The Lessons We (Don't) Learn: Counterfactual Thinking and Organizational Accountability after a Close Call

Michael W. Morris
Paul C. Moore
Stanford University

We investigate how individuals learn from imagined might-have-been scenarios. We hypothesize that individuals are more likely to learn when they have responded to an event with upward-directed, self-focused counterfactual thoughts, and, additionally, that this learning process is inhibited by accountability to organizational superiors. Support for these hypotheses was obtained in two studies that assessed learning by aviation pilots from the experience of near accidents. Study 1 analyzed counterfactual thoughts and lessons in narrative reports filed by experienced pilots after actual dangerous aviation incidents. Study 2 involved laboratory experiments in which college students operated a flight simulator under different conditions of organizational accountability.●

No topic in research on cognition in organizations is more prominent than learning from experience (March, 1994), yet in this ever-growing literature on learning, there is a curious deficit. Although many studies have analyzed how lessons are transmitted from one employee to the next (Crossan, Lane, and White, 1999), surprisingly few have addressed the prerequisite questions of how individuals draw lessons from experience in the first place and how the organizational context affects this process. Models of learning from repeated experience have been adapted to explain some cases of learning in organizations (Luthans and Kreitner, 1985; Wood and Bandura, 1989), yet these models do not explain how individuals can draw lessons rapidly without extensive experience. Organizational theorists have called for research on the cognitive strategies through which individuals learn about rare and hazardous events (Carroll, 1998) or novel and ambiguous events (Weick, 1995). Clearly, individuals do not always and cannot always wait for repeated experience with an event before drawing lessons; in these situations, learning may draw on imagination rather than experience. In developing a model of how and when people learn from imagination, we can draw on the psychological literature on counterfactual thinking—thinking about what "might have been" (Roese and Olson, 1995a, 1995b).

Research on the counterfactual thoughts that students generate in response to academic outcomes, such as exam results, indicates that certain types of counterfactual thinking engender lessons about how to improve future performance and, ultimately, performance gains (Roese, 1994). Although all counterfactual thoughts may be logically relevant to learning, only certain types (with a particular structure) have the psychological effect of promoting learning. Building on the distinctions made in prior research, we analyze the type of counterfactual thinking that should foster performance-promoting lessons in organizational settings. The type we identify as efficacious involves more complex and self-critical thinking than other types. For this reason, it is likely to be inhibited in organizational contexts in which individuals feel the pressure of accountability to hierarchical superiors (Lerner and Tetlock, 1999).

We studied learning in an aviation setting with a focus on the pilot's task of avoiding accidents in the air. This focus highlights the desirability of learning from imagination—from the counterfactual thoughts spurred by a close call—in that pilots

0001-8392/00/4504-0737/$3.00.

●

The two authors contributed equally to the research. We are grateful to Michal Tamuz and Richard Tarrell for their collaboration in the initial stages of assembling the Study 1 data set, to participants in the "New Ideas in Decision Making" conference at Northwestern University for helpful reactions to the findings, to Daniel Ames, John Jost, and Jennifer Lerner for helpful comments on a previous draft, and to Associate Editor Daniel J. Brass and three anonymous reviewers for their constructive comments on the manuscript.

do not generally survive actual accidents and hence cannot rely on learning from experience. Yet this focus also highlights the difficulty of learning from imagination. It is not inevitable that performance-promoting lessons will be drawn from a near accident because the experience is ambiguous and can be cognitively construed in various ways. A pilot can focus either on how it could have gone worse, by imagining an actual collision that could have occurred, or on how it could have gone better, by imagining a way that the close call could have been averted entirely. Moreover, in imagining how an alternative outcome could have been produced, the pilot can either focus inwardly, by imagining changes to his or her own actions, or focus externally, by imagining changes to others' actions, to aviation systems, or to the natural environment. These dimensions of ambiguity allow for different types of counterfactual thinking about near accidents. And the way people resolve the ambiguity (the type of counterfactual thoughts they generate) has consequences for whether they subsequently draw performance-promoting lessons from the experience.

In our first study, we analyzed an archival sample of narrative reports by licensed pilots about experiences of near accidents in the air. We compared responses of pilots accountable to organizational superiors (i.e., those flying commercial or military planes) with those of pilots who were not similarly accountable (i.e., those flying private planes) in terms of counterfactual thinking and lessons for future performance. Our second study was a laboratory experiment with college students, piloting a flight simulator, who had been randomly assigned to either the presence or absence of organizational accountability. Students experienced a near accident on the flight simulator and then completed a flight log that tapped their thoughts about the past experience and any lessons for the future.

LEARNING FROM COUNTERFACTUAL COMPARISIONS

In the last decade, cognitive and social psychologists have increasingly emphasized that the way individuals make sense of experienced outcomes is greatly determined by thoughts of what could have been, by comparisons of actual outcomes to counterfactual alternatives (Kahneman and Tversky, 1982; Roese and Olson, 1995a). Research on the antecedents of counterfactual thoughts suggests that they are most frequently provoked by outcomes that are unexpected and harmful or negative in their impact on the individual (Roese and Olson, 1997). Accidents, which have both of these properties, have been a common setting for the study of counterfactual thinking (e.g., Miller, Turnbull, and McFarland, 1990; Williams, Lees-Haley, and Price, 1996).

Social psychologists have distinguished types of counterfactual thoughts that have different functions or consequences. A primary distinction is between upward comparisons of reality to better possible alternatives and downward comparisons of reality to worse possible alternatives. The differential consequences of upward versus downward counterfactual comparisons are much like those of upward versus downward social comparisons (Festinger, 1954). Downward counterfactual comparisons—thoughts about how it could have been worse—improve a person's affective evaluation of the actual outcome. In a striking demonstration of this, Medvec, Madey, and Gilovich (1995) found that Olympic bronze medallists tend to feel happy because of the salient downward comparison with the might-have-been outcome of fourth place, whereas silver medallists feel less happy because of the salient upward comparison with the gold medal outcome.

If upward counterfactual comparisons reduce an individual's satisfaction with reality, why do people draw these comparisons? Research suggests that upward counterfactual comparisons help people draw preparative lessons for improved future performance. Research on academic performance has shown that students generating upward rather than downward counterfactual comparisons are more likely afterwards to articulate intentions to take actions in preparing for their next exam that are likely to increase their success (Roese, 1994, Experiment 2). Additionally, in a laboratory experiment, students who reacted to outcome feedback on the first round of a puzzle task with upward counterfactual comparisons (as opposed to downward comparisons or no comparisons) were more likely to show improved performance on subsequent rounds of the task (Roese, 1994, Experiment 3). Yet the hypothesis that upward comparison facilitates learning should not be uncritically translated from academic settings to more complex organizational settings. If we understand why this relationship holds, then we can determine how the constructs need to be refined to apply to organizational settings.

