Journal of Applied Psychology
2005, Vol. 90, No. 5, 857–871

0021-9010/05/$12.00 DOI: 10.1037/0021-9010.90.5.857

After-Event Reviews: Drawing Lessons From Successful and Failed Experience

Shmuel Ellis and Inbar Davidi
Tel Aviv University

The claim that appropriate after-event review might decrease the relative advantage of drawing lessons from failures over drawing lessons from successes was examined in a quasi-field experiment. The results show that performance of soldiers doing successive navigation exercises improved significantly when they were debriefed on their failures and successes after each training day, compared with others who reviewed their failed events only. The findings also show that, before the manipulation, in both groups, learners' mental models of failed events were richer in constructs and links than were their mental models of successful events. This gap closed gradually in subsequent measurements.

Keywords: learning from experience, mental models, cause maps, need for closure, after-event reviews

Failures have traditionally been considered as better motivators than successes for drawing lessons from experience (Sitkin, 1992; Weiner, 1985). Schank (1986), in his theory of failure-driven learning, argued that failed expectations trigger tweaking, a process that adapts an explanation pattern to an unexpected situation. To quote Weiner (2000),

> Search is not undertaken following all events, and it is particularly likely when an outcome is negative, unexpected, and/or important. Thus, if one expects to succeed and does, "*why questions*" are not likely to follow. But unexpected failure at an important exam surely will evoke attributional processes. (p. 2)

These effects have been demonstrated repeatedly by Lau and Russell (1980), Wong and Weiner (1981), Hastie (1984), and Maheswaran and Chaiken (1991). In the organizational context, Zakay, Ellis, and Shevalski (2004) asked managers to evaluate the need to initiate a learning process after a particular organizational experience. It was found that the more negative were the results of the particular event, the stronger was the managers' inclination to recommend a more intensive process of learning from experience. Likewise, a need to create control and follow-up procedures was reported only after negative outcomes.

Not surprisingly, the tendency to draw lessons from failures contributed to the development of a body of psychological and organizational literature about learning from failures or errors (Heimbeck, Frese, Sonnentag, & Keith, 2003; Seifert & Hutchins, 1992) while overlooking the potential value of learning from successes. Learning from successes is needed under two conditions: (a) when learners are not sure that the successful performance is a result of their ability or effort but might be attributed to luck (Weiner, 1985). In such cases, individuals and organizations will want to understand why events happened as they did, as a basis for planning future actions; and (b) when the cost of errors is extremely high. In organizations such as nuclear plants, pharmacological industries, aircraft companies, or particular military units, extensive efforts and resources are invested in developing efficient and stable activity patterns, that is, highly standardized routines, to prevent errors (Hannan & Freeman, 1984). But at the same time, such organizations are exposed to errors when they do the same things after circumstances have changed, and the changes go undetected because people are bored, rushed, distracted, careless, or simply ignorant (Weick, Sutcliffe, & Obstfeld, 1999). To detect potential failures and to adapt to events that require revision of old and successful routines or plans, problem solvers must understand when and why these routines still work and under what conditions they do not. They need to be able to revise their understanding of the situation, their evidence collection and evaluation tactics, or their response strategy when new events are detected and evaluated (Weick et al., 1999).

The goal of the present study was to highlight the value of systematic reviewing of successes, or in other words, to encourage asking "why" questions not only after failed events but also after successful ones. We focused on a single organizational learning mechanism, after-event review (AER), and argue that although failures have an initial motivational advantage over successes for learning, by conducting appropriate after-action reviews, organizations can benefit from helping individuals to draw important lessons from successful experiences as well.

Drawing Lessons From Experience: The Role of AERs

In the present article, we define *AER* as an organizational learning procedure that gives learners an opportunity to systematically analyze their behavior and to be able to evaluate the contribution of its various components to performance outcomes. In a broader sense, according to Busby (1999), the role of AERs is to intensify cognitive elaboration of experiential data, under the assumption that this process will ultimately promote the necessary

Shmuel Ellis and Inbar Davidi, The Leon Recanati Graduate School of Business Administration, Tel Aviv University, Tel Aviv, Israel.

Correspondence concerning this article should be addressed to Shmuel Ellis, The Leon Recanati Graduate School of Business Administration, Tel Aviv University, Ramat Aviv, Tel Aviv, 69978, Israel. E-mail: sellis@post.tau.ac.il

behavioral changes. Furthermore, one may argue that AERs help learners to switch their mode of information processing from automatic to conscious. In the automatic mode, people run well learned scripts and respond to external cues only in terms of their well established, existing mental models. In the conscious mode, by contrast, their cognitive activity is characterized by more awareness, attention, information gathering, and reflection (Chanowitz & Langer, 1980; Louis & Sutton, 1991).

AERs have three functions in learning from experience: self-explanation, data verification, and feedback. These functions are described in detail in the following sections.

Self-Explanation

During AER, individuals are asked to analyze their behavior and to suggest explanations for their successful and/or failed actions. Chi, Bassok, Lewis, Reinman, and Glaser (1989); Chi, de Leeuw, Chiu, and Lavancher (1994), and Pirolli and Recker (1994) succeeded in promoting the skill acquisition of individuals by asking them to generate explanations for their actions. Others, such as Ferguson-Hessler and de Jong (1990) and Nathan, Mertz, and Ryan (1994), found the number of self-explanations to be correlated with problem-solving successes in various content domains. As an *active process* of gathering, analyzing, and integrating data, self-explanation has been found not only to direct learners to reflect on their past behavior but also to facilitate the construction of "if–then" rules that help to improve subsequent performance and encourage the integration of newly learned materials with existing knowledge (Chi et al., 1994).

Data Verification

Insofar as learning from the past is open to errors and biases, the lessons drawn may also turn out to be irrelevant, invalid, or even misleading (Ariely & Zakay, 2001; March, Sproull, & Tamuz, 1991). Several cognitive biases have been found to be potential sources of danger to optimal learning. One such bias is the confirmation bias (e.g., Brehmer, 1980; Feldman, 1989), that is, the tendency of individuals to overlook information that is not compatible with their a-priori hypotheses.

The hindsight bias (Fischhoff, 1982), implying that knowledge of outcomes strongly affects how people view their past experience, poses another threat to the validity of lessons drawn from the past. Zakay (1984) found that when managers evaluate their decisions, they are strongly influenced by the value of known outcomes. Positive outcomes, in contrast to negative ones, increase the probability that past decisions will be evaluated as good ones. Moreover, managers in general see outcome value as the most important criterion of decision quality.

The AER is the place where learners confront different perceptions of the same data. Accepting other people's perceptions depends, of course, on the relative expertise and authority of the learners, the event reviewer, and the other participants in the AER session (Ellis & Kruglanski, 1992). However, although there is no guarantee that learners will yield to other people's views, it can be argued that they will, at least, have to relate to them and even to reelaborate this information before they reject it. In sum, one may say that AERs enable learners to cross-validate the information they hold before changing or correcting their mental models.

Feedback

Finally, feedback is an important by-product of the debriefing process. Feedback is defined as information with which a learner can confirm, add to, overwrite, tune, or restructure information in memory, whether that information is domain knowledge, metacognitive knowledge, beliefs about self and task, or cognitive tactics and strategies (Alexander, Schallert, & Hare, 1991). The main advantage of feedback that is received in AERs is that it concentrates not only on performance outcomes but also, and especially, on the process of task performance, that is, not only on the global outcome (overall success or failure) but also on local outcomes (particular aspects or parts of task performance). Various types of cognitive feedback focusing on different aspects of task performance have been suggested as mechanisms of learning improvements. For example, Balzer, Doherty, and O'Connor (1989) suggested the concepts of task validity, cognitive validity, and functional validity feedback. Thus, task validity feedback, for instance, might bring the learner's attention to the relationship between the presence and use of a compass and the probability of successful navigating performance. Cognitive validity feedback might suggest to the learner the use of the compass to plan his or her navigation route. Functional validity feedback might help learners to understand the gap between their estimates of their achievements and their actual performance. It should be noted, however, that in contrast to feedback, AER is an organizational procedure aimed at helping individuals and groups to gather and analyze data that will ultimately improve their performance. One may say that AER is a kind of guided self-explanation. Feedback, in contrast, is generally provided by an external authority and conveys already elaborated information to the learner. In a review of 19 published studies, Webb (1989) showed that giving elaborate explanations yielded more and stronger correlations with individual achievements than did receiving elaborate explanations.

Mental Models and Learning From Experience

The role of information processing in learning from experience has been emphasized by many researchers. Huber (1991), for example, defined learning as information processing that leads to behavioral change. Others, like Friedlander (1983), even considered information processing as the dependent variable of the learning process. Friedlander argued that learning results in new and significant insights and awareness or a change in one's cognitive maps. In the present article, we use behavioral changes (performance improvement) as indicators of learning, but we analyze learners' mental models to get a better understanding of the cognitive determinants of these behavioral changes.

The concept of mental models has received increasing attention in the organizational literature over the last decade (Klimoski & Mohammed, 1994; Laukkanen, 1994; Walsh, Henderson, & Delighton, 1988). It covers a wide range of similar concepts representing different types of knowledge structures, for example, cognitive maps (Axelrod, 1976; Neisser, 1976; Weick & Bougon, 1986), belief structures (Fiske & Taylor, 1991), production rules (Anderson, 1983; Ohlsson, 1996), mental models (Senge & Sterman, 1991), theories in use (Argyris & Schon, 1978), schemas (Neisser, 1976; Rumelhart, 1984), categories (Higgins & Lurie, 1983), and scripts (Abelson, 1976).

