

CLASSIC PAPER

Learning from samples of one or fewer

J G March, L S Sproull, M Tamuz

Qual Saf Health Care 2003;12:465-472

Organizations learn from experience. Sometimes, however, history is not generous with experience. We explore how organizations convert infrequent events into interpretations of history, and how they balance the need to achieve agreement on interpretations with the need to interpret history correctly. We ask what methods are used, what problems are involved, and what improvements might be made. Although the methods we observe are not guaranteed to lead to consistent agreement on interpretations, valid knowledge, improved organizational performance, or organizational survival, they provide possible insights into the possibilities for and problems of learning from fragments of history.

Organizations learn from experience, but learning seems problematic when history offers only meager samples of experience. Historical events are observed, and inferences about historical processes are formed, but the paucity of historical events conspires against effective learning. We consider situations in which organizations seek to learn from history on the basis of very small samples of experience. For example:

- *Case 1.* A military organization has rarely fought in a battle. Yet it wants to learn from its history how to improve its ability to engage in warfare.
- *Case 2.* A business firm has little experience with foreign acquisitions. Yet it wants to learn from its history whether and how to make such investments.
- *Case 3.* An airline rarely has fatal accidents. Yet it wants to learn from its history how to reduce the chances of such disasters.
- *Case 4.* A business firm rarely makes major marketable discoveries. Yet it wants to learn from its history how to increase the chances of such innovations.
- *Case 5.* A power company rarely has nuclear accidents. Yet it wants to learn from its history how to minimize the chances of such catastrophes.

In this paper we examine (1) how organizations convert meager experience into interpretations of history by experiencing infrequent events richly; (2) processes for simulating hypothetical histories; and (3) some justifications for these two learning strategies and some of the problems involved.

EXPERIENCING HISTORY RICHLY

Historical events are unique enough to make accumulating knowledge difficult. Each event is a single unrepeated data point, and accumulation seems to require pooling across diverse contexts. Organizations attempt such pooling, but they also seek to increase the information extracted from their own limited historical experience by treating unique historical incidents as detailed stories rather than single data points. They elaborate experience by discovering more aspects of experience, more interpretations of experience, and more preferences by which to evaluate experience.

Experiencing more aspects of experience

Characterizing history as small samples of unique occurrences overlooks the wealth of experience that is represented in each historical event. The apparent stinginess of history is moderated by attending to more aspects of experience (Campbell, 1979). For example, learning about a decision involves monitoring its outcomes. But long before an organization experiences many of the outcomes of a typical decision, it experiences a variety of collateral consequences associated with the making of the decision and its implementation. Learning and evaluation occur through these experiences prior to outcome based learning. For example, participants appreciate collateral experiences such as "a bold move" or "a good meeting".

When early collateral experiences are positive, organizations, like individuals, are prone to exhibit self-reinforcing decision behavior. Especially when outcome feedback is slow or unclear, an organization is likely to repeat decisions simply because it has made them before. Thus, in a study of decisions about foreign direct investments by Finnish firms, Björkmann (1989) found that, prior to receiving information on the results of their first investments and simply as a result of their experience in making the first decisions, firms increased their propensity to make more investments. This did not appear to be a consequence of any explicit intention to spread risk. Apparently, the organizations extracted lessons from the choice process itself, lessons about the competence and character of relevant actors and the pleasure of deciding to invest. Learning was embedded richly

*This is a reprint of a paper that appeared in *Organization Science*, 1991, Volume 2, pages 1-13. © 1991, The Institute of Management Sciences (now INFORMS), 901 Elkridge Landing Road, Suite 400, Linthicum, MD 21090, USA.

See end of article for authors' affiliations

in the taking of action, rather than simply in considering its ultimate consequences.

Making a decision also induces anticipations of its future costs and benefits (Merton, 1968). The anticipations are experienced prior to the consequences and are an independent basis for learning. Since expectations for chosen alternatives will generally be positive, they ordinarily reinforce a repetition of the action. On the other hand, negative expectations might be experienced in situations involving coercion or peril, or when learning takes place in an alienated part of the organization, or among opposition groups. In such cases, the effect of making a decision reduces the propensity to repeat it. In general, decision processes in organizations lead to overly optimistic expectations and thus are vulnerable to subsequent disappointment with results (Harrison and March, 1984). These optimistic errors in anticipations are likely to make the short-run lessons of experience more reinforcing of action than the long-run lessons. The inconsistency in learning is reduced by the tendency for actual experience to be both delayed and more ambiguous than anticipations, thus allowing optimistic expectations about experience to be confirmed by retrospective sense making of it (Aronson, 1968; Salancik, 1977).

Organizations also enhance the richness of history by focusing intensively on critical incidents. For example, when a large section of the metal skin of an Aloha Airlines aircraft peeled away in mid air, the event attracted considerable attention and triggered a major modification in FAA mandated maintenance programs. Close examination of what happened revealed significant features of aircraft engineering and maintenance that had not been noted earlier. By identifying those features and their implications, the organizations learned. Similarly, when a computer science graduate student propagated a "virus" among many computer networks, producing breakdowns in hundreds of systems and considerable publicity, the incident stimulated analyses that identified weaknesses in the underlying computer code and in how people and systems were organized to respond to such events.

Three aspects of an event seem to make it critical. The first is its place in the course of history. Events that change the world are critical incidents. They are branching points of historical development. From such an incident, one learns about changed implications for the future rather than about how to predict or control similar occurrences in the future. A classic example is the invention of the printing press.

The second aspect of an event that makes it critical is its place in the development of belief. Events that change what is believed about the world are critical incidents. In a way consistent with conventional ideas about the relation between surprise and information value (Raiffa, 1968), criticality is associated with the surprise an event evokes for current belief. A single incident is typically unsurprising because it can be interpreted as consistent with sampling variation within existing theories (Fischhoff, 1975; Fischhoff and Beyth, 1975). But sometimes a solitary event provides an unexpected contradiction to our beliefs as, for example, in the Aloha Airlines and computer virus cases.