In examining why there is a causal relationship between upward counterfactual thoughts and performance-promoting lessons, it is prudent first to assure ourselves that they are not identical or overlapping constructs. Sometimes learning is defined as any cognition spurred by an experience, and, of course, by this vague standard counterfactual thoughts would be a form of learning. Yet by a more appropriately restrictive definition of learning as performance-promoting lessons (i.e., specific plans about how to improve one's outcome in the future), counterfactual thoughts are not a form of learning because the object of a counterfactual thought is a *past* outcome, whereas the object of a performance-promoting lesson is a *future* outcome.

Counterfactual thinking engenders lessons for the future by focusing one's attention. Any counterfactual thought focuses attention on a factor (a condition, event, or action) temporally antecedent to the outcome that is mutable or changeable. Upward comparisons, furthermore, focus attention on a factor that has causal potency to make the difference between the actual outcome and a better outcome (Roese and Olson, 1995b). The result is a focused proposition linking a change in a given causal factor to an improved outcome, and hence it can serve as a frame for the construction of a lesson. For example, an upward counterfactual by an aviation pilot after a near accident might be, "If I had understood the controller's words accurately, I wouldn't have initiated the inappropriate landing attempt." This might give rise to a lesson such as, "From now on, when I am given verbal instructions by a controller, I will always repeat the instructions back to be sure I understand correctly." The counterfactual thought has focused the pilot's attention on a particular link (e.g., understanding instructions) in the chain of events leading to the near accident. Starting with the counterfactual statement as a frame, the lesson is constructed by substituting a controllable future action and future outcome. In short, images of how a better outcome *could have occurred* in the past make it easier to construct images of plans for changing actions so that a better outcome *will occur* next time.

By contrast, downward counterfactual comparisons do not directly engender lessons for future improvement. Downward counterfactuals identify what *not to do* in the future so as to avoid worse outcomes than a near miss, but this can be far from specifying what *to do* in the future to assure outcomes better than a near miss. An important research program by Higgins and colleagues (e.g,, Liberman et al., 1999) has found that preventing performance losses, as opposed to promoting performance gains, involves different kinds of causal attributions for outcomes. To illustrate, a common downward counterfactual statement in the aviation domain is "If I had not swerved at the last moment, we would have crashed." This does not serve as the basis for constructing a performance-promoting lesson; the causal antecedent that the pilot highlights (swerving at the last moment) is not one that has potency to bring about the better outcome of a completely smooth, danger-free flight. The relevant antecedent factors for such a lesson are located further back in the causal chain, factors related to what brought the plane onto a collision course in the first place. Although it is possible that an individual, after a downward counterfactual comparison, might be spurred into a many-step inferential process that results in a performance-promoting lesson, this is clearly less probable than such a lesson resulting from an upward counterfactual comparison because of the greater number of inferential steps required.

In considering a complex performance task, such as flying a plane safely in the highly interdependent aviation system, it quickly becomes clear that another distinction is needed to clarify the type of counterfactuals that lead individuals to performance-promoting lessons. Compared with academic tasks, aviation incidents involve many more factors outside of the individual's control, such as mechanical failures, actions by other pilots, instructions from controllers, and so forth. Hence, some upward counterfactual comparisons are not about what the pilot personally could have done differently but what could have been done differently by other pilots, air traffic controllers, or airport designers. Clearly, these other-focused upward counterfactual comparisons do not engender performance-promoting lessons for the pilot in the same direct way as do self-focused comparisons. For self-focused thoughts, there is a direct link from an upward comparison to a behavioral intention and, ultimately, to changed behavior. But for other-focused thoughts, this is not the case. For example, whereas my thought about what *I might have done better* translates directly into an intention about what to do better next time, my thought about what *someone else* *might have done better* does not translate directly into a prescription for my future behavior. In fact, an other-focused upward comparison directs the spotlight of one's attention away from the factors under one's control.

Based on the discussion above, we can hypothesize a link between a type of counterfactual thinking and learning. Of the four types of counterfactual thoughts arising from the cross of direction and subject, only one type—self-focused and upward-directed—works as a cognitive frame for constructing performance-improvement lessons. Importantly, even this type does not necessarily lead to lessons—sometimes the antecedent factors focused on (such as one's unintentional acts) do not correspond to factors that are foreseeably controllable in the future—yet most factors to which this type draws one's attention can serve as the seed of a lesson. Hence,

Hypothesis 1: Individuals are more likely to draw performance-promoting lessons from ambiguous outcomes after they have responded with a self-focused upward counterfactual comparison than after they have responded with other types of comparisons or no comparison at all.

Inhibiting Effects of Accountability

While our first hypothesis describes a cognitive process that may help individuals learn to perform well on challenging organizational tasks, other research programs in psychology suggest that features of an individual's organizational context may inhibit the cognitive processes needed for this learning. In particular, accountability to audiences of particular kinds has predicable effects on individuals' information processing strategies (for a review, see Lerner and Tetlock, 1999). This is not merely a shift in individuals' conscious impression management; careful studies have documented that the changes occur in people's private thoughts, not merely in public statements (Tetlock, 1992).

Accountability may affect an individual's likelihood of generating self-focused upward counterfactual comparisons for two reasons. First, somewhat obviously, this type of counterfactual comparison is more self-critical and self-implicating than other types. At the least, they acknowledge the possibility of a better course of action than the one that was taken. At worst, they can imply negligence or culpability for the event. Second, the generation of self-focused, upward counterfactual comparisons requires more complex cognitive processing of the event than do other comparisons. This is because accidents and close calls occur under one of two scenarios: lapses in the pilot's control or lapses in the pilot's awareness of the impending problem. When a pilot is not in *control* of the accident situation, the most proximal links to the incident are not the pilot's actions but factors in the environment. To identify a point at which the pilot could have exerted a positive influence on the outcome requires tracing the causal chain further back in time, which requires complex and imaginative reasoning. When a pilot is not *aware* of the impending problem, the causal chain is not perceived correctly until it is too late. Hence, the pilot has to reconstruct the events before constructing a counterfactual comparison, which is also a cognitively demanding process.

The self-critical and complex quality of self-focused upward counterfactual thoughts means that individuals in hierarchical organizational contexts may be less likely to generate them. Hierarchical organizational contexts make an individual accountable to organizational superiors, an audience with known views and with power. This threatening form of accountability evokes an information processing strategy of "defensive bolstering," a general tendency to avoid complex or self-critical thoughts (Tetlock, 1983; Tetlock, Skitka, and Boettger, 1989). Studies of judgment have shown that threatening accountability leads individuals to avoid complex, self-critical lines of thought in such work domains as auditing (Peecher and Kleinmuntz, 1991), performance appraisal (Antonioni, 1994), foreign policy analysis (Anderson, 1981), and medical diagnosis (Hendee, 1995).