The creation of any kind of mental model starts with a hypothesis one generates and proceeds with testing of the hypothesis against available evidence to determine the confidence one has in it (Kruglanski, 1989, 1991). Individuals tend to generate new hypotheses after facing unexpected problems, when acts are frustrated, when there is an unexpected failure or a disruption, or when there is a significant difference between expectations and reality (Pyszczynski & Greenberg, 1987). However, people may also test hypotheses suggested by others or spontaneously generate hypotheses to learn about their social (or organizational) environment (Trope & Liberman, 1996).

People use the various types of mental models to filter out inconsistent or incongruent information to attend to particular aspects of their experience (Kiesler & Sproull, 1982; Weick & Bougon, 1986). Mental models not only give meaning to the social environment but also function as a frame of reference for action and interpretation of the world (Gioia & Manz, 1985; Gioia & Poole, 1984). Research studies in organizations have demonstrated the value of eliciting mental models to understanding managerial behavior (Huff, 1990; Senge & Sterman, 1991; Starbuck & Milliken, 1988), to responding to organizational disasters (Kiesler & Sproull, 1982), and to diagnosing organizations (Bougon, Weick, & Binkhorst, 1977).

Rabkin (1995) argued that in almost any domain of business where people learn to build new products or use new process technology, they first of all try to notice the new, relevant variables and then integrate them into a model that can help guide behavior. Furthermore, as noted above, the richness of individuals' mental models reflects their mode of information processing (Chanowitz & Langer, 1980; Louis & Sutton, 1991). Whereas noticing new variables during the automatic processing mode is very unlikely, in the conscious mode, people invest time and effort to deepen their understanding of the problems they are facing. Bartunek, Gordon, and Weathersby (1983) described cognitive complexity according to the number of constructs embedded in the knowledge structures and according to the way in which they are connected. The complexity or extensiveness of a mental model reflects the depth and breadth of knowledge or expertise that it represents (Evans, 1988). Previous research has suggested that experts hold more detailed mental models (Murphy & Wright, 1984; Tanaka & Taylor, 1991). Furthermore, Carley (1997) found that members of successful groups tend to use more concepts and have larger cognitive maps (more constructs) than do members of nonsuccessful groups. In addition, Neisser (1976) showed how having complex knowledge structures to identify subtle differences among various environmental stimuli helps navigators to perform better.

If mental models guide people's behavior or responses to various social stimuli, learning may be defined as the process of formulating and updating mental models. That is, learning is the process of noticing new variables that are relevant to explaining and predicting various social phenomena or, in other words, the process of hypothesis generation and validation.

Drawing Lessons From Failed Versus Successful Events

Failed and successful events constitute two databases from which individuals can make inferences to advance the generation of new propositions or draw evidence to confirm/refute old or new ones. Naturally, the two databases differ according to the relative number of erroneous, as opposed to correct, actions within each event. The failed events include more errors (and even more serious errors) than do the successful events.

Errors are usually detected by comparing actual with expected outcomes (Allwood, 1984; Miller, Galanter, & Pribram, 1960). In other words, errors are conceived as deviations from the right solution of a problem, requiring a change in one's mental model that is supposed to lead to a solution. Explicit feedback is a major tool that individuals use for interpreting their performance as right or wrong. When explicit feedback is not available, performers can detect errors by comparing their actual performance against an internal representation of the intended action, that is, their prior plans or mental models (Adams, 1971; Ohlsson, 1996; Schmidt, 1975). It was recently found that using errors in training helps to develop good operative mental models as long as the errors do not create negative emotional effects (Heimbeck et al., 2003). Errors enlighten areas of misunderstandings and thus trigger learners to develop thoughtful strategies to learn and enhance mental models (Lord & Levy, 1994).

In contrast to failures, drawing lessons from successful actions is much more difficult for two reasons. First, whereas propositions (more specifically, mental models or behavioral rules) can be refuted by errors (Popper, as cited in Berkson & Wettersten, 1984; Kruglanski, 1980, 1989; Ohlsson, 1996), there is as yet no accepted analysis of how successful actions can contribute to refreshing or enriching existing mental models. In other words, whereas deciphering failures may ultimately change or improve mental models, successes can just increase one's confidence in the old ones (Sitkin, 1992). Second, the gaps between the original plan (mental model, rule) and performance, which are the basis of learning from failure, do not exist in successes. This is the basic reason it is so difficult to analyze successful events. Therefore, if learners still want to improve their performance, they must focus on the internal logic of their plans (prior mental models) and on the potential misfits between the existing mental model and the conditions under which performance was executed. Thus, by concentrating on the central decision-making points of the relevant performance and by systematic generation of if–then optional propositions, individuals can improve their mental models and future performance. While reviewing their successes, individuals may also be surprised to find that their outcomes have been accomplished accidentally and as a result will revise their mental model and change their behavior. Furthermore, successful events are not always errorless. Careful review of successful events may reveal information that is high in value.

The Motivational Dimensions of Learning From Experience

The process of learning from experience is not independent of motivational influences. Whereas cognitive processes have to do with how individuals formulate hypotheses regarding their social environment and how they determine their confidence in the hypotheses generated, motivation to acquire knowledge has to do with the intensity and duration of the knowledge acquisition processes, or in other words, with the braking or starting mechanisms of the knowledge acquisition sequence.

Kruglanski (1989, 1991) argued that individuals activate the cognitive process of knowledge acquisition according to their level of need for closure, that is, their desire for a clear, firm, or unambiguous answer to a question. The higher the perceived benefits of closure (e.g., meeting an important deadline, removing the necessity for further information processing), the more intensive will be the acts taken to attain closure. Conversely, a desire for judgmental noncommitment elicits avoidance of closure. Individuals who doubt the validity of their own knowledge because of the prospective high cost of a mistaken judgment will prefer to refrain from making high-commitment decisions.

Failures might evoke high fear of invalidity in individuals with experience of previous closure that proved to be inadequate (Freund, Kruglanski, & Schpitzajzen, 1985). Individuals who feel that the knowledge structure on which they based a failing performance was flawed might want to improve their knowledge so as to improve their performance (Kruglanski & Klar, 1987; Weiner, 1985). More specifically, if learners genuinely wish to improve their performance, they will increase their epistemic activity so as to revise their mental models (Hastie, 1984; Lau & Russell, 1980; Wong & Weiner, 1981), halting it as soon as they have to utilize the revised knowledge. In other words, the sooner the next experience, the higher the need to attain cognitive closure.

In contrast to failures, successes do not create an urgent need for information gathering or hypothesis generation. Because successes generally confirm prior expectancies (Weiner, 1985, 2000) and increase confidence in old routines (Weick, 1984), they generally reduce the cost of remaining in closure and tend to discourage any cognitive activity designed to invoke immediate change. According to Sitkin (1992), successful outcomes tend to restrict search, reduce attention, and increase complacency and risk aversion. Successful events provide a stable basis for future activity (Weick, 1984) and encourage people to develop expectations that these outcomes will be repeated (Herriot, Chalmers, & Wingrove, 1985; Sitkin, 1992; Weick, Sutcliffe, & Obstfeld, 1999).

Thus, if we want successful learners to draw lessons from their successes, we must induce them to do so. Because events are comprised of many acts, and as noted above, sometimes there is a lot to improve even in events that have been defined as successful, we may take measures (asking questions) to intensify attributional activity that would not have otherwise taken place (Enzle & Schopflocher, 1978). By attracting learners' attention not only to obvious failed aspects of their activity but also to successful activities or decisions, it is possible to intensify their epistemic activity, trigger their motivation to revise their knowledge structures, and ultimately improve their performance. In the organizational context, a formal AER may constitute an effective tool to help relatively successful learners to use valuable information that otherwise would be ignored, to enrich their mental models, and improve their performance. More specifically, though mental models of failed events are basically richer than those of successful events, it is expected that, under reviews of successful experience, this advantage will diminish. It is also expected that eliciting learners' mental models might give some indication of the different ways learners analyze failed as opposed to successful events.

The Study

Assuming different motivations to rethink failures and successes, we also created an experimental opportunity for learners to contemplate success, in an effort to examine the relative contribution of successes and failures to learning from experience. Soldiers from two companies of the Israel Defense Forces (IDF) taking a ground navigation course participated in the study. Because the army rejected the possibility of having experimental conditions of "AER focused only on success" or "no AER," one group learned by drawing lessons from failures, whereas the second group learned from failures as well as from successes. Thus, soldiers in one company were not encouraged to analyze successful experiences during the event reviews conducted after each navigation exercise, whereas soldiers of the second company were debriefed systematically on successful experiences as well as on failed experiences.

We expected that deliberate and systematic review of successful experience might be an effective substitute for the initial motivation to draw lessons from failed experience only and would also channel learners' epistemic activity toward the search for information that is not usually detected by reviewing failed events. Therefore, even though learners are motivated to ask more "why" questions and to search for more information after failed events (Hastie, 1984; Wong & Weiner, 1981), forcing them to also contemplate their successes might make their mental models even richer than those of learners who try to learn from their failures only. The failure-focused review might equalize the richness of the maps of the two groups on the failed aspects of their navigation assignments, whereas the debriefing of the successful aspects of the navigation assignments will add more constructs that cannot be generated by failure-focused AERs. Finally, under the assumption that richness of mental models is positively correlated with performance (Carley, 1997; Evans, 1988; Neisser, 1976), we expected more performance improvement after AERs of successful and failed events than after AERs of failed events only.