A third aspect that makes an event critical is its metaphorical power. Events that evoke meaning, interest, and attention for organizational participants are critical incidents. Anecdotes and stories are standard features of pedagogical practice. Skill at story telling is a major factor in endowing experience with metaphorical force. But the raw material of experience also affects the development of stories. Critical incidents have a quality of simplicity and representativeness that is not entirely imposed on them. Some historical events are better vehicles for meaning than others.

Experiencing more interpretations

Organizations often augment history by attending to multiple observers or interpretations. The consequences of an action are experienced differently throughout the organization. Conflicts of interest or differences in culture, in particular, stimulate multiple interpretations. Because different individuals and groups experience historical events differently, they learn different lessons from the same experience (Dearborn and Simon, 1958; Sproull and Hofmeister, 1986). As a result, organizational experience leads to a variety of interpretations, and an organization's repertoire may come to include several different, possibly contradictory, story lines. Differences in perspectives lead to differences in interpretations and create a mosaic of conflicting lessons.

To be sure, efforts to make multiple interpretations consistent are also routine in organizations. Formal proceedings, findings, informal conversations, and the diffusion of stories tend to create a shared, interpretive history. Interpretations of individual or group responsibility for mistakes or failures (or brilliant moves or successes) come to be shared. However, such efforts are not always successful (Brunsson, 1989). The structure of internal competition and conflict divides many organizations into advocates and opponents for organizational policies and actions. The contending groups interpret history differently and draw different lessons from it.

Experiencing more preferences

Organizations discover values, aspirations, and identities in the process of experiencing the consequences of their actions. They learn how to distinguish success from failure, and thus affect considerably the other lessons they take from their experience. While the interpretation of a particular outcome as a success or failure is not arbitrary, neither is it always self-evident. The preferences and values in terms of which organizations distinguish successes from failures are themselves transformed in the process of learning. By acting, reflecting, and interpreting, organizations learn what they are. By observing their own actions, they learn what they want (Weick, 1979). Whether these changes are seen as learning new implications of alternative actions for stable preferences, or as transforming preferences, is partly a matter of intellectual taste (Becker and Stigler, 1977).

For any given dimension of organizational preferences, aspiration levels change in response to an organization's own experience and to the experience of other organizations to which it compares itself (Cyert and March; Lant and Montgomery, 1987), thereby loosening the link between performance and outcomes, on the one hand, and evaluations of success and failure (and thus learning), on the other. Because experiencing an outcome as a success or failure depends on the relation between the outcome and adaptive aspirations for it, what is learned from any particular kind of experience can vary substantially across time and across organizations. Consider, for example, the efforts of a business firm to learn from its marketing experience. Whether a particular marketing strategy is viewed as a success (to be reinforced) or a failure (to be extinguished) will depend as much on the organization's aspirations as on the marketing outcomes. In a similar way, election results are experienced by political parties and movements in terms of a comparison between outcomes and aspirations rather than simply in terms of the former.

The dimensions of preferences also change. As we have noted above, a rich examination of an individual case uncovers a variety of features and consequences of action. These experiences become bases for organizational interpretations not only of the world and its rewards but also of the organization, particularly its preferences, values, and

character (Zald and Denton, 1963). For example, when the Coca-Cola Company reinstated old coke as Coca-Cola Classic (after having first withdrawn it from the market in favor of a replacement), it learned from its behavior that it was a “flexible company that listened to its customers” (Oliver, 1986). Anything that predisposes parts of an organization to find pleasure in consequences, for example, an upbeat mood (Isen *et al.*, 1978) or a sense of responsibility for action (March and Olsen, 1975) tends to increase the likelihood of identifying positive aspects of unanticipated consequences, thus of transforming preferences.

SIMULATING EXPERIENCE

In trying to understand unique experiences, organizations make implicit choices between two alternative perspectives on history. In the first perspective, realized events are seen as necessary consequences of antecedent historical conditions. In the second perspective, realized events are seen as draws from a distribution of possible events. If historical events are (possibly unlikely) draws from a wealth of possibilities, an understanding of history requires attention to the whole distribution of possible events, including those that did not occur (Fischhoff, 1980; Hogarth, 1983). The organizational problem is to infer an underlying distribution of possible events from a series of realized events having varying, but possibly quite low, probabilities. The merging of empirical and theoretical knowledge required to understand these underlying distributions raises some complicated problems of inference and method.

We consider two closely related techniques for organizational simulation of hypothetical events: the first technique is to define and elaborate a class of historical non-events that can be called near-histories—events that almost happened. The second technique is to define and elaborate a class of historical non-events that can be called hypothetical histories—events that might have happened under certain unrealized but plausible conditions.

Near-histories

If a basketball game is decided by one point, one team wins and the other team loses, with consequences that may be vital for a championship. But the outcome will normally be interpreted by experts as a draw from some probability distribution over possible outcomes rather than simply as a “win” by one team and a “loss” by the other. In general, if a relatively small change in some conditions would have transformed one outcome into another, the former will be experienced to some degree as having been the latter. In such a spirit, the National Research Council (1980) has defined a safety “incident” as an event that, under slightly different circumstances, could have been an accident.

Air traffic systems illustrate how organizations learn from near-histories or “incidents” (Tamuz, 1987). By collecting information about near-accidents from pilots and air traffic controllers, air safety reporting systems considerably enlarge the sample of events that can be treated as relevant to understanding aviation safety. Information on near-accidents augments the relatively sparse history of real accidents and has been used to redesign aircraft, air traffic control systems, airports, cockpit routines, and pilot training procedures.