Not all kinds of accountability relationships evoke a cognitive strategy of defensive bolstering. Accountability to an audience with unknown views can evoke a strategy of "pre-emptive self-criticism," as one tries to anticipate all possible audience critiques (Schlenker, 1986; Tetlock, Skitka, and Boettger, 1989). Also, accountability to an audience that is merely challenging rather than threatening (an audience without control over one's fate) can be a stimulant to self-analysis and motivation (Schlenker, 1986).[1] But our concern here is with threatening accountability to hierarchical superiors, who have known views and control over one's fate, which should inhibit complex or self-critical thoughts. In aviation, this accountability pressure should be felt by pilots embedded in organizational hierarchies but not by private pilots. Hence, these groups should differ in their information processing strategies in response to a near accident: organizational pilots will be more likely to generate simple, self-protective interpretations as opposed to complex, self-critical interpretations and, as a part of this bolstering strategy, should be less likely to generate self-focused upward counterfactual comparisons and, ultimately, performance-promoting lessons. We contrast individuals in organizational and private roles to test the following hypothesis:

Hypothesis 2: Individuals performing under organizational accountability will be less likely than individuals in private settings to draw performance-promoting lessons, an effect mediated by the rate of generating self-focused upward counterfactual comparisons.

To our knowledge, there has been no previous research on the effects of organizational accountability on counterfactual thinking or learning in regard to accidents, but there is evidence that defensive motives shape the counterfactual thoughts that people construct. In studies of the counterfactual conjectures about world history made by political forecasters, Tetlock (1998) has found that defensiveness about being wrong leads these experts to counterfactual conjectures about how they were "almost right." Thus, there is some suggestive evidence consistent with our second hypothesis, although there have been no direct tests.

1
There is also a distinction between retrospective and prospective accountability (see Lerner and Tetlock, 1999). Retrospective accountability applies to accounts given after an event or decision point, and the accounts may show bolstering against criticism, since the past cannot be changed. Prospective accountability applies when the accountable person has knowledge of and control over the upcoming decision point, which is not the case in an accident situation, but it would increase willingness to learn.

To test our hypotheses, we relied on two complementary methods. Study 1 analyzed an archival data set of pilots' reports about near accidents. Although accountability is not manipulated in this field study, its effects can be assessed by comparing reports of incidents by pilots flying privately to reports of similar events by those flying in an organizational context. Study 2 was a laboratory experiment conducted with a flight simulator apparatus, which allowed us to randomly assign participants to the conditions of private or organizational pilot while controlling the flight incident and flight conditions. This laboratory experiment allows findings with greater internal validity than is possible in the archival analysis of naturally occurring incidents. The experiment also provided a conservative test of the effects of accountability: if the pressure created by a minimal laboratory manipulation changes the way people process information, then the effects of a real and enduring relationship to superiors are, no doubt, much stronger. Study 2a was a follow-up with the same apparatus in which we manipulated whether or not participants were instructed to draw self-focused upward counterfactual comparisons and then measured performance improvement on their next flight to test learning.

STUDY 1: ARCHIVAL STUDY OF NEAR ACCIDENTS

We analyzed narrative reports about near accidents submitted by pilots to the Aviation Safety Reporting System (ASRS), a NASA program that compiles rich data about near accidents and other aviation incidents to enable studies of risk factors. This program requires that pilots who experience a near accident complete a reporting form. Initially it queries pilots for an open-ended description of the incident. We coded this part of the report for statements about counterfactual alternatives to the actual incident (references to what *might have been*, *would have been*, or *should have been*). A subsequent optional question asks, "What can be done to prevent a recurrence?" We coded this part for whether the pilot states a specific intention about how to achieve better performance in the future. These two measures allowed a test of hypothesis 1, that pilots who make self-focused upward counterfactual comparisons are more likely to draw lessons. To test hypothesis 2, that accountability to organizational superiors suppresses the crucial type of counterfactual thinking and thereby inhibits learning, we also used a variable in the ASRS record indexing whether pilots were flying their own *private* planes or were flying for commercial or military *organizations*. Another variable was a categorization of the severity of the near accident. By selecting records within levels of this variable we were able to (a) ensure that the private and organizational pilots' reports concerned the same type of incident and (b) explore whether the degree of danger is associated with the rate of counterfactual comparisons and learning.

It is worth saying a bit more about the ASRS to address an obvious concern about whether our distinction between organizational versus private pilots maps onto the presence and absence, respectively, of accountability to superiors. A critic might ask the following: Are not both private and organizational pilots accountable for their behavior in near accidents to aviation authorities, such as the FAA? Does a pilot's felt accountability toward legal authorities after a near accident overwhelm accountability toward organizational authorities? Fortunately, a feature of the ASRS eliminates the problem of legal accountability. As an incentive to file an ASRS report after an incident, pilots who file are given immunity from prosecution for non-criminal violations involved in the incident and are assured that any identifying information is removed from the report before it is added to the data base. Hence, there is no legal accountability to muddy the private versus organizational distinction. The ASRS itself is not a known, controlling audience that would evoke defensive bolstering. Anecdotal reports by pilots suggest that they see the ASRS reporting process more like therapy or confession than like an inquisition. Thus, we can be reasonably sure that the only accountability that affects pilots in the ASRS data set is the accountability of organizational pilots to their superiors. In addition, the incentive of immunity from prosecution leads to a very high rate of reporting by pilots who have been involved in near accidents, mitigating the typical problem of limited and biased response rates in survey data sets. Furthermore, the fact that pilots know that anonymity is preserved mitigates the problem of impression management or self-censorship that plagues many self-report data sets.

Method

Sample of incidents. Our data came from an archival data base compiled by the ASRS. Records of near accidents are compiled by aviation analysts, beginning with the receipt of a report from a pilot. The pilot-report form begins with a few factual questions, such as flight conditions and plane type, and then has open-ended queries for a description and any lessons drawn. Pilots are first supposed to describe the chain of events and human actions that "really caused the problem" and then to state any insights about "what can be done to prevent a recurrence or correct the situation."

The ASRS reports from each year are recorded on a CD ROM. All of the reports we analyzed were from the first six months of 1989.[2] In the selection of ASRS narratives for our sample, we wanted to have a sufficient number of narratives with counterfactual thoughts to permit our planned analyses. Our goal was to gather as many consecutive narratives as necessary to get approximately 80 that contained counterfactual thoughts. To accomplish this, we did an automated search of the narratives for phrases distinctly marking counterfactual statements: "would have," "should have," "ought to have," "may have," and sentences beginning with the word "Had." We found 80 narratives meeting this criterion after searching the first 252 consecutive narratives. Then, we searched manually through the narratives to check and categorize counterfactual thoughts. Through the manual categorization process, we observed that even complex counterfactual statements had very little ambiguity in terms of our coding dimensions: antecedent subject and the direction of the imagined outcome. The manual check revealed, in the end, that 89 narratives contained counterfactual statements and 163 narratives did not, a 96 percent correct identification rate. Although the combination of mechanical and manual search was not strictly necessary, their convergence demon-

2
We chose 1989 reports because that year immediately preceded the implementation of the Traffic Collision Avoidance System (TCAS). In the opinion of a former ASRS analyst who acted as a consultant to this project, TCAS has changed the way pilots think about some categories of dangerous incidents. It is an alarm system that alerts pilots to other planes within a specified distance. It gives the FAA a comprehensive means to police the skies, which has increased the number of reports about incidents that were not in fact dangerous and that are filed out of self-protective motives.

strated that counterfactual statements often take a standardized form that can be codified for machine search.