It should be emphasized that the participants were soldiers in an elite military unit who were aware of the high importance of navigation training and knew that failure in the final navigation exam might lead to their dismissal from the unit. In other words, they were genuinely motivated to learn. Because such individuals are less inclined to accountability biases and do not tend to attribute their failures to situational or ad hoc causes (weather or ground conditions as opposed to knowledge or planning), we expected that they would try to acquire information that could make a valuable and complementary contribution to their current knowledge. We made the following hypotheses:

> *Hypothesis 1:* Mental models of failed events are richer than those of successful events.

> *Hypothesis 2:* Mental models of learners after reviews of failed and successful events are richer than mental models of learners after reviews of failed events only.

> *Hypothesis 3:* Performance improvement is greater when reviews are concentrated on both successful and failed experience than when they are concentrated on failed experience only.

Method

Overview

The study was conducted at the training base of an elite unit of the IDF, the spectacular achievements of which have traditionally been attributed to, among other things, the hallowed institution of intense and rigorous AERs (Marinko, 1991; see also Lipshitz & Popper, 2000).

From direct observations and many informal talks with commanders and training personnel, we realized that the AERs conducted during the basic training courses of this unit (where soldiers are supposed to adopt a stringent event-reviews culture) were usually unbalanced; soldiers were practically encouraged to analyze failed events but not successful ones. Thus, we decided to conduct a field study comparing the traditional method of failure-focused after-event reviews (FAERs) to a new method that strikes a more even balance between reviews of failed and those of successful events, that is, failure- and success-focused after-event reviews (FSAERs). As previously noted, the army rejected the possibility of having an experimental condition of success-focused after-action reviews.

The participants in the study were 98 male soldiers, aged 18–20 years. The soldiers were in the middle of their basic field training and, more importantly, were about to finish the 1st week of their navigation course. The navigation course was selected for the present study for several reasons: (a) Navigation is a repetitive task (across different sites and times); (b) each navigation exercise is routinely reviewed by soldiers and their supervising commanding officers; and (c) AERs, as previously noted, usually focus on questions relating to failures (e.g., "What went wrong?" or "Why did not you succeed in reaching the specified coordinates?").

All soldiers in the two companies of this elite unit participated in the study. We were not allowed to assign soldiers individually to the experimental conditions, so we assigned one entire company to each and attempted to ensure equalization of the two conditions except for the experimental manipulation. In such quasi-experimental designs, which are very common in educational and social research (Campbell & Stanley, 1966), the groups are naturally assembled collectives such as classrooms, companies, or teams, as similar as availability permits but yet not so similar that one can dispense with a pretest. The two companies that participated in the present study were similar in three respects: (a) All soldiers were of high quality and were randomly assigned to the two companies at the beginning of their military service; (b) they went through precisely the same training schedule; and (c) they were exposed to the same event-reviews culture and military education.

The study was carried out in three stages, corresponding to the three stages of the training program schedule. In the first stage (the 1st week of the navigation training camp), the two companies took classes in various topics of ground navigation and had three preliminary navigation exercises. At the end of the first stage, one company was assigned to the traditional FAER experimental condition and the second to the FSAER condition. In the second stage (2nd training week), both companies had four navigation exercises. In the third stage (the 3rd week), both groups were assigned to the FSAER experimental condition. This allowed us to test the effect of the new review method twice and to cross-validate its effect. Two navigation exercises were included in the training program of the third stage. The level of difficulty of the exercises increased gradually throughout the navigation training course.

Procedure

In the first and second stages of the experiment, the two companies had the same training program, stayed in the same training camp, used the same training facilities, and performed similar assignments in the same navigation area—but 1 week apart. In To prevent spillover of information from one company to the other, Company B (the FSAER experimental group) arrived at the navigation training camp after Company A (the FAER experimental group) had already left. The two groups took the third stage of the navigation training at the same time, 2 months later. In the third stage, they were located in separate training areas with differing topographical characteristics and were given completely different assignments. To prevent exposure of the real objective of the study, the soldiers and their direct commanders were told that they were going to participate in a study on soldiers' learning skills.

Navigation exercises. The navigation exercise consisted of two parts. First, each soldier had to study the route and learn it by heart, so as to be able to arrive at five or so points along the way. In the second part, which took several hours, each soldier had to individually traverse the route, trying to reach each point as quickly as possible without using a map. As noted, in the first stage of the experiment, soldiers had three preliminary navigation exercises (three navigation days); in the second stage, they had four navigation exercises (four navigation days), and in the third stage, they had two navigation exercises (two navigation days). It should be noted that on each navigation day it was possible for soldiers to experience several events that could be failed, successful, or both.

Manipulating the independent variable. At the beginning of the 2nd training week, the supervising commanders of Company B started to debrief the soldiers about their failed and successful experiences. The soldiers of Company A, in contrast, continued to be debriefed according to the traditional after-event method—the FAER. It should be noted that, in each group, three supervising commanders conducted the AERs. These commanders had at least 8 years of experience with AERs (5 years as soldiers and at least 3 years as commanders). All soldiers completed a personal (one-on-one) AER with one of the commanders after each navigation exercise. About 15 min were allocated to debriefing each soldier at the end of each navigation exercise. To ensure that the AERs were performed according to the assigned method and to prevent experimental bias, Inbar Davidi observed all reviews of both experimental groups and measured the time devoted for each.

In the FSAER group, the soldiers were first asked to identify the positive and negative aspects of the navigation exercise. Then, while they were being debriefed about each part of the route, they were asked about the decisions that led to successful performance as well as those that led to failure. Finally, at the end of the event review, the soldiers were asked to note three aspects of the navigation exercise that needed improvement or that should be retained unaltered. In the FAER group, the commanders focused, as usual, on the failures. At the beginning of the debriefing session, the soldiers were asked about the problems they encountered during the navigation exercise. Then, during the detailed debriefing, they were asked about failures or problems that emerged along the route. There was no debriefing on the successful parts of the exercise. At the end of the event review, the soldiers were asked to note three points that needed improvement. It should be noted that it was not possible to fully structure the AERs. For example, a successful assignment for one soldier could be a failed one for another and vice versa. All the supervisors were aware of the major decision-making points along each of the navigation routes, so they were able to focus the debriefing on them if necessary.

Senior command of the military unit ascribed great importance to the implementation of AERs as an organizational routine, and commanding officers were expected to devote time and effort to them as an educational tool. Supervisors of the FSAER group had no greater motivation to work harder than did supervisors of the FAER group, and as was found later, there was (a) no significant difference, across navigation exercises, between the two companies in the mean time spent on event reviews, $M_A = 13.91$, $M_B = 14.62$, $F(1, 68) = 1.25$, $p < .26$, $\eta^2 = .018$; (b) no significant difference between the three AERs, $M_1 = 14.45$, $M_2 = 14.00$, $M_3 = 14.27$, $F(2, 136) = 0.41$, $p < .66$, $\eta^2 = .006$; and (c) no significant interaction of type of AER by time of AER, $F(2, 136) = 0.66$, $p < .51$, $\eta^2 = .001$.

Preparation workshop for supervisors. Four days before the beginning of the second stage of the experiment (the soldiers' 2nd week of navigation training), the commanding officers of both companies, all of whom were well versed in the army's routine procedures for debriefing, participated in one of two training workshops set up as part of the experimental design. One workshop stressed the importance of drawing lessons from successes as well as from failures (the innovative approach), and the other repeated the standard approach of learning from failures only. Before the beginning of the third stage of the experiment (the soldiers' 3rd week of training), the group of commanding officers who had participated in the workshop that had repeated the standard approach (learning from failed events only) underwent the innovative workshop, which emphasized the importance of learning lessons from successes as well. Commanding officers who had already done the innovative workshop did not do it again; they performed regular duties at this time.

The two workshops were presented as purporting to improve commanding officers' debriefing capabilities, but they differed in content. The training sessions of the innovative workshop started with a simulation game that exemplified the need to reconstruct the successful event so as to be able to replicate the successful results. Participants discussed the importance of reviewing successes and learning from them. They generated ideas on how to debrief after successful events and what questions to ask. The workshop ended with precise instructions on how to debrief soldiers during the next navigation-training week. The workshop repeating the standard approach was conducted simply to prevent experimental bias due to mutual suspicion of the two groups. The leader of this workshop discussed with participants the importance of learning from experience and how to cover the problematic aspects of a failed event and refreshed their memories on the standard procedures for debriefing after a navigation exercise. It should be noted that none of the commanding officers who participated in the FSAER workshop expressed any objection to the procedural change of the AER. On the contrary, after they were instructed on how to implement it, they welcomed the change.

Interviews for drawing cognitive cause maps. To minimize experimental biases, the data needed for the preparation of the cognitive cause maps of successes and failures were gathered in structured interviews. Ten soldiers randomly selected from each group were invited for personal interviews three times during the navigation training: before the manipulation (about 2 hr after the AER at the end of last navigation exercise of the last training session of the 1st training week) and 2 hr after the AERs of the last training session of the 2nd and 3rd training weeks. In these interviews, the soldiers were asked to select two navigation assignments of the training week they had just completed, one considered a success and the other a failure. The interviewer then asked them to describe each of the events in as much detail as possible and to list all the reasons for their success or failure. Each interview lasted about 90 min. To reduce random error and to avoid various biases, probe questions were prepared in advance, and identical procedures were adopted for the order of the questions and for wrapping up the interviews (for all interviewees).