Near-histories are useful antidotes to a tendency to overgeneralize from the drama of great disasters or victories (Fischhoff, 1982). For example, students of the Battle of Midway have suggested a number of quite likely alternative scenarios for that battle that would have led to notably different outcomes (Prange, 1982). Future admirals learn not only from the battle but also from its near-histories. Standard folkloric observations that great failures often are the consequence of bad luck or timing and great successes the

consequence of good luck or timing suggest an implicit distribution of possible outcomes around the observed outliers. They emphasize that the near-histories of genius and foolishness are more similar than their realized histories.

Hypothetical histories

Near-histories are a special case of a more general approach—the construction of hypothetical histories. Hypothetical histories play a role in organizational learning similar to that of mental models or simulations in studies of individual learning (Kahneman and Tversky, 1982; Johnson-Laird, 1983). Organizations use small samples of specific historical events to construct theories about events, and then simulate hypothetical histories that can be treated as having interpretive significance comparable to, or even greater than, the history actually experienced. In this process, the analysis of unique historical outcomes emphasizes identifying the underlying distribution from which that realization was drawn rather than explaining the particular draw (Stinchcombe, 1978).

A pervasive contemporary version of hypothetical histories is found in the use of spread sheets to explore the implications of alternative assumptions or shifts in variables in a system of equations that portrays organizational relations. More generally, many modern techniques of planning in organizations involve the simulation of hypothetical future scenarios, which in the present terms are indistinguishable from hypothetical histories (Hax and Majluf, 1984). The logic is simple: small pieces of experience are used to construct a theory of history from which a variety of unrealized, but possible, additional scenarios are generated. In this way, ideas about historical processes drawn from detailed case studies are used to develop distributions of possible futures.

Organizations expand their comprehension of history by making experience richer, by considering multiple interpretations of experience, by using experience to discover and modify their preferences, and by simulating near-events and hypothetical histories. They try to learn from samples of one or fewer. Many of the techniques organizations use to learn from small samples of history, however, are clearly suspect. They can lead to learning false lessons, to superstitious learning, to exaggerated confidence in historical understandings. As a result, discussions of organizational learning from small samples tend to be framed by a mood of despair over the futility of the effort. In the next section we turn to an examination of how such learning might be understood, evaluated, and improved.

ASSESSING AND IMPROVING LEARNING FROM SMALL HISTORIES

In the previous sections we have described some of the ways organizations learn from fragments of history. In this section we assess their effectiveness in terms of two common criteria: reliability and validity. A reliable learning process is one by which an organization develops common understandings of its experience and makes its interpretations public, stable, and shared. A valid learning process is one by which an organization is able to understand, predict, and control its environment. Neither reliability nor validity is assured. Because different people and groups in an organization approach historical experience with different expectations and beliefs, shared understandings cannot be assumed. And because historical events are produced by particular (and often complicated) combinations of factors occurring in non-repetitive contexts, learning validly from small samples of historical experience is difficult (Brehmer, 1980; Kiesler and Sproull, 1982).

If individual beliefs converge to an accurate understanding of reality, then they become simultaneously shared and valid. Such a convergence might be expected in worlds of stable knowledge and cumulative discovery. Alternatively, if socially constructed beliefs are enacted into reality (Weick, 1979), the enactment brings high levels of both reliability and validity. For example, organizational beliefs about power, legitimacy, competence, and responsibility are based upon interpretations of shared experiences that are themselves considerably affected by the beliefs.

The more general situation, however, is one of partial conflict between the dual requirements of reliability and validity. Stable, shared knowledge interferes with the discovery of contrary experience from which valid learning arises, and the exploration of novel ideas interferes with the reliable maintenance and sharing of interpretations (March, 1990). As a result, organizational learning involves balancing the two. The trade-offs between reliability and validity made by the learning procedures we observe are less a result of conscious choice than a collection of evolving practices, imperfectly justified and incompletely comprehended. For example, the extent and character of subgroup differentiation in organizations and the mechanisms by which they are sustained or changed are important means by which the trade-off between reliability and validity is made.

The reliability of learning: the construction and sharing of belief

Stability in shared understandings is important for organizational effectiveness and survival (Beyer, 1981; March and Olsen, 1989), as it is for social systems (Durkheim, 1973) and knowledge in general (Rorty, 1985). However, the ambiguities of history make common understandings of organizational experience difficult to sustain. Meaning is not self-evident but must be constructed and shared. Many different interpretations are both supportable and refutable.

Some standard mechanisms of individual sense-making contribute to reliability. The retrieval of history from memory exaggerates the consistency of experience with prior conceptions (Fischhoff, 1975; Pearson *et al.*, 1990). Incorrect predictions are not noticed or are interpreted as irrelevant anomalies or measurement errors (Einhorn and Hogarth, 1978; Lord *et al.*, 1979). Missing data are experienced as consistent with the model and are remembered as real (Cohen and Ebbesen, 1979; Loftus, 1979).

The apparatus of organizational information processing and decision-making supplements these individual and social cognitive processes. Information is gathered and distributed more to interpret decisions than to inform them (Feldman and March, 1981; March and Sevon, 1984). Meetings are organized more to share stories and explanations than to take action (Brunsson, 1985, 1989). Organizations develop robust understandings that are resilient to contradictory information (Sproull *et al.*, 1978; Starbuck, 1983). Conceptions of identity in organizations tend also to be conserved by interpretations of experience. Decision makers in organizations discover what they are and how they should behave by taking actions (Weick, 1979). Making a decision leads to defining a personal and organizational identity consistent with that decision. Similarly, the social construction of aspirations tends to be conservative, to reinforce shared behavioral preferences.

Pressures toward reliability are easily orchestrated within an emphasis on critical incidents. Defining an event as critical focuses attention on interpreting and responding to the event. Because of the ambiguities associated with any single incident, responses and interpretations tend to be adopted more as a result of their temporal proximity, cognitive

availability, or political convenience than by virtue of their obvious validity (Cyert *et al.*, 1958; Cohen *et al.*, 1972).