Independent variables. The ASRS employs analysts who input reports from pilots into the data base. The analysts are aviation experts who review the reports submitted by pilots and code each on certain dimensions. These preexisting variables in the data set are the source of our variables for organizational accountability and outcome negativity. *Organizational accountability* was revealed most precisely by the variable indicating the type of plane. A former ASRS analyst who consulted with our project distinguished airline and military pilots from private pilots on this basis. *Outcome negativity* was coded as the severity of the close call between two planes. "Near Mid-Air Collisions" are those conflicts in which the distance between two planes was less than 500 feet. Conflicts more than 500 feet (but less than the FAA mandate of 3000 feet) are coded by the analysts as "Less Severe." It is appropriate to consider outcome negativity in this study because it has been discussed in the counterfactual-thinking literature, but we treat it more as a control, because one could propose conflicting hypotheses about valence in this setting (see Roese, 1997).

Dependent variables. After we extracted pilots' statements describing counterfactual outcomes, we categorized their *counterfactual thinking* in terms of direction and the antecedent subject. Here are examples of each kind of statement:

Self/Up: I feel that had I reviewed the approach better, I would have been more alert to the difference between the "cleared to" altitude and published intercept altitude.

Other/Up: He did not declare an emergency to my knowledge and this incident could have been avoided by canceling takeoff clearance and to clear runway by tower controller.

Self/Down: If I had not taken evasive action, a collision with either aircraft would have been a certainty.

Other/Down: If aircraft b had followed the locator to 24r, this situation could have been disastrous since he was at a higher speed than we—he may have descended into us from above.

Our theoretical concern with self-focused upward counterfactual thoughts (Self/Up CFTs) led us to isolate this type of counterfactual thought as one category. We created a residual category lumping together other types—upward counterfactual comparisons based on changes to external factors and downward comparisons. One quarter of the narratives contained multiple counterfactual statements. An examination of these cases revealed that pilots were usually expressing the same counterfactual idea in multiple ways. If any of the counterfactual statements were self-focused upward comparisons, the narrative was counted in the first category. This classification yielded 24 narratives with Self/Up CFTs and 65 narratives with other types of CFTs.

To measure learning, pilot narratives were scored for the presence of specific lessons for their own future performance. Two judges, naive to the accountability condition from which the narrative originated, rated each narrative on a three-point scale: 0 = no lesson was drawn; 1 = a vague, nonspecific observation; and 2 = explicit, specific statement about an action that will be taken in the future. The two coders were reliable (α = .74), and their codes were combined into a composite measure of learning. Two examples of explicit, specific learning statements are "I know I will never take a visual approach based on seeing another aircraft unless ATC [Air Traffic Control] can verify that that aircraft is the only one there that I could possibly see," and "When you call your intentions you should state what airport you are at because no one knows whether you are at Colebrook, Lyndonville, or Cape Cod." A broad, all-inclusive definition of learning might include all observations about the relationships between antecedents and outcomes. This would clearly make counterfactual statements a subset of learning. In this paper, however, we use the term individual learning in a more focused sense, referring to the drawing of lessons that will affect the pilot's future behavior. Counterfactual statements have a distinct orientation to the past that may or may not lead to learning that has a material effect in the future. As such, learning is not a subset or a superset of the counterfactual observations we make about the past. The two constructs are distinct, and importantly, counterfactual thoughts do not necessitate learning. First, the situation spurring the counterfactual thought may not present itself again in the future, so the counterfactual thought is an end in itself. For example, its purpose may be to cope with the stressful event, and the pilot may not go to the cognitive effort to draw a lesson. Second, the antecedents of the counterfactual comparison may be beyond the person's control, thus making the learning non-actionable. For example, even if the pilot blames the co-pilot or air traffic controller and draws a lesson about not working with those people again, a pilot would rarely have control of these assignments.

Covariates. We included several other variables in the ASRS records as covariates to control their possible effects on the likelihood of generating particular kinds of counterfactual thoughts and on generating lessons. First, an indicator variable distinguished flights flown in visual meteorological conditions (VMC) from flights flown in instrument meteorological conditions (IMC) and mixed conditions. Flying under visual flight rules is qualitatively different than flying under instrument flight rules (e.g., at night or in bad weather), and we wanted to control for the possible added complexity and stress created by the meteorological conditions.

A second indicator variable distinguished near accidents that were detected by Air Traffic Control (ATC) from incidents that were not detected. ATC keeps a watchful eye over the activities of planes, but it does not monitor all planes at all times. This variable was included because knowing that a flight incident was detected by ATC may affect the way the pilot reports and learns from the incident. Two predictions about the direction of this effect could be drawn. If pilots perceived ATC as a threatening audience to which they were accountable, then its effects should be to reduce learning by inhibiting Self/Up CFTs. Yet if pilots perceived ATC as a non-threatening but stimulating audience, then contact with ATC might lead to more complex information processing (Tetlock, Skitka, and Boettger, 1989; Lerner and Tetlock, 1999). This second possibility is more likely, because ATC does not have a direct ability to impose sanctions on pilots.[3] As such, ATC may stimulate and improve one's ability to take an external and self-critical perspective (Schlenker, 1986), an information processing strategy likely to increase learning. We cannot be certain about how pilots perceive ATC, so we do not state a concrete hypothesis, but we include the indicator variable in our analyses because of its potential power to reduce extraneous variance.

Results

In general, many pilots included counterfactual thoughts in their description and diagnosis of the near accident. Thirty-five percent of all narratives contained some kind of counterfactual statement. Consistent with literature indicating that counterfactual thinking is spurred by negative outcomes, we found that outcome valence made a significant difference in the overall rate with which CFTs were evoked: 43 percent of "More Severe" incidents and only 30 percent of "Less Severe" incidents [χ^2 (1, N = 252) = 4.90, $p < .05$]. As shown in table 1, however, organizational accountability did not have a significant effect on the overall rate of counterfactual thinking: 33 percent of organizational pilots stated a CFT compared with 39 percent of private pilots [χ^2 (1, N = 251) = 1.11, n.s.]. The proportion of Self/Up CFTs, however, was twice as high among private pilots as organizational pilots. As percentages, 38 percent of private pilots' CFTs were Self/Up CFTs versus only 18 percent for organizational pilots' CFTs. A directional trend that fell short of significance was a higher frequency of Self/Up CFTs in more severe than less severe incidents (33 percent versus 20 percent, n.s.).