Finally, the two companies were thoroughly debriefed at the end of the second stage of their navigation training. At this time, the true aims of the study were revealed, and potential applications were presented.

Measures

Performance. To evaluate soldiers' performance in the navigation training, the present study used the three measures that were routinely used by the unit's commanding officers: (a) number of specified points along the route that were accurately reached, (b) time to complete the task (navigation pace), and (c) number of times trainees needed to consult the map. Each navigation exercise during the two stages of training was evaluated according to these three dimensions. A weighted average of the three measures (45%, 45%, and 10%, respectively) comprised the final score for each soldier for each navigation exercise. It should be noted that soldiers very rarely open maps during navigation. In fact, almost all the soldiers obtained the maximum score for the third criterion.

Cognitive cause maps. To reflect learners' mental models of their successes or failures, we used cognitive cause maps, a pictorial device, consisting of nodes connected by arrows (Weick & Bougon, 1986). The nodes represent constructs (either causes or outcomes) that are enacted or defined by individuals with respect to a particular event, phenomenon, or domain of interest. The arrows represent the subjects' beliefs about the causal relationships among a particular group of constructs. A construct may have a direct or indirect causal link with other constructs. Thus, the cognitive cause map may show not only direct causal relations between constructs but also mediation effects of some constructs on others.

The interviewees' cognitive cause maps yielded two quantitative measures of mental model richness: number of constructs (causes) within each individual map and number of causal links in each map (and the actual constructs contained in the map).

Six soldiers were unable to report any failed events for the first and third measurements, and 3 soldiers dropped out of the course during the third stage of the navigation course, so there were nine missing observations. Thus, the total number of cognitive maps was 111, out of a possible 120. The first step in drawing the maps was to carefully convert the interviewees' natural language into standard content categories to enable comparison, on the basis of the raw constructs (causes or outcomes) described by each subject in his own language and the expressions reflecting the causal links among the constructs, as recorded in the interviewer's notes. The two researchers judged each of the constructs of the 111 cause maps. The maps of 2 subjects (1 from each of the two groups) were used by the two judges as a basis for determining the various content categories. There was no need to alter this basic list of categories later. Then, the construct lists of the 20 soldiers who participated in this part of the study were shuffled. Each judge assigned each of the constructs extracted from the failed and success scripts of the 20 soldiers (766 constructs in total) to one or another of the 16 content categories. The list of the content categories and examples of expressions reflecting them appear in Table 1. The interrater reliability coefficient (according to Tinsley & Weiss, 1975) was high (0.96), due to the fact that most of the soldiers used many of the same expressions repeatedly across the three measures. Table 2 displays the number of times each content category appeared in each of the 12 experimental conditions. In the few cases of disagreement, consensus was reached after a short discussion. At this point, the two judges returned to the protocols and determined the cause–effect relationships of each of the constructs. This task was quite simple, as each construct was phrased as a cause or an effect for example, "I selected the wrong route because I did not have enough time to study yesterday" (therefore, learning time → performance) or "I did not find all the points along the route because I did not feel well" (therefore, feelings → performance), and so on.

To facilitate comprehension of the data pattern in each of the two experimental groups, group maps were constructed that reflected both the structure and content of the individual maps. Regarding structure, we characterized the general structure of each individual map in terms of number of direct or indirect cause–effect relations and number of indegrees and outdegrees of each cause and effect. We tried to ensure that the group maps would reflect the average scores of each of these characteristics. Direct cause–effect relations are links between constructs that are not mediated by any other construct. For example, in Figure 1 (FAER group), studying has a direct link to failure (studying → failure). Indirect links are links between constructs that are mediated by other constructs. In the same table, it is shown that studying also has indirect relations with failure (studying → performance → failure). Indegrees are the paths leading from all the constructs to a particular construct. For example, failure in Figure 1 (FAER group) has 4 indegrees. Outdegrees are the paths leading from a particular construct to other constructs. For example, in Figure 1 (FAER group) studying has 2 outdegrees (see Bougon et al., 1977). Regarding content, according to Laukkanen (1994), a group map is actually the intersection of all the individual maps on which it is based. In other words, the group map contains all the causal links that appear in all the individual maps. Because complete overlap of all the individual maps is practically impossible, and for the group maps to be meaningful, we decided to include in them only those causal links that appeared in at least three of the individual maps; thus, we omitted any links that appeared in fewer than three individual maps. When the number of links that appeared three times did not match the average number of causes mentioned by the particular group, causal links that appeared only twice were added to the map as well. These links are represented in the figures by dotted lines. For example, the

Table 1
Content Categories of Causal Explanations of Failed and Successful Navigation Events

Category	Examples
Planning	1. I selected a wrong route . . . 2. My plan was based on using the most salient point along the navigation route.
Studying	1. I changed the learning method of the navigation route details. 2. My learning method suited the topographical conditions.
Feelings	1. I had a feeling of low self-confidence. 2. I felt like I was not sure that . . .
Weather	1. The day was foggy . . . 2. It was a full-moon night.
Ground conditions	1. The area was quite easy to navigate . . . 2. It was a mountainous area . . .
Navigation conditions	1. We were ordered to navigate without a compass . . . 2. It was my first experience of navigating alone . . .
Performance	1. I tried to take a shortcut . . . 2. I missed an important curve . . .
Navigation facilities	1. I had to open a map. 2. I did not use the compass at that point . . .
Experience	1. I used my experience from the previous navigation day. 2. I was already familiar with the area . . .
Physical conditions	1. I was very tired . . . 2. I hurt my leg.
Motivation	1. It was very important to us to find the ordinate. 2. I tried hard to be on time.
Pace	1. I proceeded very slowly. 2. I increased the walking pace . . .
Knowledge	1. I knew this route very well. 2. I had no knowledge of how to solve this problem.
Luck	1. I found 2 ordinates by chance . . . 2. It was just luck that I took this direction . . .
Learning time	1. I invested a lot of time in learning the route. 2. I did not ascribe importance to this issue and did not spend much time on studying it.
Ability	1. I am an expert. 2. I have excellent spatial ability . . .

average number of causal links of the FAER group in the second measurement was seven, but only six links appeared three or more times. Therefore, the links between planning and failure, feeling and performance, and weather and performance were also included (see Figure 1). In the statistical analysis, these links were counted only twice. Finally, in some cases, the content analysis created incongruence between the structure of the group map and the individual maps. For example, we had individual maps containing two or three performance nodes that were included in one category following the content analysis. Because we wanted the group map to reflect the structure of the original individual maps, it contained more than one node from the same category (e.g., performance); that is, two or more nodes in such a map carried the same name (performance → performance).

Results

Performance of Navigation Exercises

Not all the soldiers completed the full set of navigation exercises; 76 performance scores were available for inclusion in the statistical analysis. Table 3 shows performance scores for the three navigation exercises held during the 2nd training week (the second stage of the study, after the new method of event reviews had been presented to Company B, the FSAER group).

Performance data of the FSAER and FAER groups from the third stage could not be meaningfully compared because of the different training areas and assignments. Therefore, only the measurements taken during the second stage were included in the statistical analysis of the performance data. The three before-treatment measures (taken during the 1st training week) were averaged and used as a covariate. As the AER manipulation could not affect the performance scores of the first navigation of the 2nd training week (the AER took place at the end of the navigation day), we added it to the three measures of the 1st week. The mean score of these four measures was used as a covariate. A 2 (between-groups [FSAER vs. FAER]) × 3 (within-group [three measurements from the second stage]) analysis of covariance and trend analysis were performed. The analysis of covariance yielded a significant linear training effect, $F(2, 74) = 7.74$, $p < .007$, $\eta^2 = .095$, which means that soldiers in both groups made significant progress during the 2nd navigation training week. No main effect of group was found. The research hypothesis regarding performance (Hypothesis 3) predicted that the progress of the FSAER group in the 2nd training week would be greater than that of the FAER group. The significant linear Training × Group interaction on the performance data, yielded by trend analysis, confirmed this hypothesis, $F(1, 74) = 3.795$, $p < .05$, $\eta^2 = .049$. More specifically, the linear trend of the performance scores of the FSAER group in the 2nd training week was significantly stronger than that of the FAER group. The fact that the navigation assignments became more difficult from day to day made these results even more impressive.

Analysis of the Cognitive Cause Maps

Table 4 presents the mean number of nodes in the maps of the subjects within each of the 12 experimental conditions, and Table 5 displays the number of links. Because the starting number of observations in each experimental condition was only 10, and because of the missing observations, we used nonparametric tests to analyze the data presented in Tables 4 and 5. We used the Friedman analysis of variance by ranks. Kendall's statistic (W), the coefficient of concordance, was used to represent effect size. The coefficient W is closely related to the average *r*s among *m* rank orders (see Hays, 1988).