The learning process is generally conservative, sustaining existing structures of belief, including existing differences, while coping with surprises in the unfolding of history. Organizations create the same kinds of coherent systems of belief that in science are called knowledge, in religion are called morality, and in other people's societies are called myths. Experience is used to strengthen and elaborate previously believed theories of life.

Such a description is, however, incomplete. There are limits to the conservation of belief (Martin and Siehl, 1983; Higgins and Bargh, 1987). Both success and failure contain the seeds of change. A persistent subjective sense of success leads to a sense of competence and a willingness to experiment (March and Shapira, 1987). A persistent subjective sense of failure produces instability in beliefs and disagreement among organizational participants with respect to both preferences and action (Sproull *et al.*, 1978). In addition, pressures toward subgroup homogeneity lead to internal differentiation and limit organizational homogeneity (Cangelosi and Dill, 1965; Lawrence and Lorsch, 1967).

The validity of learning: the construction of causal belief

The confusions of history often obscure what happened, why it happened, and how we should learn from it. The general problem is not simply one of eliminating known biases in historical interpretation by organizations. The experimental design and sample sizes of history are inadequate for making inferences about it (Lounamaa and March, 1987; Levitt and March, 1988). Estimation from historical events is subject to two major kinds of variability. The first stems from the fact that some of the processes by which history is produced may be genuinely stochastic. Understanding those processes requires approximating the underlying distributions from which a realized event is drawn. The expected error in estimation can be decreased by aggregating over several events, but history limits the number of comparable events. Lacking multiple events, organizations use whatever information they can extract from single cases to discern the historical processes that determine those underlying distributions.

The second kind of variability in estimation stems from the measurement and interpretation of historical events by observers. Measurement error, model misspecification, and system complexity introduce substantial noise into observations and interpretations. With large samples of events, organizations can tolerate a relatively large amount of noise, aggregating over events to extract a signal. With small samples, however, aggregation is a less powerful procedure. Two organizational responses are common. First, since variability in interpretation with respect to any information is partly a function of the effort expended in examining it, the expected error can be decreased through a more intense examination of the individual case. Second, since the processes of measurement and interpretation yield a distribution of possible observations, the expected error can be decreased by aggregating over multiple observers.

Most of the ways organizations increase the validity of learning from historical experience can be seen as reflecting such considerations. Organizations attempt to overcome the limitations in the experimental design and sample sizes of history by enhancing the knowledge they have. They attempt to experience history more richly, to formulate more interpretations of that experience, and to supplement history by experiencing more of the events that did not occur but could have.

Consider, first, efforts to experience history more richly. Every unique historical event is a collection of micro events, each of which can be experienced. In this sense, the learning potential of any historical event is indeterminate. Because both the scope of an event and the depth of its decomposition into elements are arbitrary, so also is the richness of experience. By considering additional aspects of experience and new dimensions of preferences, an organization expands the information gained from a particular case. The pursuit of rich experience, however, requires a method for absorbing detail without molding it. Great organizational histories, like great novels, are written, not by first constructing interpretations of events and then filling in the details, but by first identifying the details and allowing the interpretations to emerge from them. As a result, openness to a variety of (possibly irrelevant) dimensions of experience and preference is often more valuable than a clear prior model and unambiguous objectives (Maier, 1963; March, 1978, 1987).

Moving from rich experiences of history to valid inferences about history involves a logic that is not very well defined but is different from the logic of classical statistical inference. It assumes that the various micro events associated with an event are in some way interconnected. They are clearly not independent samples of some universe in the standard statistical sense. But they provide scraps of information about an underlying reality that cumulate, much the way various elements of a portrait cumulate to provide information about its subject.

Consider, second, efforts to interpret experience in more ways. Imagination in generating alternatives of interpretation reduces the standard confirmatory bias of experience, at the cost of also reducing the speed at which a correct interpretation is recognized and confirmed. Janis and Mann (1977) and George (1980) have each pointed out the advantages of pooling perceptions and judgments across individuals who interpret history differently. A similar argument is made into a methodological point by Allison (1971) and Neustadt and May (1986). Such observations suggest an important trade-off in attempts to improve the precision of estimates: an organization can opt to increase the number of events to be observed and interpreted or, alternatively, to increase the number and diversity of observers for a single event. Whether it is better to invest in additional events or in additional observers depends both on the relative cost of the two and on the relative magnitudes of the two sources of variability. But in pursuing an understanding of organizational history, greater reduction in uncertainty can often be achieved by pooling observers than by pooling events, particularly if the observers are relatively independent.

Consider, third, efforts to experience more of the events that did not happen. The presumption is that the processes of history are both more stable than their realized outcomes and more susceptible to understanding through rich descriptions. Near-histories and hypothetical histories produce distributions of unrealized, possible events. By treating events that did not occur as having significance similar to those that did occur, hypothetical histories exploit the information contained in rich descriptions of historical processes to provide a more judicious assessment of the probability distribution of future events.

In addition to providing a wider range of experience from which to draw, near-histories may be more easily interpreted than realized history. Tamuz (1988) suggests that understanding actual aviation accidents is heavily compromised by the legal and financial contexts which provide individual and organizational incentives for discovering particular self-interested interpretations. She argues that although the reporting of near-accidents is affected by publicity, politics,

and perceived reporting incentives, the analysis of these near-accidents often introduces fewer biases than those of accidents, thus producing understandings that are more consistent with broader social constructions and theories of evidence.

Although near-histories make useful contributions to learning, supplementing realized events with hypothetical ones introduces certain complications. First, constructing hypothetical histories can be expensive. Sometimes the substantial costs of such activities are shared by professional associations and governmental agencies, as they are in the air transportation industry. But often the ordinary branchings of history make it difficult to gather and interpret information on consequences of hypothetical histories that are not immediate. Imagine an organization that wanted to compare the ultimate careers and productivity of its employees with those job applicants it almost hired, or of applicants to whom it offered employment but who chose to work elsewhere. Assembling information on such a collection of historical branches involving outsiders, and interpreting the information, are substantial tasks.