To test the first hypothesis, that Self/Up counterfactual thinking is a mechanism for learning (more so than other types of counterfactual thinking), we examined how the degree of learning varied as a function of the presence versus absence of a Self/Up counterfactual thought. Results showed a strong effect in the anticipated direction: the degree of learning was higher in the presence of Self/Up counterfactuals [t(251) = 7.08, $p < .001$]. To confirm the legitimacy of isolating narra-

Table 1

Frequency of Types of Counterfactual Comparisons as a Function of Organizational Accountability

	Organizational Accountability		
	Absent (Private pilot)	Present (Organizational pilot)	Total
Self/Up	15 (38%)	9 (18%)	24 (27%)
Other CFTs	24 (62%)	41 (82%)	65 (73%)
Other/Up	11 (29%)	17 (34%)	28 (32%)
Self/Down	6 (15%)	13 (26%)	19 (21%)
Other/Down	7 (18%)	11 (22%)	18 (20%)
Total	39 (100%)	50 (100%)	89 (100%)

χ^2 (1, N = 89) = 4.66, $p < .05$.

3
It is true that pilots must obey ATC directives and can be held responsible for not doing so. The incidents analyzed here, however, are either cases in which ATC was an observer not directly involved or in which ATC provided corrective directions for the pilot to follow.

tives with Self/Up CFTs, we conducted post hoc tests among the other subgroups of narratives. The learning means of each group are shown in table 2. Tukey's Honestly Significant Difference pairwise comparison tests do not indicate differences among the Other/Up, Self/Down, Other/Down, and No CFT groups, but the Self/Up CFT group differs significantly from all of the others.

Our second hypothesis was that Self/Up CFTs mediate the relationship between organizational accountability and learning from experience. Before testing the hypothesis with regression analysis, we checked the degree of collinearity among the independent variables. Partial correlations between the frequency of Self/Up CFTs and our two independent variables, controlling for flight conditions and for detection of the incident by ATC, revealed that self-focused upward counterfactual statements are less frequent under conditions of organizational accountability ($r = -.179$, $p < .01$) and more frequent after incidents in the more severe category ($r = .188$, $p < .01$). Also, private pilots tended to be involved in more severe incidents, meaning that the two planes passed more closely together ($r = -.237$, $p < .01$), which may reflect that smaller planes operate in greater proximity to each other or that their flight paths are less planned in advance. Importantly, however, the degree of collinearity is modest and hence does not threaten the regression analysis.

To assess whether Self/Up CFTs mediate accountability effects on learning, we regressed learning on different combinations of our independent variables. Results are shown in table 3. In an initial model, equation 1, we found a significant effect of organizational accountability but not negativity of outcome. Private pilots drew more specific lessons for future performance than organizational pilots, an effect that can also be seen in a test between mean levels of learning [private mean = .72 versus organizational mean = .52, $t(251) = -2.19$, $p < .05$]. Next, we ran a model testing if the mediating variable predicts the dependent variable (equation 2). Finally, we included all variables in a single model to test whether self-focused upward counterfactual comparisons mediate the relationship between accountability and learning. Equation 3 shows that the coefficient for Self/Up CFT remains highly significant and is virtually unchanged relative to equation 2. The coefficient for organizational accountability is still significant, although it is decreased in magnitude and significance rela-

Table 2

Pilot Learning as a Function of Counterfactual Thinking*

Counterfactual direction and subject	Mean	S. D.	N
Self/Up	1.48_a	.77	24
Other counterfactuals	$.62_b$	.65	65
Other/Up	$.66_b$	.72	28
Self/Down	$.61_b$	.64	19
Other/Down	$.56_b$	.59	18
None	$.46_b$	.61	163
Total	.60	.70	252

*Rows sharing the same subscript are not significantly different by Tukey's Honestly Significant Difference pairwise comparison tests.

Table 3

Regressions on Individual Learning in Study 1 (ASRS Data)*

	1	2	3
Independent variables			
Organizational accountability	–.253••		–.174•
(private plane = 0; commercial or military plane = 1)	(.093)		(.087)
Negativity of outcome	–.030		–.118
(Less Severe = 0; More Severe = 1)	(.094)		(.089)
Covariates			
Detection by controller	.278••	.175•	.170•
	(.091)	(.082)	(.086)
Flight conditions (visual meteoric conditions = 1;	–.021	.034	.011
instrument and mixed conditions = 0)	(.117)	(.108)	(.108)
Mediating variable			
Self/Up CFT		.938•••	.918•••
		(.139)	(.142)
Adjusted R^2	.046	.172	.182
F	4.03••	18.39•••	12.13•••
D.f.	(4,250)	(3,251)	(5,250)

•p <.05; ••p <.01; •••p <.001.
* Standard errors are in parentheses.

tive to equation 1. This does not definitively demonstrate mediation, but it suggests that Self/Up counterfactual thoughts play some mediating role (Baron and Kenny, 1986).

Discussion

Study 1's results reveal that organizational pilots differ from private pilots in their rate of Self/Up CFTs and, consequently, in their rate of learning from experience. One strength of this archival study is its high external validity. The data indicate that responses by experienced licensed pilots to a representative sample of near accidents depend on the presence or absence of organizational accountability. Another virtue of this archival data is that the Aviation Safety Reporting System (ASRS) narratives, although minimal, provide insights by their rich qualitative nature, providing a window into the chain of inferences with which a pilot reacts to a dangerous incident. Excerpts from these narratives provide insight into the processes confirmed by our quantitative analyses. Our central claim that Self/Up CFTs provide a frame for the construction of a lesson can be seen in the following example of a Self/Up CFT: ". . . I was confused, knew something was wrong, but not quite sure what. In any case, I should have executed a miss on my own accord immediately once I thought something was wrong." This pilot later stated that, in the future, he would act as he wished he had done during the incident being reported.

Though a counterfactual thought can give rise to a lesson, it is worth emphasizing again that the connection between Self/Up CFTs and lessons is not a logical tautology. Self/Up CFTs are neither sufficient nor necessary to arrive at a personal lesson for future behavior. In the following example, the pilot makes a clear Self/Up CFT, but the lesson at which he arrives is a systemic solution beyond his control:

Description: This was a very serious mistake that possibly happened because of an inexperienced radar controller working too much traffic for the weather conditions. He also was possibly burdened by working traffic on separate frequencies.

Self/Up CFT: If I could have heard him working the helicopter, maybe I would have questioned the controller about another aircraft at 7000 feet.

Stated lesson: Controllers need to work only one frequency at a time. When a pilot or controller makes a mistake, instead of an FAA that has an intimidation stance, we need to work through the mistake and help everyone involved learn from it to make flying safer.

It is also interesting to note that while the counterfactual statement is very likely true, it does not make for a compelling personal lesson. The antecedent condition identified is beyond the pilot's control at the time of the incident and in the future.

A final example shows that a Self/Up CFT is not necessary for a pilot to learn a lesson about how to act in the future. The narrative providing the example below had no Self/Up CFT and concluded with an external attribution for the near accident. The lesson, however, is not external. Instead, it focuses squarely on the pilot's personal behavior:

Description: The more oblique angle [of the flight path] relative to the final approach to 16 of ITH R135, plus greater distance, makes it inherently less accurate.