To test the first hypothesis, namely, that mental models of failed events are basically richer than those of successful events, we used the data from the first measurement (taken before the experimental manipulation) to make a comparison within each of the two groups (FSAER and FAER) between the number of nodes and links in the cause map of failed events and successful events. In both groups, the Friedman tests yielded significant results, demonstrating that the cause maps of the failed events, as expected according to Hypothesis 1, were richer in nodes and links than were those of the successful events (FSAER: $Fr_{(1)} = 8.0$, $p <. 01$, $W = 1.00$, and $Fr_{(1)} = 8.0$, $p < .01$, $W = 1.00$; FAER: $Fr_{(1)} = 5.44$, $p < .05$, $W = .60$, and $Fr_{(1)} = 5.44$, $p < .05$, $W = .60$).

Table 2
Frequency of Constructs Within Each Content Category[a]

	Measurement											
	1st				2nd				3rd			
	FSAER		FAER		FSAER		FAER		FSAER		FAER	
	S	F	S	F	S	F	S	F	S	F	S	F
Performance	13	16	7	25	14	32	10	31	15	20	24	31
Studying	3	7	4	8	8		10	7	12	9	6	
Feeling	7	7	6	11	7	9		3	4			3
Planning		3		5	5	5		4	4	3	7	5
Topographical conditions	3			5	4	5	3	4	6	3	4	
Experience	5	5			6	4			5		7	
Studying time		4				4			4	3		
Knowledge		4			3	5			7			
Navigation conditions					8		3		3	4	3	3
Physical conditions						4		3		3		
Weather								3	4			
Navigation facilities					3			4				
Time										5		
Pace									3			
Motivation									5			

Note. FSAER = failure-focused and success-focused after-event review; FAER = failure-focused after-event review; S = successful navigation; F = failed navigation.
[a] Frequency of at least three constructs.

Tables 4 and 5 also reveal, in accordance with our second hypothesis, that only in the FSAER group were significant differences found between the first and second measurements of the complexity of the cause maps of successful events, $Fr_{(1)} = 9.00$, $p < .01$, W = .90, and $Fr_{(1)} = 9.0$, $p < .01$, W = .90, respectively. Significant differences in the number of nodes and links were also found between the second and the third measurements, $Fr_{(1)} = 4.5$, $p < .05$, W = .50, and $Fr_{(1)} = 5.44$, $p < .05$, W = .60, respectively, which means that the cause maps of the FSAER group continued to improve with time. No significant differences between the first and second measurements were found in the FAER group. As the FAER group was also exposed to the manipulation after the second measurement, the cause maps of soldiers of that group were also expected to be enriched. In actuality, a significant increase occurred only in the number of links, $Fr_{(1)} = 7.00$, $p < .01$, W = .78; the difference in the number of nodes was not significant at the traditional value of .05, $Fr_{(1)} = 3.57$, $p < .058$, W = .39. In the third measurement, the FSAER group actually showed significantly more nodes and links in their cause maps of successful events than in their cause maps of failed events, $Fr_{(1)} = 4.5$, $p < .034$, W = .56, and $Fr_{(1)} = 4.5$, $p < .034$, W = .56, respectively.

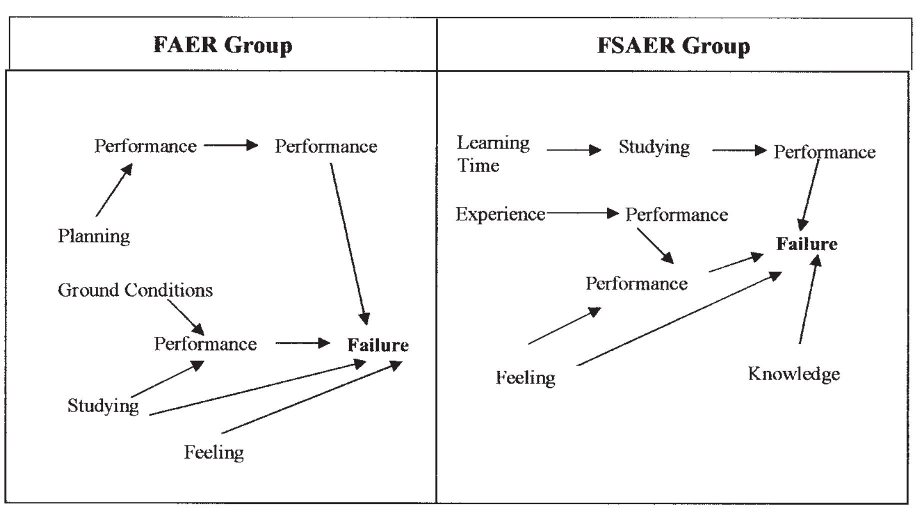

Figure 1. Cause maps pattern of failed navigation: first measurement. FAER = failure-focused after-event review; FSAER = failure- and success-focused after-event review.

Table 3
Means and Standard Deviations of Performance of Three Successive Navigation Exercises

	Kind of AER			
	FSAER		FAER	
Measurement	*M*	*SD*	*M*	*SD*
1st	46.40	20.06	54.41	28.21
2nd	63.92	19.74	65.66	30.71
3rd	76.42	18.82	71.30	26.06

Note. AER = after-event review; FSAER = failure-focused and success-focused after-event review; FAER = failure-focused after-event review.

Qualitative Analysis of the Cause Maps

The final 12 group maps representing the "mental model pattern" of the subjects within each of the experimental conditions are presented in the figures. These group maps were drawn for qualitative analysis only and not for any kind of statistical analysis. The statistical analysis was performed on the individual maps as described above.

Qualitative analysis of the data corroborated the findings of the quantitative analyses. More specifically, as expected, the cause maps of failed events of the first measurement were richer than those of successful events (see Tables 4 and 5 and Figures 1 and 2). Their nodes were more numerous and varied, and they had a greater number of direct and indirect causal links.

Figure 3 and Figure 4 describe the second measurement (taken at the end of the first stage of the navigation course). They demonstrate that although there were still substantial differences between the success and failure cause maps of both groups, the cause maps of the two groups were already clearly different from each other. The maps of the FSAER group contained more nodes and more causal links (especially indirect ones) between the various causes and the successful outcome. Also, note that the new categories in the success maps were associated with three basic elements affecting the learners' mental models: planning, experience, studying (as opposed to situational explanations such as performance and feeling). Furthermore, the path structure shows that subjects were looking for root causes reflecting issues of planning and studying. As shown in Figure 3, this picture was not so clear for the FAER group.

The cause maps of the failed events revealed a similar pattern. The maps of the FSAER group showed a clearer path structure than did those of the FAER group; the majority of the root causes related, as expected, to the building blocks of their mental models (planning, learning time, and studying) rather than to performance or feeling.

Figure 5 and Figure 6 depict the cause maps of the two groups after both had been exposed to the innovative training method. They demonstrate the progress made by both the FSAER and FAER groups. The success maps of both groups were quite complex and included causes for successful events, such as planning and experience, as well as situational causes. Furthermore, for both groups, the successful event maps were richer and more complex than were the maps of failed events.

It should be noted that differences between the mental models of successes and failures were relatively small and nonsignificant in the first two measurements of the FSAER group. By contrast, in the FAER group, there were more constructs in the failure maps in the two first measurements, but the gap decreased in the third measurement, after this group had also been exposed to the manipulation.

It is interesting that the FAER group's cause maps of failed events became quite similar to those of the successful events, demonstrating the impact of contemplating failed events as well. In sum, the qualitative analyses corroborated the quantitative results.

Discussion

Our main thesis in the present study was that lessons might be learned not only from failed events but also from successful ones. Though both are valuable, learning from failed experience and learning from successful experience differ in two respects: (a) Whereas failures motivate epistemic processes like hypothesis generation and information gathering, successes tend to halt this process; and (b) the identification of flaws in learners' mental models might be easier in learning from failed events than in learning from successful events. In our field study, we created an experimental opportunity for learners to systematically contemplate their successes as well as their failures. We argued that debriefing of successful events would elicit more systematic think-

Table 4
Means and Standard Deviations of Number of Constructs in the Cause Maps of Failed and Successful Events

	FSAER group				FAER group			
	Failed event		Successful event		Failed event		Successful event	
Measurement	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Before treatment	7.12	1.80	4.20	1.31	6.78	1.20	3.80	1.68
After 1st week	9.20	2.44	7.30	2.11	6.70	2.02	3.99	1.37
After 2nd week	7.37	1.68	9.11	1.53	6.87	2.74	7.11	2.74

Note. FSAER = failure-focused and success-focused after-event review; FAER = failure-focused after-event review.

Table 5
Means and Standard Deviations of Number of Causal Links in the Cause Maps of Failed and Successful Events

	FSAER group				FAER group			
	Failed event		Successful event		Failed event		Successful event	
Measurement	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Before treatment	8.37	2.92	4.50	1.50	7.11	1.45	3.80	1.68
After 1st week	10.00	2.94	7.70	2.40	7.10	2.28	4.00	1.49
After 2nd week	8.00	2.56	10.22	1.71	9.00	2.44	7.66	2.82

Note. FSAER = failure-focused and success-focused after-event review; FAER = failure-focused after-event review.

ing (that would yield richer mental models) and eventually improve performance.

The results supported the hypotheses. First, a significant learning effect was found across experimental conditions. The level of difficulty of the experimental tasks (the navigation assignments) increased steadily from day to day, and subjects' performance improved in spite of this inhibiting factor; we may thus conclude that the learning effect is even stronger than that indicated by the data. Second, we found that the learners' performance improvement was even greater in the FSAER group. It should be noted that the starting point of the two groups was different; that is, the FAER group performed better in the 1st week of the first stage of the study. This artifact might have jeopardized the internal validity had we been looking for significant main effects of groups (FSAER vs. FAER). However, we neither expected nor found such a significant main effect. In addition, because the FSAER group's starting point was lower than that of the FAER group, the differences in improvement between the two groups could be attributed to artifacts like regression to the mean or to initial motivational differences between them. It should be noted that the members of the FSAER group not only equaled those of the FAER group during the navigation course but also gained better scores. Because the two groups worked completely independently and did not communicate along the first phase of the course, the motivational artifact does not hold either.