Second, the impact of hypothetical histories ordinarily cannot compare with the dramatic power of realized history (Fischhoff and Beyth, 1975; Kiesler and Sproull, 1982). It is difficult to match the powerful effect of actual events (for example, the 1987 Challenger explosion) on beliefs. As a result, a vital part of the telling of history is the evocation of imagination (Tolstoy, 1869). The probable dependence of imagination on vividness (Shedler and Manis, 1986) and rich detail (Krieger, 1983) provides at least a partial reason for emphasizing such stories in organizations (Clark, 1972; Martin, 1982). In the stories of Three-Mile Island, the Aloha Airlines flight, and the Cornell computer virus, vivid historical events were used to dramatize a hypothetical story of even greater potential disaster. The drama mobilized attention and learning across a wide spectrum of groups.

Third, hypothetical histories may be ambiguous and thus unpersuasive. Where organizations face possible events of great consequence but small likelihood, the use of near-histories to augment simple experience is sometimes controversial. If the probability of disaster is very low, near-histories will tend to picture greater risk than will be experienced directly by most organizations or individuals in a reasonable length of time. In such case, near-histories are likely to be treated as generating too pessimistic a picture. For example, long before the fatal Challenger flight, the spacecraft flew a series of successful missions despite its faulty O-rings. Some engineers interpreted the indications of O-ring problems during these early flights as symptoms that past successes had been relatively lucky draws from a distribution in which the probability of disaster was relatively high (Bell and Esch, 1987; Boisjoly, 1987). Others, including some key personnel in NASA, considered these estimates of danger as exaggerated because, in the realized history, the system had been robust enough to tolerate such problems (Starbuck and Milliken, 1988).

Conversely, if the probability of success is very low, most short sequences of realized experience will contain no successes. The direct experience of most organizations and individuals with projects offering very low probability of very high return will be less favorable than will an analysis of near-successes. In such cases, near-histories are likely to be treated as providing too optimistic a picture. One such case involves organizations searching unsuccessfully for major innovations and treating assertions of "near-discoveries" as an unduly optimistic basis for sustaining investment in research.

As these examples suggest, the most obvious learning problem with near-histories is the necessary ambiguity of

their interpretation. If an organization is concerned with product quality and uses an inspection system to reject items that do not meet standards, every rejected item provides information on two things—the likelihood of substandard production and the likelihood of discovering the inadequacy. Each event, therefore, is both a failure and a success. Similarly, every time a pilot avoids a collision, the event provides evidence both for the threat and for its irrelevance. It is not clear whether the learning should emphasize how close the organization came to a disaster, thus the reality of danger in the guise of safety, or the fact that disaster was avoided, thus the reality of safety in the guise of danger.

FOUR QUESTIONS

Organizational efforts to learn reliably and validly from small histories are marked by two conspicuous things: first, we try to learn from them, often believing that we do so (Allison, 1971; George and McKeown, 1985), or can do so (Fischhoff, 1982; Kahneman and Tversky, 1982), and often believing that we do not and cannot (Fischhoff, 1980; Dawes *et al.*, 1989). Second, we do not have a shared conception of how we learn from small histories or what distinguishes single cases that are informative from those that are not (Herbst, 1970; Mohr, 1985).

We have not invented a general logic for learning from history that can fully rationalize what we have described, nor do we imagine that such a total rationalization is possible. Many of the ways in which organizations treat small histories are difficult to justify as either leading to shared beliefs, exhibiting intelligence, or producing competitive advantage. Learning processes sometimes result in confusion and mistakes.

Nevertheless, we are disposed to see elements of intelligence in organizational efforts to organize, construct, and interpret experience, so as to move toward a shared understanding of it. We think organizations learn from their histories in ways that are, at times, remarkably subtle adaptations to the inferential inadequacies of historical experience. We recognize some advantages in having stable, shared beliefs about experience even if misinterpretations are embedded in those beliefs. We see possibilities for expanding and enhancing unique ambiguous events so as to learn more richly and validly from them. We believe that usable knowledge can be extracted from fragments of history and that intensive examinations of individual cases can be used imaginatively to construct meaningful hypothetical histories.

Such beliefs depend ultimately on confidence in being able to resolve some fundamental issues in historical inference. These include four critical questions:

- What is the evidential standing of imagination? Organizations use near-histories and hypothetical histories to learn from samples of one or fewer. The procedures seem to have elements of intelligence in them, but they mix theoretical and empirical knowledge in ways that are not considered comprehensively in our theories of inference.
- What is a proper process for combining prior expectations and interrelated, cumulated aspects of a rich description into an interpretation of history? Organizations develop and modify stories about history on the basis of detailed examinations of individual cases. It is clear that radically different stories may be told about the same history. But it also seems clear that the evaluation of stories is not arbitrary, that there are criteria for differentiating between good and bad stories.
- What is the proper trade-off between reliability and validity in historical interpretation? As organizations develop theories of their experiences, they balance gains

and losses in validity against gains and losses in reliability. The metric and the procedures for the trade-off are ill defined, but there seems little doubt that an intelligent organization will sometimes sacrifice conventional notions of validity in order to achieve or sustain reliability in interpretation.

- What are the relative values of multiple observations of events and multiple interpretations of them? Improving precision in estimates involves pooling over observations and over observers. Theories of historical inference tend to emphasize pooling over observations. Pooling over observers appears to have advantages in some common situations, but in the absence of a clearer formulation of the gains and losses involved, it is hard to specify the precise conditions favoring one strategy or the other.

These questions invite heroic philosophical and methodological efforts to clarify and extend the uses of historical experience in the construction and sharing of meaning. The problems involved are not trivial. Nevertheless, we think modest progress can be made without waiting for a revolution in epistemology and within reasonably conventional modes of thinking about historical inference and learning from experience. The present paper is in that spirit. By examining the ways organizations actually seek to learn from small histories, and by trying to make sense of some of the things they do, we have tried to suggest some possible directions for understanding how meaning is extracted from sample sizes of one or fewer.