Stated lesson: In the future, I will avoid defining points with NAVAIDS that are distant and subtend a small angle (less than 45 degrees) if at all possible, or request another approach.

Despite the external validity and richness of archival data, there are always limitations that arise from possible third variables and potential alternative explanations. Some of the obvious alternative explanations can be ruled out by the data. First, a critic could argue that even though we controlled for the severity of the near-miss incidents that happened to private and organizational pilots, the severity of the possible accidents is a confounding variable. For example, in a worst-case scenario, larger planes place more people at risk in an emergency landing. So, organizational pilots may be less likely to learn from the experience of a near accident because the pilot's feeling of relief may overwhelm the goal of learning. To test whether this variable determined the difference between conditions, we had our hypothesis-naive ASRS expert rate a large subsample of near-accident narratives (60 percent) for how catastrophic the (avoided) accident would have been, on a scale from "minor accident" to "worst possible aviation accident." We found that catastrophe-level ratings were, in fact, correlated with the size of the plane involved ($r = .666$, $N = 150$, $p < .001$). But there was no relationship between catastrophe level and the pilot's level of learning ($r = -.080$, $N = 150$, n.s.) or pilot's Self/Up CFTs ($r = -.093$, $N = 150$, n.s.), indicating that the greater catastrophe level of the accidents avoided by organizational pilots was not a critical factor in their failure to learn from dangerous incidents.

Another alternative interpretation of our finding is that organizational pilots are affected not by accountability to a known threatening audience but, rather, by a simple audience effect—they are being watched by others in their airline or military organization during a flight. Consciousness of this may take attention away from learning and counterfactual thinking. Some relevant evidence is the effect that Air Traffic Control's (ATC's) detection of the incident had on counterfactual thinking and pilot learning. We found a positive effect of ATC detection on self-focused upward counterfactual statements and on pilot learning. Controlling for other variables (organizational accountability, negativity of outcome, and flight conditions), ATC incident detection still had a significant partial correlation with the presence of Self/Up CFTs ($r = .191$, $N = 246$, $p < .01$). This effect is independent of the organizational accountability effect; the mean learning rating increases equally with ATC detection for both private and organizational pilots. Finally, equation 3 in table 3 indicates that the effect of ATC incident detection is mediated in part by the mechanism of Self/Up CFTs. Although there is some ambiguity with regard to the ATC variable—it may be correlated with unobserved factors such as proximity to airports—the findings are intriguing because they allow us to dismiss an interpretation in which all audiences have the same general effect. Schlenker (1986) has argued that accountability to a non-threatening but challenging audience can spur complex analysis, and this may describe the impact of ATC observation on pilots.

While no field data have perfect internal validity, we can rule out the most obvious alternative explanations for Study 1's results, namely, an effect of accident severity or an effect merely of having an audience. Other possible confounds still need to be examined and may be of interest in future studies of learning in an organizational setting. First, in an organizational setting there is greater interdependence between actors. Counterfactual statements, if they are to be the foundation for drawing lessons, may not work in these settings because they fail to express interaction effects between changes to different links in the causal chain. Second, there may be motivational differences between private pilots, who fly for recreation, and organizational pilots, who fly for employment. Study 2 was designed to establish the internal validity of our causal claims in a laboratory experiment in which we randomly assigned participants to conditions. In the experiment, the alternative explanations discussed above are ruled out because all subjects experience the same simulated flight task.

STUDY 2: DANGEROUS INCIDENTS ON A FLIGHT SIMULATOR

Conducting experiments on pilots in their natural environment would be difficult if not impossible. There are simply too many variables beyond control. Further, the dangerous experiences required as stimuli would be unlikely to get ethics board approval, even if consenting pilots could be found. Fortunately, the existence of flight simulator technology allowed us to create a realistic approximation of the experience of danger in a laboratory experiment. We constructed a task in which college students were asked to land their plane on a small runway in bad weather. The chief advantage of this method, of course, is internal validity. We randomly assigned participants to the presence or absence of organizational accountability, and all other variables were held constant. In the accountable condition, participants were put in the role of an organization member, an operationalization used in past experimental research on the effects of organizational contexts (see Zucker, 1977; Fox and Staw, 1979). Our dependent measures were taken from a pilot log modeled on the ASRS reporting form so that these variables could be constructed in the same way as in Study 1.

Method

Participants. Participants were 42 Stanford students who were paid for their participation. They ranged in age from 18 to approximately 30, and 74 percent were female. We screened for students without previous experience using flight simulators so that the task would present the same degree of difficulty to all participants. A previous experiment with a flight simulator task found that college students and experienced pilots are prone to the same patterns of errors (Skitka, Mosier, and Burdick, 1996).

Materials. We used Microsoft's Flight Simulator 5.0 and Microsoft's Sidewinder joystick. Running on a PC with attached speakers, this combination provided excellent visual cues and feedback, and participants easily learned it.

Procedure. We conducted the experiment in three phases: training, trial, and pilot log completion. In the training period, the experimenter taught participants how to operate the controls and encouraged them to get a feel for the plane's responsiveness by taking a practice flight. This lasted between 5 and 10 minutes. It ended when the participant indicated comfort flying the plane and was ready to move on to the real flight.

The trial required that participants finish a short flight that culminated in a landing at Meigs Field, Chicago. This task was designed to make it virtually certain that the novice pilot would experience a rough landing. The program realistically simulated the geography of the area and the weather conditions. In the instructions before the trial flight, the experimenter warned participants about environmental conditions (which were modeled in the simulation) that could affect their performance, such as airport characteristics, fog, wind, and so forth. Specifically, participants were told the following:

> Meigs Field is on an island with a short runway, so it is crucial to gracefully touch down at the beginning of the runway and to quickly but smoothly come to a complete stop. Meigs Field has a large concentration of private pilot and charter aircraft which can create traffic problems. The visibility is poor: only one mile due to fog. But, it is daytime so there is plenty of light. There are light crosswinds blowing off Lake Michigan which may account for some of the plane's erratic movements.

The purpose of these instructions about external conditions affecting the flight was to make it plausible to participants that they could "cognitively undo" a rough landing by imagining changes of external factors. To mirror the ambiguity of the actual dangerous incidents that pilots experience, there needed to be external as well as internal factors that seemed to be plausible causes of the outcome.

Participants were asked to imagine themselves as real pilots and were assigned one of two roles, private pilot or organizational pilot. In both cases, participants were told that they were experienced pilots flying a small plane. Private pilots were flying for recreational purposes, and organizational pilots were flying in the context of work. Organizational pilots worked for a company that taxied executives on commuter flights. This company was very concerned about smooth landings because of a desire for repeat business, and it evaluated its pilots on this basis. Pilots taking this role should feel the pressure of accountability to a threatening audience with known views and hence should interpret the events of their flight with a defensive-bolstering style rather than a complex self-critical style.