Analysis of the learners' mental models provided us the opportunity to take a deeper look at the process of drawing lessons from experience and helped us to understand the differential functions of failure and success in the learning process. The analysis confirmed our expectation that people usually tend to generate more causes for their failures than for their successes. These findings also corroborate previous findings that people's information-seeking responses are oriented toward explanation-relevant information when unexpected events occur (Hastie, 1984; Wong & Weiner, 1981).

In accordance with our expectations, throughout the navigation course, the FSAER group invested increasingly more systematic cognitive effort than did the FAER group in investigating their successes as well as their failures, and their performance continued to improve. Furthermore, when individuals in the FAER group were also debriefed on their successes, their mental models changed accordingly. It seems that encouraging contemplation of successful events stimulated the learners to generate more hypotheses regarding their performance. In sum, in light of these findings,

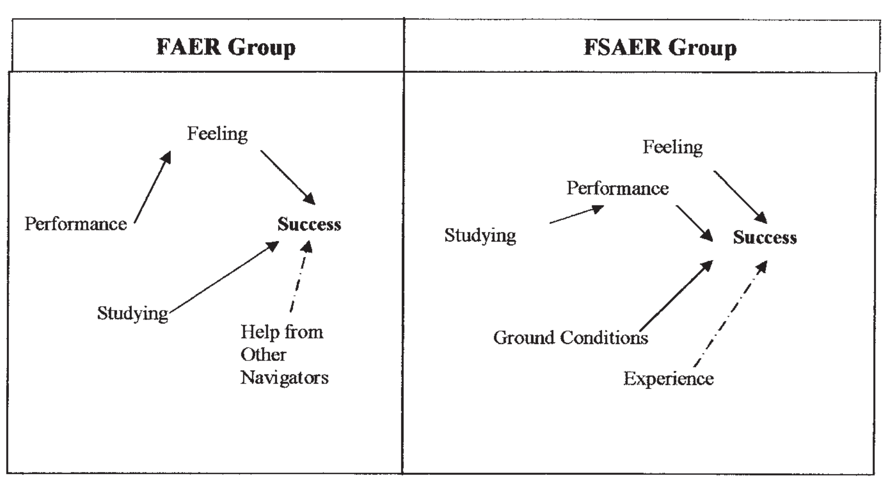

Figure 2. Cause maps pattern of successful navigation: first measurement. FAER = failure-focused after-event review; FSAER = failure- and success-focused after-event review.

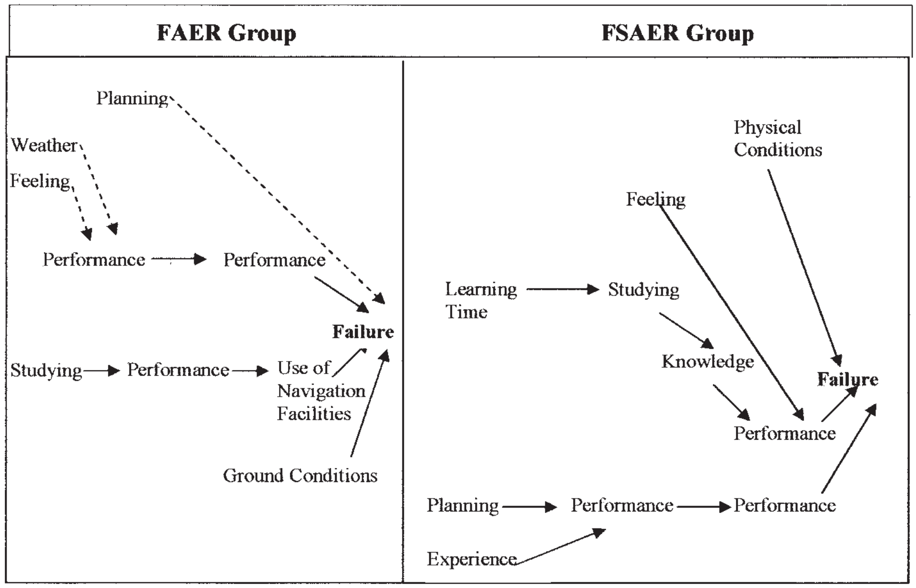

Figure 3. Cause maps pattern of failed navigation: second measurement. FAER = failure-focused after-event review; FSAER = failure- and success-focused after-event review.

one may argue that the characteristics of mental models (number of nodes and links) might mediate the effect of type of AER on learning from experience. Because we had only a sample of maps from each experimental condition, we could not test it statistically in the present study.

The fact that people usually invest more effort in trying to analyze their behavior after failure was reflected not only in the greater number of constructs in the maps of the failed events (first measurement) but also in their structure and kind. In addition, qualitative differences were also detected. In the first measurement (before exposure to the experimental manipulation), there were qualitative differences between maps of failures and maps of successes. Whereas successful events were explained mainly by direct causal effects, failed events were associated with more complicated mental models, that is, models characterized by longer causal paths and causal explanations relating to the preparation phase of the navigation exercise (initial mental model). In the second and third measurements, participants' mental models of successful events became increasingly more complex and included increasingly more causal explanations that did not reflect situational reasons but instead reflected issues relating to their prior knowledge and task planning. It has been shown, for example, that

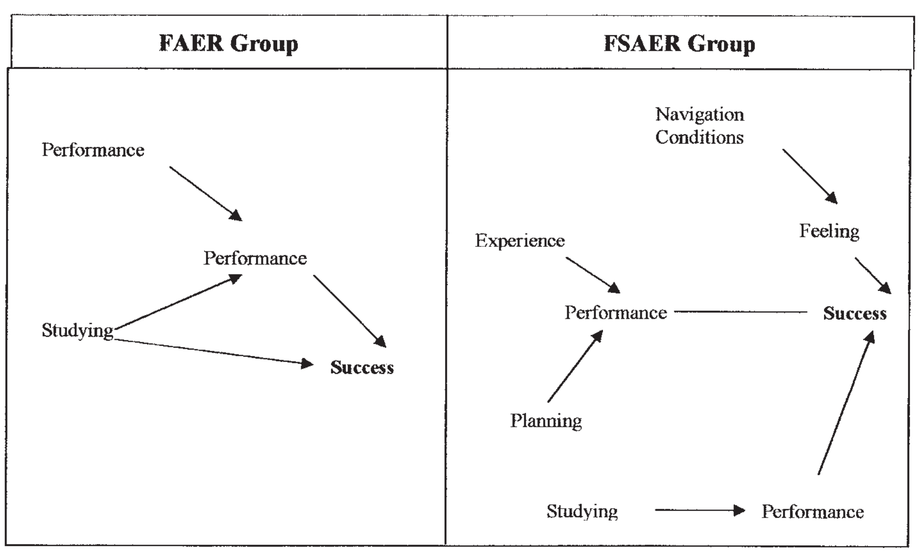

Figure 4. Cause maps pattern of successful navigation: second measurement. FAER = failure-focused after-event review; FSAER = failure- and success-focused after-event review.

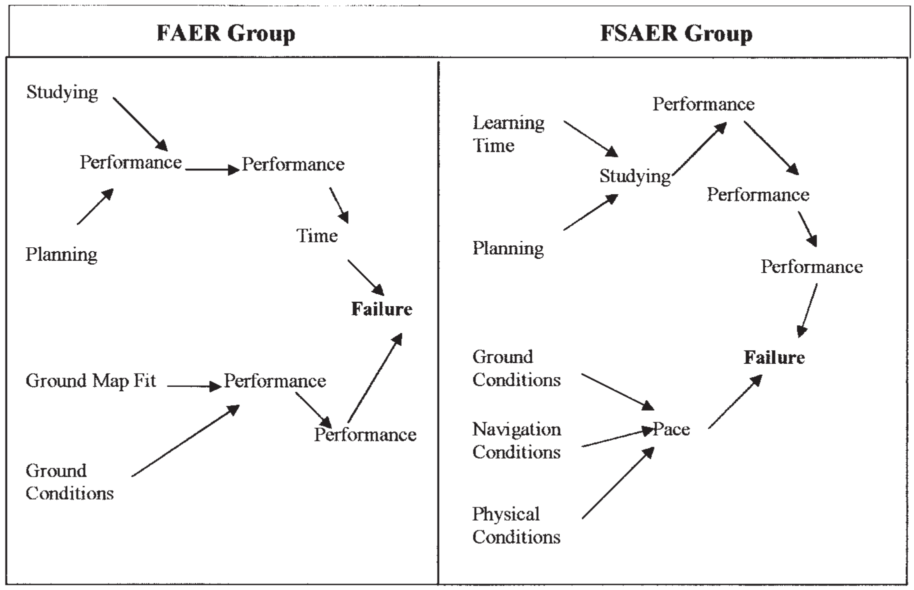

Figure 5. Cause maps pattern of failed navigation: third measurement. FAER = failure-focused after-event review; FSAER = failure- and success-focused after-event review.

relative number of constructs reflecting planning, experience, and studying (as opposed to context explanations such as ground conditions and feeling) in the success maps versus the failure maps was higher in the FSAER group than in the FAER group, which means that when individuals try to draw lessons from successful experience, they must figure out whether their performance was consistent with their prior plans and investment.