ACKNOWLEDGEMENTS

The research has been supported by the National Science Foundation, the Spencer Foundation, the Stanford Graduate School of Business, the System Development Foundation, and the Xerox Corporation. We are grateful for the comments of Michael Cohen, Robyn Dawes, Kristian Kreiner, Arie Lewin, Allyn Romanow, Sim Sitkin, and Suzanne Stout.

Authors' affiliations

J G March, Stanford University, Stanford, California 94305, USA
L S Sproull, Boston University, Boston, Massachusetts 02215, USA
M Tamuz, Rutgers University, New Brunswick, New Jersey 08903, USA

REFERENCES

- Allison GT. *Essence of decision*. Boston, MA: Little, Brown, 1971.
- Aronson E. Disconfirmed expectancies and bad decisions—discussion: expectancy vs. other motives. In: Abelson RP, Aronson E, McGuire W, *et al.* *Theories of cognitive consistency*. Chicago, IL: Rand McNally, 1968:491–3.
- Becker GS, Stiller GJ. De Gustibus non est Disputandum. *Am Econ Rev* 1977;67:76–90.
- Bell TE, Escu K. The fatal flaw in flight 51-L. *IEEE Spectr* 1987;24:36–51.
- Beyer JM. Ideologies, values and decision-making in organizations. In: Nystrom PC, Starbuck WH, eds. *Handbook of organizational design*, Volume 2. Oxford: Oxford University Press, 1981:166–202.
- Björkman I. *Foreign direct investments. an empirical analysis of decision making in seven Finnish firms*. Helsinki: Svenska Handelshögskolan, 1989.
- Boisjoly R. Ethical decisions—Morton Thiokol and the space shuttle Challenger disaster. *Am Soc Mech Eng J*, 87-WA/TS-4, 1987:1–13.
- Brehmer B. In one word: not from experience. *Acta Psychol* 1980;45:223–41.
- Brunsson N. *The irrational organization: irrationality as a basis for organizational action and change*. Chichester, England: Wiley, 1985.
- Brunsson N. *The organization of hypocrisy*. Chichester, England: Wiley, 1989.
- Campbell D. Degrees of freedom and the case study. In: Cook TD, Reichardt CS, eds. *Qualitative and quantitative methods in evaluations research*. Beverly Hills: Sage, 1979.
- Cangelosi V, Dill WR. Organizational learning: observations toward a theory. *Admin Sci Q* 1965;10:175–203.
- Clark BR. The organizational saga in higher education. *Admin Sci Q* 1972;17:178–84.
- Cohen CE, Ebbesen EB. Observational goals and schema activation: a theoretical framework for behavior perception. *J Exp Soc Psychol* 1979;15:305–29.
- Cohen MD, March JG, Olsen JP. A garbage can model of organizational choice. *Admin Sci Q* 1972;17:1–25.
- Cyert RM, Dill W, March JG. The role of expectations in business decision making. *Admin Sci Q* 1958;3:307–40.
- Cyert RM, March JG. *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice-Hall, 1963.