The trial flight involved guiding the plane from a set position above a small runway to the ground, a challenging task. As one might expect, not all participants landed successfully on their first try. In the event of a crash, the simulation was reset and the participant repeated the landing sequence until successful (14 of 42 participants required a second attempt, and 2 required a third). Since the experience of a crash might affect a participant's subjective experience of the successful landing, the number of crashes prior to a successful landing was included in our analyses as a second covariate. Another measure of the trial flight was the time elapsed from the start to their coming to a full stop. Neither the number of crashes nor total flight time varied by condition, but they were included as covariates to control for any extraneous variance owing to differences in the severity of problems encountered in the practice and trial flight. After recording the flight time, the experimenter gave moderately positive feedback to participants so as to reduce any tendency for the experimenter to be seen as a threatening, critical audience.

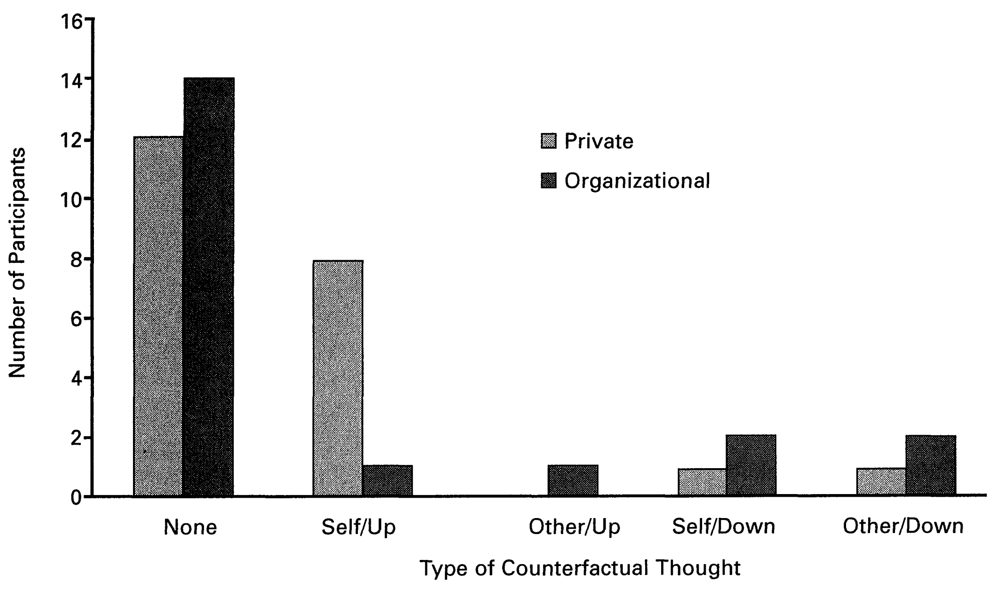
After landing, participants were told to fill out their flight log. The log format was based on the ASRS reporting form. It initially requested that pilots describe and diagnose any dangerous incidents in their flight. Included for reference were instructions suggesting a list of possible contributing factors, some human, some environmental—the same list of factors that appear on the ASRS reporting form. The last section asked subjects to list any lessons or recommendations about how to improve performance in the future: "Write any lessons that you learned. Include any of the following: future actions you would or would not take, rules for other pilots or airport traffic, or lessons about airports." Upon completion of this section, participants were debriefed and paid.

Dependent measures. The first dependent variable was the presence of spontaneous counterfactual comparisons (CFTs) in participants' open-ended description of their flight. A judge who was naive to the experimental conditions and hypotheses scored counterfactual comparisons for direction (upward versus downward) and subject (self versus other). The other dependent variables were measures of learning. As mentioned above, the last section of the report explicitly asked participants to write any lessons they learned. By scoring this section, we were interested in variation in the quality, quantity, and target of the learning statements. To be sure that evaluation of the learning statements was completely unbiased by the experimental conditions or other cues in the pilots' reports, the lesson section of each report was transcribed and put on a form for scoring. Two independent judges rated the lesson statements on four criteria: number of lessons, number of personal lessons, specificity of personal lessons, and an overall rating of personal lessons (see the Appendix for coding instructions). The judges were very reliable, α = .88, .96, .81, and .90, on the four criteria, respectively, so judges' scores were averaged.

Results

At the time participants wrote their flight description, there had been no mention at all of counterfactual comparisons. Nonetheless, spontaneous references to counterfactual outcomes occurred frequently. Of the 42 participants, 16 made references to counterfactual outcomes. The overall rate of CFTs did not significantly vary across the conditions. Private pilots had a 45 percent CFT rate, compared with 30 percent for organizational pilots. Consistent with the first hypothesis, however, the frequency of particular kinds of CFTs varied across conditions. Figure 1 shows the distribution of response types by pilot condition, which illustrates that private pilots were the primary contributors of Self/Up CFTs—eight of the nine Self/Up CFTs were made by private pilots. Looking at the proportion of Self/Up CFTs to all CFTs within each pilot group, 80 percent of the counterfactual statements made by private pilots were Self/Up CFTs, and 17 percent of the counterfactual statements made by organizational pilots were Self/Up CFTs. Overall, the distribution of pilot narratives

Figure 1. Distribution of participants' responses in the pilot log as a function of whether organizational accountability was absent (private pilots) or present (organizational pilots).



with no CFTs, Self/Up CFTs, and Other CFTs is significantly different between commercial and private pilots [χ^2 (1, N = 42) = 6.80, $p < .05$].

Our analyses of organizational accountability, the two controls, and counterfactual thinking on the overall rating of personal learning are summarized in table 4. We report the dependent measure of overall rating of learning, yet the results are substantively the same using both specificity of personal lessons and the number of personal lessons. Organizational accountability had a large impact: private pilots had an average overall learning rating of 2.94, while organizational pilots averaged only 1.75 [$F(1,40) = 6.45$, $p <. 05$].

In support of the first hypothesis, one particular kind of counterfactual comparison enhanced the drawing of lessons. People who generated Self/Up CFTs had an average learning rating of 3.88, much higher than people who did not [$F(2,40) = 6.42$, $p <. 01$]. People who did not make any counterfactual comparisons had a rating of 2.02, and people who made other kinds of counterfactual comparisons had a rating of 1.80.

To test the claim in hypothesis 2 that the counterfactual thinking mechanism mediates the effect of accountability, we included the independent variable of organizational accountability and the presence of Self/Up CFTs in the same model predicting learning. The effect of organizational accountability was reduced to a nonsignificant level, and only Self/Up CFT remained significant [$F(1,40) = 2.10$, n.s., and $F(1,40) = 7.89$, $p < .01$, respectively]. This indicates that the effect of the manipulation of organizational accountability on lesson drawing was fully mediated by counterfactual thinking. Results using the number and specificity of personal lessons as dependent measures are substantively the same, including the mediation relationship.