It should be noted that there are more attributional dimensions that we can apply to characterize the content of the learners' knowledge structures to better understand the learning process. Wong and Weiner (1981) argued that because individuals generally have better control over their own actions than their physical or social environment, those who behave adaptively will probably have causal explanations comprising relatively more internal than external causes to explain their performance (success or failure). These people take more responsibility for their performance and can face the next challenge more adequately. Thus, we can argue that causal structures of successes will be comprised of relatively more internal (as opposed to external) causes. Furthermore, under the assumption that AER is a kind of a guided investigation of past experience that directs learners to understand the real causes for their successes or failures, one may expect that, following AER,

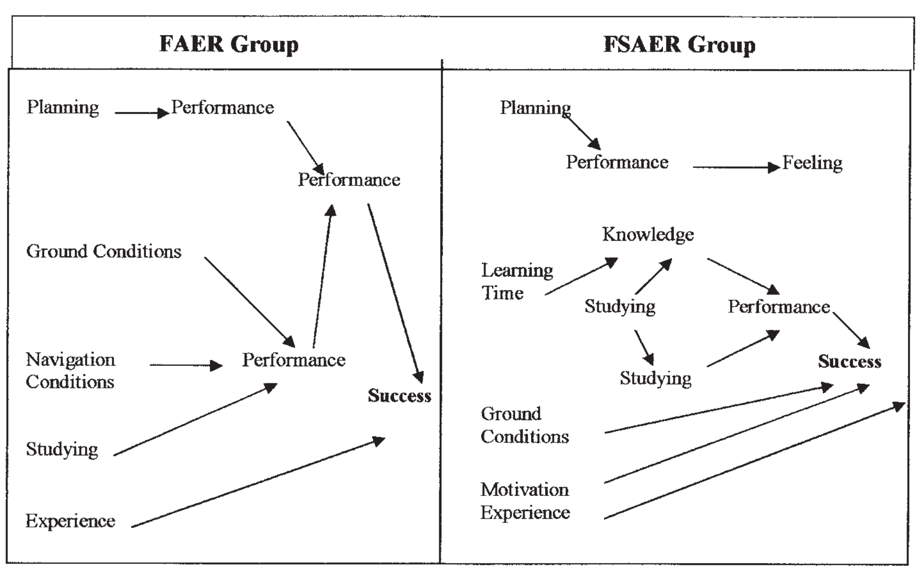

Figure 6. Cause maps pattern of successful navigation: third measurement. FAER = failure-focused after-event review; FSAER = failure- and success-focused after-event review.

learners' causal explanations will contain even more internal causes (as opposed to external causes) than will the causal explanations of learners who did not participate in after-action review. Finally, we can also argue that a general cause of an outcome is less informative than a specific cause, or that knowledge of specific factors that lead to a specific performance are more useful for guiding behavior in a subsequent performance (Gronhaug & Falkenberg, 1998). Therefore, we can argue that causal structures of successes will be comprised of more specific (as opposed to general) causes than will causal structures of failures and that, after AER, learners' causal explanations will contain even more specific than general causes. All these important arguments should be systematically investigated in future research.

Limitations

Although the results generally support our basic thesis that reviews of successful experience may contribute to the learning process and thereby promote performance, the study suffered from several limitations and left us with a number of questions that can be resolved only in true laboratory experiments.

First, the quasi-experimental setting of our study did not allow us to examine our basic claim that inducing individuals to think harder on their previous experience might start a learning process that ultimately improves their performance. Including a control group (no AER) in the experimental design, against which we can test the contribution of AERs in general to learning, may enable us to test this claim. Second, because the field study was conducted according to rules dictated by the hosting organization (IDF), we could not directly test the relative contribution of drawing lessons from "successes only" (as compared with "failures only" or with "no AER") to performance improvement. Third, the findings of the present study are limited to situations in which learners have definitively right and wrong answers. It is our belief that the present findings apply also to judgmental situations as long as learners receive definitive feedback on their performance or are able to interpret their performance in terms of success or failure (Ellis & Kruglanski, 1992). According to Schul (1992), "When individuals know how they have performed, as when they are given accurate feedback about their performance, their assessment of performance is likely to coincide with their actual performance" (p. 167). Fourth, as previously noted, although our theory strongly implies that richness of cognitive maps mediates the effect of AER on performance, our data did not allow us to test this effect statistically. Finally, one more comment regarding the procedure: In our study, Inbar Davidi observed all AERs of both the FAER and the FSAER groups to ensure that the AER was performed according to the assigned method. Although Inbar Davidi was a passive observer, the fact that she was familiar with the hypotheses could have jeopardized the validity of the findings.

Conclusions

In spite of its weaknesses, the study provides some important indications of how successful events, as opposed to failures, are processed; the questions of exactly how leaning occurs and how to conduct the AER to maximize learning effectiveness are still open and call for further research. The present study constitutes a complementary contribution to Sitkin's (1992) suggestion for organizations to develop a strategy of learning from small losses. Sitkin (1992) argued that learning from successful outcomes is hampered by liabilities such as reduced motivation to learn and adapt, complacent behavior, reduced attention and information gathering, risk aversion, and tendencies to develop homogeneous attitudes or behaviors on the basis of successful outcomes. To prevent organizations from resting on their laurels or developing learning patterns reflecting failure avoidance, Sitkin called to encourage learning by experimentation that produces new ideas leading to organizational renewal and readjustment without invoking harmful frustrations. According to his plan, modest levels of failure can promote willingness to take risks and foster resilience-enhancing experimentation.

Although the idea to design an organizational strategy of learning from carefully planned small losses makes sense and seems promising, its implementation is neither easy nor risk free, especially in high-risk/high-reliability organizations. Such organizations emphasize total elimination of errors and the absence of trial-and-error learning (Weick, 1987). If such organizations do not allow for inevitable errors, they limit the degree of trial-and-error learning that is based on these errors (LaPorte & Consolini, 1991). If organizations that try to learn must identify events that require triggering revision of situation assessment and plans or refreshing their evidence collection and evaluation tactics (Woods, 1988), asking the right questions about successful experience might suffice. Reliable outcomes can result from stable processes of cognition directed at varying processes of production that uncover and correct unwanted consequences (Weick et al., 1999). Mindful organizational thinking may be stimulated not only by learning from failed experience but also by consistent elaboration of successes. As was found in the present study, individuals' need for closure can be increased not only by manipulating the situation (according to Sitkin, 1992) but also by conducting other types of AER.

References

Abelson, R. P. (1976). Script processing in attitude formation and decision making. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior* (pp. 33–46). Hillsdale, NJ: Erlbaum.

Adams, J. A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behavior, 3,* 111–149.

Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research, 61,* 315–343.

Allwood, C. (1984). Error detection processes in statistical problem solving. *Cognitive Science, 8,* 413–437.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Argyris, C., & Schon, D. A. (1978). *Organization learning: A theory of action perspective.* Reading, MA: Addison Wesley.

Ariely, D., & Zakay, D. (2001). A timely account of the role of duration in decision making. *Acta Psychologica, 108*(2), 187–207.

Axelrod, R. M. (1976). *Structure of decision: The cognitive maps of political elites.* Princeton, NJ: Princeton University Press.

Balzer, W. K., Doherty, M. E., & O'Connor, R. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin, 106,* 410–433.

Bartunek, J. M., Gordon, J. R., & Weathersby, R. P. (1983). Developing "complicated" understanding in administrations. *Academy of Management Review, 8*(2), 273–284.

Berkson, W., & Wettersten, J. (1984). *Learning from error: Karl Popper's psychology of learning.* La Salle, IL: Open Court.

Bougon, M., Weick, K., & Binkhorst, D. (1977). Cognitions in organizations: An analysis of the Utrecht Jazz Orchestra. *Administrative Science Quarterly, 22,* 606–639.

Brehmer, B. (1980). In one word: Not from experience. *Acta Psychologica, 45,* 223–241.

Busby, J. S. (1999). The effectiveness of collective retrospection as a mechanism of organizational learning. *The Journal of Applied Behavioral Science, 35*(1), 109–129.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Carley, K. M. (1997). Extracting team mental models through textual analysis. *Journal of Organizational Behavior, 18,* 533–558.

Chanowitz, B., & Langer, E. J. (1980). Premature cognitive commitment. *Journal of Personality and Social Psychology, 41,* 1051–1063.

Chi, M. T. H., Bassok, M., Lewis, M., Reinman, P., & Glaser, R. (1989). Self-explanation: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145–182.

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439–477.

Ellis, S., & Kruglanski, A. W. (1992). Self as epistemic authority: Effects on experiential and instructional learning. *Social Cognition, 10,* 155–177.

Enzle, M. E., & Schopflocher, D. (1978). Instigation of attribution process by attribution questions. *Personality and Social Psychology Bulletin, 4,* 595–599.

Evans, J. St. B. T. (1988). The knowledge elicitation problem: A psychological perspective. *Behavior and Information Technology, 7,* 111–130.

Feldman, J. (1989). On the difficulty of learning from experience. In H. P. Sims & D. A. Gioia (Eds.), *The thinking organization* (pp. 263–292). San Francisco: Jossey-Bass.

Ferguson-Hessler, M. G. M., & de Jong, T. (1990). Studying physics texts: Differences in study processes between good and poor performers. *Cognition and Instruction, 7,* 41–54.

Fischhoff, B. (1982). For those condemned to study the past: Heuristics and biases in hindsight. In D. Kahanman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases.* Cambridge, England: Cambridge University Press.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition.* Reading, MA: Addison Wesley.

Freund, T., Kruglanski, A. W., &. Schpitzajzen, A. (1985). The freezing and unfreezing of impressional primacy: Effects of need for structure and fear of invalidity. *Personality and Social Psychology Bulletin, 11,* 479–489.

Friedlander, F. (1983). Patterns of individuals and organizational learning. In S. Srivastava & Associates (Eds.), *The executive mind: New insights on managerial thought and action.* (pp. 192–220). San Francisco: Jossey-Bass.

Gioia, D., & Manz, C. C. (1985). Linking cognition and behavior: A script processing interpretation of vicarious learning. *Academy of Management Review, 10,* 527–539.

Gioia, D. A., & Poole, P. P. (1984). Scripts in organizational behavior. *Academy of Management Review, 9,* 449–459.

Gronhaug, K., & Falkenberg, J. S. (1998). Success attributions within and across organizations. In C. Eden & J. C. Spender (Eds.), *Managerial and organizational cognition: Theory, methods and research* (pp. 93–107). Thousand Oaks, CA: Sage.

Hannan, M. T., & Freeman, J. (1984). Structural inertia and organizational change. *American Sociological Review, 49*(2), 149–164.

Hastie, R. (1984). Causes and effects of causal attribution. *Journal of Personality and Social Psychology, 46,* 44–56.

Hays, W. L. (1988). *Statistics.* New York: Holt, Rinehart, & Winston.

Heimbeck, D., Frese, M., Sonnentag, S., & Keith, N. (2003). Integrating errors into the training process: The function of error management instructions and the role of goal orientation. *Personnel Psychology, 56,* 333–361.

Herriot, P., Chalmers, C., & Wingrove, J. (1985). Group decision making in an assessment centre. *Journal of Occupational Psychology, 58,* 309–312.

Higgins, E. T., & Lurie, L. (1983). Context, categorization, and memory: The "change of standard" effect. *Cognitive Psychology, 15,* 525–547.

Huber, G. P. (1991). Organizational learning: The contributing processes and literatures. *Organization Science, 2,* 88–115.

Huff, A. S. (Ed.). (1990). *Mapping strategic thought*. Chichester, England: Wiley.

Kiesler, S., & Sproull, L. (1982). Managerial response to changing environments: Perspectives on problem sensing from social cognition. *Administrative Science Quarterly, 27,* 548–570.

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of Management, 20,* 403–437.

Kruglanski, A. (1980). Lay epistemologic process and contents: Another look at attribution theory. *Psychological Review, 87,* 70–87.

Kruglanski, A. W. (1989). *Lay epistemics and human knowledge: Cognitive and motivational bases.* New York: Guilford Press.

Kruglanski, A. W. (1991). Lay epistemic theory in social-cognitive psychology. *Psychological Inquiry, 1,* 181–197.

Kruglanski, A. W., & Klar, Y. (1987). A view from a bridge: Synthesizing the consistency and attribution paradigm from a lay epistemics perspective. *European Journal of Social Psychology, 17,* 211–241.

LaPorte, T. R., & Consolini, P. (1991). Working in practice but not in theory: Theoretical challenges of high reliability organizations. *Journal of Public Administration Research and Theory, 1,* 19–47.

Lau, R. R., & Russell, D. (1980). Attributions in the sports pages: A field test of some current hypotheses in attribution research. *Journal of Personality and Social Psychology, 39,* 29–38.

Laukkanen, M. (1994). Comparing cause mapping of organizational cognitions. *Organization Science, 5,* 322–344.

Lipshitz, R., & Popper, M. (2000). Organizational learning in a hospital. *Journal of Applied Behavioral Science, 36,* 345–361.

Lord, R. G., & Levy, P. R. (1994). Moving from cognition to action: A control theory perspective. *Applied Psychology: An International Review, 43,* 335–365.

Louis, M. R., & Sutton, R. I. (1991). Switching cognitive gears: From habits of mind to active thinking. *Human Relations, 44,* 55–76.

Maheswaran, D., & Chaiken, S. (1991). Promoting systematic processing in low-motivation settings: Effect of incongruent information on processing and judgment. *Journal of Personality and Social Psychology, 61,* 13–25.

March, J. G., Sproull, L. S., & Tamuz, M. (1991). Learning from samples of one or fewer. *Organizational Science, 2,* 1–12.

Marinko, Y. (1991, December). A culture of after-action review. *Israel Air Force Magazine*, 36–41.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior.* London: Holt, Rinehart, & Winston.

Murphy, G. L., & Wright, J. C. (1984). Changes in the conceptual structure with expertise: Differences between real-world experts and novices. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 144–155.

Nathan, M. J., Mertz, K., & Ryan, B. (1994, April). *Learning through self-explanations of mathematical examples: Effects of cognitive load.* Paper presented at the Annual Meeting of the American Educational Research Association. New York, NY.

Neisser, U. (1976). *Cognition and reality.* San Francisco: Freeman.

Ohlsson, S. (1996). Learning from performance errors. *Psychological Review, 103,* 241–262.

Pirolli, P. L., & Recker, M. (1994). Learning strategies and transfer in the domain of programming. *Cognition and Instruction, 12,* 235–275.

Pyszczynski, T., & Greenberg, J. (1987). Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis-testing model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, pp. 297–341). San Diego, CA: Elsevier Academic Press.

Rabkin, D. G. (1995). *The role of explication in noticing: The discovery and use of new variables in decision making.* Unpublished doctoral dissertation, Massachusetts Institute of Technology, Sloan School of Management.

Rumelhart, D. E. (1984). Schemata and the cognitive system. In R. S. Wyer & T. K. Srull (Eds.), *The handbook of social cognition* (Vol. l, pp. 161–168). Hillsdale, NJ, Erlbaum.

Schank, R. (1986). *Explanation patterns: Understanding mechanically and creatively.* Hillsdale, NJ: Erlbaum. fs;65Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82,* 225–260.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*, 225–260.

Schul, Y. (1992). The effect of expectations on satisfaction when feedback about performance is not provided. *Social Cognition, 10*(2), 165–188.

Seifert, C. M., & Hutchins, E. L. (1992). Error as opportunity: Learning in a cooperative task. *Human-Computer Interaction, 7,* 407–435.

Senge, P. M., & Sterman, J. D. (1991). Systems thinking and organizational learning: Acting locally and thinking globally in the organization of the future. In T. Kochan & M. Useem (Eds.), *Transforming organizations.* Oxford, England: Oxford University Press.

Sitkin, S. B. (1992). Learning through failure: The strategy of small losses. *Research in Organizational Behavior, 14,* 231–266.

Starbuck, W. H., & Milliken, F. J. (1988). Challenger: Fine tuning the odds until something breaks. *Journal of Management Studies, 24,* 319–340.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology, 23,* 457–482.

Tinsley, H. A., & Weiss, D. J. (1975). Interrater reliability and agreement of subjective judgments. *Journal of Counseling Psychology, 22,* 358–376.

Trope, Y., & Liberman, A. (1996). Social hypothesis testing: Cognitive and motivational mechanisms. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 493–520). New York: Guilford Press.

Walsh, J. P., Henderson, C. M., & Delighton, J. (1988). Negotiated belief structures and decision performance: An empirical investigation. *Organizational Behavior and Human Decision Making Process, 42,* 194–216.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Education Research, 13,* 21–39.

Weick, K. E. (1984). Small wins: Redefining the scale of social problems. *American Psychologist, 39,* 40–49.

Weick, K. E. (1987). Organizational culture as a source of high reliability. *California Management Review, 29,* 112–127.

Weick, K. E., & Bougon, M. (1986). Organizations as cognitive maps: Charting ways to success and failure. In H. P. Sims, Jr., D. A. Gioia, and Associates (Eds.), *The thinking organization.* San Francisco: Jossey-Bass.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (1999). Organizing for high reliability: Processes of collective mindfulness. In B. Staw & R. Sutton (Eds.), *Research in organizational behavior.* (pp. 23–81). Grenwich, CT: JAI Press.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92,* 548–573.

Weiner, B. (2000). Intrapersonal and interpersonal theories of motivation from an attributional perspective. *Educational Psychology Review, 12,* 1–14.

Wong, P. T. P., & Weiner, B. (1981). When people ask "why" questions, and the heuristics of attributional search. *Journal of Personality and Social Psychology, 40,* 650–663.

Woods, W. A. (1988). What's in a link: Foundations for semantic networks. In A. M. Collins & E. E. Smith (Eds.), *Readings in cognitive science: A perspective from psychology and artificial intelligence* (pp. 102–125). San Mateo, CA: Morgan Kaufmann.

Zakay, D. (1984). The evaluation of managerial decisions quality. *Acta Psychologica, 56,* 49–57.

Zakay, D., Ellis, S., & Shevalski, M. (2004). Outcomes' values and early warning indications as determinants of willingness to learn from experience. *Experimental Psychology, 51*(2), 150–157.

Received June 19, 2003
Revision received August 31, 2004
Accepted September 1, 2004 ■