- Dawes RM, Faust D, Meehl PE. Clinical versus actuarial judgment. *SCI* 1989;**243**:1668–74.
- Dearborn DC, Simon HA. Selective perception: a note on the departmental identification of executives. *Sociom* 1958;**21**:140–4.
- Durkheim E. *On morality and society*, Translated by R N Bellah. Chicago, IL: University of Chicago Press, 1973.
- Einhorn H, Hogarth R. Confidence in judgment: persistence in the illusion of validity. *Psychol Rev* 1978;**85**:395–416.
- Feldman MS, March JG. Information as signal and symbol. *Admin Sci Q* 1981;**26**:171–86.
- Fischhoff B. Hindsight ≠ foresight: the effect of outcome knowledge on judgment under uncertainty. *J Exp Psych: Human Perception and Performance* 1975;**1**:288–99.
- Fischhoff B. For those condemned to study the past: reflections on historical judgment. In: Shweder RA, Fiske DW, eds. *New directions for methodology of behavioral science*. San Francisco: Jossey-Bass, 1980:79–93.
- Fischhoff B. Debiasing. In: Kahneman D, Slovic P, Tversky A, eds. *Judgment under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press, 1982:422–44.
- Fischhoff B, Beyth R. 'I knew it would happen'—remembered probabilities of once-future things. *Organ Behav Human Performance* 1975;**13**:1–16.
- George AL. *Presidential decision making in foreign policy: the effective use of information and advice*. Boulder, CO: Westview, 1980.
- George AL, Mckeown T. Case studies and theories of organizational decision making. In: Coulam RF, Smith RA, eds. *Advances in information processing in organizations*, Volume 2. Greenwich, CT: JAI Press, 1985:21–58.
- Harrison JR, March JG. Decision-making and postdecision surprises. *Admin Sci Q* 1984;**29**:26–42.
- Hax AC, Majluf NS. *Strategic management: an integrated perspective*. Englewood Cliffs, NJ: Prentice Hall, 1984.
- Herbst PG. *Behavioral worlds. The study of single cases*. London: Tavistock, 1970.
- Higgins ET, Bargh JA. Social cognition and social perception. *Ann Rev Psychol* 1987;**38**:369–425.
- Hogarth R. *Small probabilities: imagination as experience*, Working Paper, University of Chicago Center for Decision Research, 1983.
- Isen AM, Schalker TE, Clark M, et al. Affect, accessibility of material in memory, and behavior: a cognitive loop? *J Personality Soc Psychol* 1978;**36**:1–12.
- Janis IL, Mann L. *Decision making: a psychological analysis of conflict, choice and commitment*. New York: Free Press, 1977.
- Johnson-Laird PN. *Mental models: towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press, 1983.
- Kahneman D, Tversky A. The simulation heuristic. In: Kahneman D, Slovic P, Tversky A, eds. *Judgment under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press, 1982:201–8.
- Kahneman D, Tversky A. Intuitive prediction: biases and corrective procedures. In: Kahneman D, Slovic P, Tversky A, eds. *Judgment under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press, 1982:414–21.
- Kiesler S, Sproull LS. Managerial response to changing environments: perspectives on problem sensing from social cognition. *Admin Sci Q* 1982;**27**:548–70.
- Krieger S. Fiction and social science. In: Krieger S, eds. *The mirror dance: identity in a women's community*. Philadelphia, PA: Temple University Press, 1983:173–99.
- Lant TK, Montgomery DB. Learning from strategic success and failure. *J Bus Res* 1987;**15**:503–18.
- Lawrence P, Lorsch J. *Organization and environment: managing differentiation and integration*. Boston: Harvard Graduate School of Business Administration, 1967.
- Luvitt B, March JG. Organizational Learning. *Ann Rev Sociol* 1988;**14**:319–40.
- Loftus E. *Eyewitness testimony*. Cambridge, MA: Harvard University Press, 1979.
- Lord C, Lepper MR, Ross L. Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *J Personality Soc Psychol* 1979;**37**:2098–110.
- Lounamaa PH, March JG. Adaptive coordination of a learning team. *Manage Sci* 1987;**33**:107–23.
- Maier NRF. *Problem-solving discussions and conferences: leadership methods and skills*. New York: McGraw-Hill, 1963.
- March JG. Bounded rationality, ambiguity, and the engineering of choice. *Bell J Econ* 1978;**9**:587–608.
- March JG. Ambiguity and accounting: the elusive link between information and decision making. *AOS* 1987;**12**:153–68.
- March JG. Exploration and exploitation in organizational learning. *Organ Sci* 1990;**2**:71–87.
- March JG, Olsen JP. The uncertainty of the past: organizational learning under ambiguity. *Eur J Polit Res* 1975;**3**:147–71.
- March JG, Olsen JP. *Rediscovering institutions: the organizational basis of politics*. New York: Free Press, 1989.
- March JG, Sevón G. Gossip, information, and decision making. In: Sproull LS, Crecine JP, eds. *Advances in information processing in organizations*, Volume 1. Greenwich, CT: JAI Press, 1984:95–107.
- March JG, Shapira Z. Managerial perspectives on risk and risk taking. *Manage Sci* 1987;**33**:1404–18.
- Martin J. Stories and scripts in organizational settings. In: Hasdorff AH, Isen AM, eds. *Cognitive social psychology*. New York: Elsevier-North Holland, 1982:255–305.
- Martin J, Siehl C. Organizational culture and counter culture: an uneasy symbiosis. *Organ Dynam*, 1983 Autumn, 52–64.
- Merton R. *Social theory and social structure*. New York: The Free Press, 1968.
- Mohr LB. The reliability of the case study as a source of information. In: Coulam RF, Smith RA, eds. *Advances in information processing in organizations*, Volume 2. Greenwich, CT: JAI Press, 1985:65–94.
- National Research Council (Assembly of Engineering Committee on FAA Airworthiness Certification Procedures). *Improving aircraft safety: FAA certification of commercial passenger aircraft*. Washington, DC: National Academy of Sciences, 1980.
- Neustadt RE, May ER. *Thinking in time: the uses of history for decision-makers*. New York: Free Press, 1986.
- Oliver T. *The real Coke, the real story*. New York: Penguin Books, 1986.
- Pearson RW, Ross M, Dawes RM. Personal recall and the limits of retrospective questions in surveys. In: Tanur J, eds. *Questions about survey questions*. Beverly Hills, CA: Sage, 1990.
- Prange GW. *Miracle at Midway*. New York: McGraw-Hill, 1982.
- Raiffa H. *Decision analysis*. Reading, MA: Addison-Wesley, 1968.
- Rorty R. Solidarity and objectivity. In: Rajchman J, West C, eds. *Post-analytic philosophy*. New York: Columbia University Press, 1985:3–19.
- Salancik GR. Commitment and control of organizational behavior and belief. In: Staw BM, Salancik GR, eds. *New directions in organizational behavior*. Chicago, IL: St. Clair, 1977:1–54.
- Shedler J, Manis M. Can the availability heuristic explain vividness effects? *J Personality Soc Psychol* 1986;**51**:26–36.
- Sproull LS, Hofmeister KR. Thinking about implementation. *J Manage* 1986;**12**:43–60.
- Sproull LS, Weiner S, Wolf D. *Organizing an anarchy: belief, bureaucracy, and politics in the National Institute of Education*. Chicago, IL: University of Chicago Press, 1978.
- Starbuck WH. Organizations as action generators. *Am Social Rev* 1983;**48**:91–102.
- Starbuck WH, Milliken FJ. Challenger: fine-tuning the odds until something breaks. *J Manag Stud* 1988;**125**:319–40.
- Stinchcombe A. *Theoretical methods in social history*. New York: Academic Press, 1978.
- Tamuz M. The impact of computer surveillance on air safety reporting. *Columbia J World Bus* 1987;**22**:69–77.
- Tamuz M. Monitoring dangers in the air: studies in ambiguity and information. PhD thesis, Stanford University, 1988.
- Tolstoy LN. *War and peace*. Translated by R Edmonds. Harmondsworth, UK: Penguin, 1869.
- Weick K. *The social psychology of organizing*, 2nd ed. Reading, MA: Addison-Wesley, 1979.
- Zald MN, Denton P. From evangelism to general service: the transformation of the YMCA. *Admin Sci Q* 1963;**8**:214–34.