Discussion

Study 2's results supported both our hypotheses. The experimental method greatly increases evidence for the internal

Table 4

Regressions of Independent and Mediating Variables on Learning (Overall Rating) in Study 2*

	1	2	3
Independent variable			
Organizational accountability	–1.194•		–.680
(private pilot = 0; commercial pilot = 1)	(.470)		(.469)
Covariates			
Number of training crashes	–.067	.068	.023
	(.410)	(.382)	(.378)
Time to complete landing	–.047	–.081	–.068
	(.078)	(.073)	(.072)
Mediating variable			
Self/Up CFTs		1.903•••	1.584••
		(.527)	(.564)
Adjusted R^2	.094	.213	.236
F	2.39	4.62••	4.09••
D.f.	(3,40)	(3,40)	(4,40)

•$p < .05$; ••$p < .01$; •••$p < .001$.
* Standard errors are in parentheses.

validity of these claims in that all variables other than the presence versus absence of the pilot's embeddedness within an organizational hierarchy are controlled. Another way that Study 2 increases the evidence for our model is that variations of the learning measure all produced the same pattern of results, supporting the mediation relationship.

Two critiques might still be levied against our claims of learning. A first issue is whether the learning produced as a consequence of Self/Up CFTs is manifest merely in the form of stated plans for better performance or if it is also manifest in actual performance gains. A second, related issue is whether the link between this type of CFT and learning holds up when they are not both assessed with verbal measures. Fortunately, the experimental flight simulator method developed for Study 2 allowed us to check these concerns.

Study 2a: Performance Improvement in Repeated Trials

In a second flight simulator experiment, we further tested the claim that people learn from Self/Up CFTs by using a behavioral measure of learning instead of a cognitive, lesson statement measure: we measured actual performance improvement on a future trial after manipulating counterfactual comparison generation. The ancillary experiment was run with the same basic procedure except that the 21 participants had multiple trials at landing the plane. Participants were randomly assigned to one of two conditions that differed only in the kind of counterfactual comparison they were asked to make between the two attempted landings. In one condition, participants were asked to generate a self-focused upward counterfactual thought about their landing, and in the other condition, they were asked to generated a self-focused downward comparison. We did not prompt them to make counterfactual comparisons about "other" antecedents, up or down, because "other" antecedents would not change between the two trials, nor would such antecedents be obviously consequential in the controlled experimental setting. After the manipulation of counterfactual thinking, participants completed a second trial of the landing task. Both landings by participants were recorded and edited onto a videotape.

The tape of participants' landings was shown to two judges (in separate sessions) who were naive to the hypotheses and to the experimental manipulation. The judges rated the quality of each landing on an absolute scale (from 0 = crash, to 5 = perfect landing), and their ratings were highly reliable ($\alpha = .90$). We combined the coders' ratings into a composite measure for our analyses. The difference between the average rating of the second landing and that of the first was our measure of improvement.[4] We also had the judges categorize each set of pilot trials as having demonstrated performance improvement or decline. The two judges agreed on this measure for 91 percent of the participants and resolved discrepancies through later discussion.

Consistent with our expectations, participants who made self-focused upward counterfactual comparisons showed more improvement from the first to the second trial. They had ratings 1.22 points higher in the second trial than pilots who were asked to make Self/Down CFTs (one-tailed

4
This measure yields the same results as a repeated measure ANOVA with trial being the within-subject factor.

t = 1.76, $p < .05$), and they also showed greater improvement from the first to the second trial (difference score of 1.55 versus –.36, one-tailed t = 1.87, $p < .05$). The qualitative measure of performance change shows the same pattern. Self/Up CFTs were more likely than Self/Down CFTs to engender improvement from the first trial to the second [χ^2 (1, N = 21) = 5.74, $p < .02$]. Admittedly, this study focused on a narrow set of conditions (the effect of self-focused counterfactual thoughts on non-accountable pilots) and involved only a small number of participants, and hence the evidence must be regarded as preliminary. Nevertheless, these findings allay the concern that the learning resulting from Self/Up CFTs is mere talk rather than action.

GENERAL DISCUSSION

Taken together, the results of the current studies support our argument that individuals learn from imagined experiences and that certain organizational contexts inhibit this process. The consistency in results is all the more compelling in light of the diversity of methods and procedures used. In support of our first hypothesis, both experienced pilots reflecting on naturally occurring near accidents (Study 1) and college students reflecting on flight simulator incidents (Study 2) were more likely to draw lessons for improved performance when they had initially responded to their experience with a self-focused, upward counterfactual thoughts. An ancillary study using a flight simulator suggested that this learning from imagined counterfactual scenarios occurred not only at a cognitive level (lesson production) but also at a behavioral level (performance gains). In support of our second hypothesis, in both field and experimental studies, pilots flying under the condition of accountability to organizational superiors were less likely to generate self-focused, upward CFTs and, consequently, were less likely to learn from the experience. This convergence of results across studies using different, complementary methods attests to the validity of the findings.

Theoretical Implications

Counterfactual thinking. The current research contributes to basic psychological theory about the function of counterfactual thinking by clarifying that learning depends not only on the direction of a thought (upward vs. downward) but also its subject (self vs. other). The current studies also redress several methodological shortfalls that have been noted in reviews of the CFT literature (Roese and Olson, 1995a) in that, first, CFTs were measured in a natural, real-world context rather than solely in a laboratory, and, second, the mediating role of CFTs was tested rather than merely inferred. Furthermore, by demonstrating that CFTs are shaped by an individual's position of accountability, the current research challenges the dominant conception of counterfactual thinking as a purely intra-psychic phenomenon. In other words, even a process as personal and private as imagination about what might have been is importantly constrained by the social contexts in which the person is embedded. In sum, although the primary contribution of the current studies is applying psychological theory to organizational questions in novel ways, the current studies also contribute to the basic psychological literature by refining constructs, advancing methods, and challenging theoretical assumptions.

Learning from experience. In organizational behavior, one contribution of the current studies is to the literature on individual learning or sensemaking in organizations (Weick, 1995). The issue here is not how organizations "learn," in the sense of updating their policies; rather, the issue is how individuals learn, in the sense of updating their knowledge. Our research offers a novel model of individual learning from self-focused upward CFTs that differs from extant models.

Organizational research on individual learning has primarily drawn on classic psychological models of learning, beginning with behaviorist models of learning from *repeated* experience through reinforcement, which describe many kinds of learning in organizations quite adequately. For example, assembly-line workers in a piece-rate compensation system may adapt their work strategies over time in response to reinforcement (Luthans and Kreitner, 1985). Alternatively, workers might learn from multiple experiences through more cognitively mediated processes, ones in which they do not have to try different work strategies directly, learning vicariously from observing other workers' experiences (Wood and Bandura, 1989) or observing many other workers and inducing causal rules from the covariation of work strategies and compensation (Cheng and Novick, 1992; March, 1994).

Models of learning from repeated experience, however, fail to explain other types of learning in organizations. Often, employees and managers need to learn without undergoing repeated experiences of the event. For example, if the event is hazardous, a person who relies on trial-and-error reinforcement to learn about aviation accidents, for instance, would be unlikely to survive very long. Learning vicariously from observing peers or studying historical cases is also precluded in many cases, such as when the event is highly rare or when it is kept private by those who experience it. Another barrier is when events hinge on historically novel conditions. