members to characterize their organizations' structures and cultures. Their data indicated that different members of an organization disagree so strongly with each other that it makes no sense to talk about an average belief, and members' beliefs about their organizations correlate very weakly with measurable characteristics of their organizations. In other words, the properties of organizations do not even have the support of consensus.

Peter Grinyer and David Norburn (1975) added another small nail to this coffin by observing that firms' profitability correlates inconsistently and meaninglessly with the degrees to which senior executives agree about their firms' objectives or their personal responsibilities.

Sometime in the late 1970s, I gave a talk that contrasted subjective perceptions with objective data. Afterward, Karl Weick asked me: 'What if there are no objective data?' I found this a puzzling, almost incomprehensible question. But I have great respect for Karl, and I began to experiment with interpreting supposedly 'objective' data as arising from mental or social processes.

Then, in 1981, I read a manuscript in which Nils Brunsson (1982) argued that a perception held by only one person has the status of being subjective, and its effects are limited to that person's actions. On the other hand, he said, a widely shared perception acquires the status of being 'objective'; not only can it affect the actions of many people, but the actions of these have the support of objective fact.

Around that same time, Meyer and Rowan (1977) interpreted the administrative structures of schools as having negligible effects on what happens in classrooms, but instead reassuring taxpayers that their schools are being managed responsibly. Of course, if administrative structures have weak effects on classroom activities, they do not really influence the degrees to which students or teachers behave responsibly. In this interpretation, the administrative structures of schools basically create false impressions about what goes on in schools. Likewise, said Meyer and Rowan, hospitals and governmental agencies increase their chances of survival by mirroring rules valued by their societies. This notion led Paul Nystrom and me (1984) to survey a variety of ways in which business managers create facades. We argued that business managers conform to rules that their environments cherish and such conformity may produce either desirable or undesirable consequences. Among these consequences has been some rather silly research by organization theorists who failed to recognize the superficiality of some behaviors or organizational properties.

When reporting my study of a very successful law firm (Starbuck 1993b), I described it mainly through quotations from the lawyers, their secretaries, their clients, and their competitors. This description demonstrated, to me at least, that various participants in and observers of the law firm see different organizations. This ambiguity arises not merely from the perceivers' different viewpoints, but also the law firm's complexity and internal inconsistencies. The law firm truly does have properties that appear logically contradictory to some. Moreover, the firm has been simultaneously a reflection of its environment and its historical era and an influential shaper of its environment and its era. In other words, the law firm has been an important component of its environment.

Quite recently, in a history of the origins of organization theory (Starbuck 2003a), I investigated the origins of the term 'organization'. Romans gradually converted this term from medical usage and gave it the broader meaning of 'to endow with a coordinated structure'. Then around 1800, some writers began to discuss the 'organization' of societies, organization being contrasted with unfettered individualism. Late in the 19th century, some people used this term to denote voluntary associations at universities formed for athletic, intellectual, literary, religious, or social purposes. For a couple of decades, people contrasted such voluntary 'organizations' with social systems such as armies, churches, companies, and governmental bodies. In other words, before 1930 approximately, people did not perceive armies and churches as belonging to a single general category called 'organizations'.

I speculate that generalizations about organizations resulted from social and technological changes during the last half of the 19th century and first half of the 20th century. Changes in education, occupational and task specialization, and technologies caused a sudden increase in the numbers of large, formalized organizations, they made organizations relevant to many more people, and they made many more people interested in and capable of understanding abstract generalizations. Legal concepts also evolved and endowed corporations with a 'personhood' that confers legal rights independent of the rights of their stakeholders. A distinct legal entity has to have definitive boundaries. Thus, both the similarity and distinctness of 'organizations' are social constructions that reflect large-scale social and technological changes.

Potential Benefits from Natural Experiments and Explicit Designs

My chapter in Dunnette's handbook also began my advocacy of design efforts, by which I mean explicit attempts to change situations for the better. As I had read more and more social science research, I had grown increasingly skeptical that it was yielding useful knowledge. My adventures with opportunities for self-deception had convinced me that human minds have the ability to generate seemingly credible explanations for virtually any data, including data that were nonsense. Yet almost all social science studies were generating retrospective explanations of previous events that had occurred spontaneously. Both retrospection and spontaneity of data invite agile minds to deceive themselves.