COMMENTARY

STILL LEARNING HOW TO LEARN

The emergence of safety in health care as a legitimate systemic public issue has turned on its head a good deal of our traditional thinking. One area where this has been particularly acute is what we could call “medical epistemology”—how we know we know something in health care. When health professionals first encounter the study of safety in health systems, they frequently feel a tension between the familiar world of epidemiology and bioscience and this strange new world, with its roots in psychology and engineering and its methods seemingly subjective and anecdotal. Cook¹ pictures them caught between these two worlds, trying to learn new ways of learning about safety but holding on to the security blanket of more familiar evidentiary methods. Their problem goes beyond simple intellectual assent; to some extent there seems to be an emotional and aesthetic—one might even say visceral—discomfort with these new methods, a fear that letting go of the evidence-based life ring will inevitably lead to superstition, myth, and chaos. This tension has led to debates about proper methods,² with both sides largely preaching to the already converted.³

The classic paper by March *et al*⁴ republished here should help redirect this conversation into more productive areas. By describing in detail ways that organizations learn, or attempt to learn, from history when history has not been “generous with experience”, they remove some of the mystery that has been associated with learning from what have often been disparaged as mere anecdotes.

Most of the events of interest in patient safety are relatively rare—in fact, we wish they were even rarer. The traditional biomedical approach to the problem of scant data is pooling experience across multiple events, but for safety this is often not practical and, even more often, not desirable for two reasons: (1) the events of interest are so highly contingent on a specific context that pooling loses rather than adds information, and (2) some failures such as death from mistaking potassium chloride for furosemide are so devastating that we cannot afford to wait for more events to draw the appropriate safety lessons. March *et al* point out that organizations in this situation can improve their learning in three ways. Firstly, they can experience events more richly by attending to more aspects of them and by allowing for more interpretations, viewpoints, and preferences on them. Secondly, they can enrich their experience by attending to near-events—failures narrowly averted or successes improbably achieved—and hypothetical events, such as by simulation or predicting possible modes of failure. Finally, they can develop better understandings of events by allowing them to emerge from the details of a variety of aspects of the experience rather than by formal analysis of unambiguous objectives in a priori models.

This latter process contrasts sharply with familiar inferential methods. Although it is clearly not as well specified or understood as traditional verification and validation research, it is well established in the process of building scientific knowledge,³ even though it is not commonly admitted in health care. (It is, however, familiar, because it is how we teach clinical medicine.) This is not to say that the sort of inductive theory building described by March *et al* is by any means perfect; it faces challenges to reliability and validity and would appear to be more difficult to reduce to a “method” or to teach to others. However, by describing some of the details of the process and posing some critical questions about it, March *et al* have made this sort of learning more accessible to researchers and managers, and raised hopes that progress can be made in finding the proper relationship between the two research worlds.

While it seems clear that organisations sometimes do learn, what is even more striking in health care is how often they do not learn at all. Tucker and Edmondson⁶ have pointed out how failures of organizational learning in hospitals result from the confluence of organizational and psychological factors. Their analysis reinforces the points made by March *et al*. They found that healthcare organizations have difficulty learning from experience because, in a sense, they have no experience. Problems and failures encountered by front line practitioners are commonly resolved on the front lines by a process of empirical patching, so their very existence is never known to much of the organization. What knowledge is gained is shared primarily among similar social groups—for example, nurses speaking primarily only to other nurses—so the number of viewpoints, interpretations, and values

available is limited. Because of production pressure, front line workers have no time or energy to invest in examining near-events or to engage in “what if” projections of possible failures. Finally, when data do reach managers or investigators situated to respond less reactively, their training has not provided them with the analytical skills suitable for abstraction and inductive empirical generalization from rare events and, in fact, may have biased them against it.

If health care is to become safer, we will need to find ways to enhance our learning from small samples. This may not be as difficult as it seems at first, as it may involve no more than bringing into the open processes that we have traditionally suppressed but which were operating nonetheless. For example, a recent review of the management of shock syndromes in children in an attempt to develop practice guidelines deplored the paucity of high quality level I evidence while simultaneously noting that mortality had fallen by a factor of 10.⁷ Clearly, some learning had taken place despite the lack of “evidence”! It is interesting to speculate that much of the success of “scientific” medicine might depend more on the subjective decisions of thoughtful clinician researchers about what would be worth trying than on the rigour of its inferential methods.

Admitting new sorts of evidence and new ways of thinking is always risky. We could clearly learn the wrong lessons some of the time and we need to learn a great deal more about what distinguishes good from poor quality inferences drawn from samples of one or fewer. But the alternative—to discount this method of learning altogether—seems much less palatable. The paper by March *et al* should move us to learn more about how to learn about safety.

R L Wears

Department of Emergency Medicine and Center for Safety in Emergency Care, University of Florida Health Science Center, 655 West Eighth Street, Jacksonville, FL 32209, USA; wears@ufl.edu

REFERENCES

- 1 Cook RI. Two years before the mast: learning how to learn about patient safety. In: Scheffler AL, Zipperer LA, eds. *Proceedings of the Second Annenberg Conference on Enhancing Patient Safety and Reducing Errors in Healthcare*. Rancho Mirage, CA: National Patient Safety Foundation, 1998:61–4.
- 2 Shojania KG, Duncan BW, McDonald KM, *et al*. Safe but sound: patient safety meets evidence-based medicine. *JAMA* 2002;**288**:508–13.
- 3 Leape LL, Berwick DM, Bates DW. What practices will most improve safety? Evidence-based medicine meets patient safety. *JAMA* 2002;**288**:501–7.
- 4 March JG, Sproull LS, Tamuz M. Learning from samples of one or fewer. *Organ Sci* 1991;**2**:1–13.
- 5 Handfield RB, Melnyk SA. The scientific theory-building process: a primer using the case of TQM. *J Operations Manage* 1998;**16**:321–39.
- 6 Tucker AL, Edmondson AC. Why hospitals don't learn from failures: organizational and psychological dynamics that inhibit system change. *Calif Manage Rev* 2003;**45**(2).
- 7 Carcillo JA, Fields AI. Clinical practice parameters for hemodynamic support of pediatric and neonatal patients in septic shock. *Crit Care Med* 2002;**30**:1365–78.