Triangulation (investigating a situation with more than one type of data) offers mild protection against self-deception. For example, Sutton and Rafaeli (1988) examined relationships between sales volume and the emotions that employees display to customers. A chain of convenience stores attempted to gain a competitive advantage by persuading employees to smile and act friendly toward customers. Then the company observed surreptitiously the behavior of 576 employees: the desired behaviors correlated negatively with store sales; where employees were smiling, stores had lower sales. So Sutton and Rafaeli gathered qualitative data through interviews and through working in stores. They inferred that store sales reflect the flows of customers though stores; where there are many customers, stores have high sales, but employees have no time to exchange pleasantries with customers. Thus, surprising findings from statistical analyses became understandable through direct observation. However, triangulation may be like the blind men studying not an elephant, but six different animals. In particular, triangulation seems always to involve different levels of analysis, for instance individual employees talking to individual customers versus sales by a store over several months and observations about many employees. An average of many instances may describe very few of them, possibly none. A correlation across a population may occur in none of the subpopulations. A true statement about a population may be false for every member of the population.

In the 1970s, I could see five reasons why it is important for the social sciences to start diverging from a passive-retrospective mode (also see the Appendix). First, researchers have weak incentives to eliminate poor theories when poor theories exact little cost. Second, because retrospective theories are consistent with the prominent stylized facts, they all appear to perform well. To expose differences between theories, prediction must replace retrospection. Of course, there are also many problems with prediction, including the possibility of making accurate predictions on the basis of erroneous assumptions, but the problems associated with retrospection appear insurmountable. Third, whereas scientific disciplines develop social structures and codes of behavior that stifle innovation and progress, active design efforts provoke reactions that can inject new ideas. Fourth, spontaneous phenomena produce data dominated by uninteresting events — nearly every adult has brown eyes, nearly all rock formations are stable, and nearly all prices are the same as last week. To acquire data that sharpen comparisons among theories, researchers must exert some control over what they observe. Lastly, the passive-retrospective studies understate the potentialities of flexible, adaptive, and reactive systems. Because systems are almost always close to their equilibria, they do not have to display the capabilities that they would have when displaced from their equilibria.

Thus, I began urging myself and my colleagues to search for natural experiments and to become engaged in efforts to improve social systems. Natural experiments occur when exogenous events displace social systems from their normal equilibria. In these situations, one can see some of the systems' adaptive and reactive capabilities, which opens the possibility of discovering why equilibria exist. When researchers attempt to improve social systems, they must acknowledge the values guiding their proposals, use their theories to predict outcomes, and revise their theories when the predicted outcomes do not occur.

A conversation with a medical doctor reinforced my conviction that design efforts would strengthen the social sciences generally, and management and organization theory more specifically (Starbuck 1993a). The doctor's views basically challenged the conventional idea that one needs to understand a system before one dares to try to change it. Indeed, the doctor's views, which I later adopted as my own, suggest that to understand a system, one must try to change it and observe how it reacts. What follows approximates my memory of our conversation, which is, I am sure, not what we said.

I told this doctor that I had been trying to create a computer program to make medical diagnoses because I wanted to improve medical care. He responded, 'But, diagnosis is not important to *good* medical care ... *Good* doctors do not rely on diagnoses.'

'But, medical schools teach doctors to translate symptoms into diagnoses, and then to base treatments on diagnoses,' I protested.

'That's right. Medical schools do teach that,' he conceded, 'but the doctors who do what they were taught never become good doctors. There are many more combinations of symptoms than there are diagnoses, so translating symptoms into diagnoses discards information. And there are many more treatments than diagnoses, so basing treatments on diagnoses adds random errors. Doctors can make more dependable links between symptoms and treatments if they leave diagnoses out of the chain.

'However, the links between symptoms and treatments are not the most important keys to finding effective treatments. Good doctors pay careful attention to how patients respond to treatments. If a patient gets better, current treatments are heading in the right direction. But, current treatments often do not work, or they produce side-effects that require correction. The model of symptoms–diagnoses–treatments ignores the feedback loop from treatments to symptoms, whereas this feedback loop is the most important factor.

'Doctors should not take diagnoses seriously because strong expectations can keep them from noticing important reactions. Of course, over time, sequences of treatments and their effects produce evidence that may lead to valid diagnoses.'

Sometime in the late 1970s or early 1980s, I became aware of Pygmalion effects, in which predictions affect outcomes. Predictions may become either self-fulfilling or self-denying. These effects weaken even further the usefulness of retrospective research. Although explaining the past may reassure us and comfort us, it may do little to help us influence our futures. These effects also confront us with the issue of what realities we wish to understand — the ones that did exist when we gathered the data or the ones that might exist after we attempt to exert influence.

So now I advocate design in the belief that efforts to design better organizations can manufacture both greater understanding and better realities. The systems we are trying to understand are much more complex and flexible than prevalent research methods (rooted in spontaneous data and static analyses) are capable of comprehending. The phenomena that I once called 'realities' are and ought to be partly products of our research because to obtain useful understanding of these phenomena, we must attempt to change them.

There is also a possibility that we might help to create a better world.

Appendix: The Marketing Science Institute

I speculate that my advocacy of organization design is too large a pill for many to swallow. There is a way in which management and organizational research could be strengthened more incrementally.

In 1961, Wharton's dean asked several business firms to become ongoing members of a new organization. This new organization, the Marketing Science Institute (MSI), would encourage academic research in marketing that related to contemporary managerial problems and have potential for application.

MSI has had ups and downs over the years. It moved from Wharton to Harvard, and later established separate offices. It incorporated, expanded, and then spun off the PIMS database project. Its membership has fluctuated.

However, during recent years, MSI has had more than 60 member firms and its influence in the field of marketing has been remarkable. Over a recent 10-year period, projects that MSI sponsored have won every award for outstanding research in marketing and they comprised 60 percent of the articles in the *Journal of Marketing* and the *Journal of Marketing Research*. Probably more important, MSI has sponsored two-way dialogues between executives and professors. MSI's premise is that theory and practice should reinforce each other: good theory leads to good practice and good practice leads to good theory.

MSI's member firms identify research priorities, which it circulates to 2,000 academic researchers. Approved research projects receive modest financial support. Around 30–35 new projects begin each year, so about 90 projects are under way at any time. Because the projects generate written reports that tend to be technical and academic, MSI restates the reports as executive summaries that draw out their implications for practice. MSI's mailings reach around 2,000 professors and 2,000 executives. Member-only conferences and implementation workshops also discuss research findings and explore their practical implications. A typical conference involves 40 executives and 20 professors and emphasizes dialogue among the participants.

Member firms receive early access to cost-effective research. They have opportunities to participate in research projects that will be highly visible to professors and students. MSI's meetings and reports pinpoint major trends and new ideas, and they involve the leading academic and executive thinkers in marketing. Thus, MSI fosters the professional development of executives.

MSI influences the careers and attitudes of professors by sponsoring research, by making awards for the 'best dissertation' and for the 'best paper' published in the *Journal of Marketing*, and by creating interaction between professors and executives. MSI sponsorship helps research projects to gain access to firms and to obtain higher quality data.

My own experiences suggest that the goals of nearly all researchers can be rendered compatible with the goals of business firms that serve as research sites. For example, John Mezias and I wanted realistic data about the accuracy of managers' perceptions (Mezias and Starbuck 2003). We approached a senior corporate executive in one of the world's largest companies, and after eight months of negotiations, we gained the company's support for our project. The senior corporate executive said the company's top priority was quality improvement and we could gather data if the data would tell the company how it was doing in that domain. Had we been designing a study without concern for its relevance to anyone else, we would not have chosen quality improvement as the target subject. But managers have perceptions in this domain and the company was spending a lot of resources trying to measure quality, so we would have access to good measures of 'objective reality' to compare with the managers' perceptions. Personnel in each of four large divisions helped us to design questionnaires that suited the managers in their divisions. The senior corporate executive personally delivered our questionnaires to the top managers in these four divisions; we had a 100 percent response rate, and the respondents completed our questionnaires themselves instead of delegating them to their secretaries.

I spent more than two years investigating the feasibility of an institute similar to MSI to foster applied research in management. I found that many senior scholars are willing to contribute time to such an enterprise, and a large number of senior executives were willing to appropriate US$25,000 per year for such a purpose. Few executives expressed interest in sponsoring research as such; nearly all new ideas about how to manage arise from actual practice, as do the challenges that business people confront. But executives were willing to sponsor research as a way to influence education. Executives understood my argument that the best way to get current and useful content into courses is to facilitate good research on contemporary issues. The findings of such research go into textbooks and teaching cases and they set the priorities for curricula. Also, executives are looking for inexpensive, time-efficient ways to obtain professional development and continuing education for themselves and their colleagues. Business people do not have enough time to keep up with the journals and books that constantly pass across their desks, yet professional development and intellectual rejuvenation are critical to personal and organizational performance. Seminars and two-way conversations with professors about the results of research can serve this purpose.

Note

This article has benefited from questions and useful suggestions from Hari Tsoukas, Jane Webster, Joan Dunbar, John Dutton, Raghu Garud, and Roger Dunbar.

References

Ames, Edward, and Stanley Reiter
1961 'Distributions of correlation coefficients in economic time series'. *Journal of the American Statistical Association* 56: 637–656.

Bonini, Charles P.
1963 *Simulation of information and decision systems in the firm*. Englewood Cliffs, NJ: Prentice Hall.

Brunsson, Nils
1982 'The irrationality of action and action rationality: Decisions, ideologies, and organisational actions'. *Journal of Management Studies* 19: 29–44.

Campbell, John P., Richard L. Daft, and Charles L. Hulin
1982 *What to study: Generating and developing research questions*. Beverly Hills, CA: Sage.

Faust, David
1984 *The limits of scientific reasoning*. Minneapolis: University of Minnesota Press.

Fitter, Mike J., and Max E. Sime
1980 'Creating responsive computers: Responsibility and shared decision making' in *Human interaction with computers*. H. T. Smith and Thomas R. G. Green (eds), 39–66. London: Academic Press.

Fromkin, Howard L.
1969 'The behavioral science laboratories at Purdue's Krannert School'. *Administrative Science Quarterly* 14: 171–177.

Garvey, William D., Nan Lin, and Carnot E. Nelson
1970 'Some comparisons of communication activities in the physical and social sciences' in *Communications among scientists and engineers*. Carnot E. Nelson and Donald K. Pollock (eds), 61–84. Lexington, MA: Heath Lexington.

Gottfredson, Stephen D.
1978 'Evaluating psychological research reports: Dimensions, reliability, and correlates of quality judgments'. *American Psychologist* 33/10: 920–934.

Gottfredson, Don M., and Stephen D. Gottfredson
1982 'Criminal justice and (reviewer) behavior: How to get papers published'. *Criminal Justice and Behavior* 9/3: 259–272.

Green, Thomas R. G.
1982 'Pictures of programs and other processes, or how to do things with lines'. *Behaviour and Information Technology* 1: 3–36.

Green, Thomas R. G., Max E. Sime, and Mike J. Fitter
1980 'The problems the programmer faces'. *Ergonomics* 23: 893–907.

Grinyer, Peter H., and David Norburn
1975 'Planning for existing markets: Perceptions of executives and financial performance'. *Journal of the Royal Statistical Society, Series A* 138: 70–97.

Hayek, F. A. von
1975 *The pretence of knowledge*. Stockholm: Nobel Foundation.

Hubbard, Raymond, and J. Scott Armstrong
1992 'Are null results becoming an endangered species in marketing?' *Marketing Letters* 3/2: 127–136.

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

McClain, David
1981 'Understanding model acceptance: For many users, "small is beautiful"'. *Simulation* 36/9: ix–xii.

Mahoney, Michael J.
1977 'Publication prejudices: An experimental study of confirmatory bias in the peer review system'. *Cognitive Therapy and Research* 1: 161–175.

Mahoney, Michael J.
1979 'Psychology of the scientist: An evaluative review'. *Social Studies of Science* 9/3: 349–375.

March, James G., and Herbert A. Simon
1958 *Organizations*. New York: Wiley.

Meyer, John W., and Brian Rowan
1977 'Institutionalized organizations: Formal structure as myth and ceremony'. *American Journal of Sociology* 83/2: 340–363.

Mezias, John M., and William H. Starbuck
2003 'Studying the accuracy of managers' perceptions: A research odyssey'. *British Journal of Management* 14: 3–17.

Nylenna, Magne, Povl Riis, and Yngve Karlsson
1994 'Multiple blinded reviews of the same two manuscripts: Effects of referee characteristics and publication language'. *JAMA* 272: 149–151.

Nystrom, Paul C., and William H. Starbuck
1984 'Organizational facades'. *Academy of Management, Proceedings of the Annual Meeting, Boston*: 182–185.

Payne, Roy L., and Derek S. Pugh
1976 'Organizational structure and climate' in *Handbook of industrial and organizational psychology*. Marvin D. Dunnette (ed.), 1125–1173. Chicago, IL: Rand McNally.

Peters, Douglas P., and Stephen J. Ceci
1982 'Peer-review practices of psychological journals: The fate of published articles, submitted again'. *Behavioral and Brain Sciences* 5: 187–255.

Pfeffer, Jeffrey
1993 'Barriers to the advance of organization science: Paradigm development as a dependent variable'. *Academy of Management Review* 18: 599–620.

Simon, Herbert A.
1950 'Modern organization theories'. *Advanced Management* 15/10: 2–4.

Simon, Herbert A.
1952 'Comments on the theory of organizations'. *American Political Science Review* 46: 1130–1139.

Starbuck, William H.
1973 'Tadpoles into Armageddon and Chrysler into butterflies'. *Social Science Research* 2: 81–109.

Starbuck, William H.
1976 'Organizations and their environments' in *Handbook of industrial and organizational psychology*. Marvin D. Dunnette (ed.), 1069–1123. Chicago, IL: Rand-McNally.

Starbuck, William H.
1983 'Computer simulation of human behavior'. *Behavioral Science* 28: 154–165.

Starbuck, William H.
1993a '"Watch where you step!" or Indiana Starbuck amid the perils of Academe (rated PG)' in *Management laureates*, Vol. 3. Arthur Bedeian (ed.), 63–110. Greenwich, CT: JAI Press.

Starbuck, William H.
1993b 'Keeping a butterfly and an elephant in a house of cards: The elements of exceptional success'. *Journal of Management Studies* 30: 885–921.

Starbuck, William H.
2003a 'The origins of organization theory' in *The Oxford handbook of organization theory: Meta-theoretical perspectives*. Haridimos Tsoukas and Christian Knudsen (eds), 143–182. Oxford: Oxford University Press.

Starbuck, William H.
2003b 'Turning lemons into lemonade: Where is the value in peer reviews? *Journal of Management Inquiry* 12: 344–351.

Starbuck, William H., and John M. Dutton
1971 'The history of simulation models' in *Computer simulation of human behavior*. John M. Dutton and William H. Starbuck (eds), 9–102. New York: Wiley.

Starbuck, William H., Arent Greve, and Bo L. T. Hedberg
1978 'Responding to crises'. *Journal of Business Administration* 9/2: 111–137.

Sutton, Robert I., and Anat Rafaeli
1988 'Untangling the relationship between displayed emotions and organizational sales: The case of convenience stores'. *Academy of Management Journal* 31: 461–487.

Thibault, John W., and Harold H. Kelley
1954 'Experimental studies of group problem solving and process' in *Handbook of social psychology, Vol. 2*. Gardner Lindzey (ed.), 735–785. Reading, MA: Addison Wesley.

Thune, Stanley S., and Robert J. House
1970 'Where long-range planning pays off: Some research findings'. *Business Horizons* 13/4: 81–87.

Webster, E. Jane, and William H. Starbuck
1988 'Theory building in industrial and organizational psychology' in *International review of industrial and organizational psychology 1988*. Cary L. Cooper and Ivan T. Robertson (eds), 93–138. London: Wiley.

Wolff, W. M.
1970 'A study of criteria for journal manuscripts'. *American Psychologist* 25: 636–639.

William H. Starbuck

William H. Starbuck is the ITT Professor of Creative Management at the Stern School of Business, New York University. He currently serves on the editorial boards of the *Asian Case Research Journal*, *British Journal of Management*, *Information and Organization*, *International Journal of Management Reviews*, *Journal of Leadership Studies*, *Journal of Management Inquiry*, *Journal of Management Studies*, *Journal of Socioeconomics*, *Knowledge Management and Information Studies*, *Organization*, *Organization Management Journal*, and the *Scandinavian Journal of Management*. He has published numerous articles on accounting, bargaining, business strategy, computer programming and simulation, forecasting, decision-making, organizational design, organizational growth and development, scientific methods, and social revolutions. He has also edited four books, including the *Handbook of Organizational Design*.
Address: Stern School of Business, New York University, 40 West Fourth Street, New York, NY 10012-1118, USA.
E-mail: wstarbuc@bellatlantic.net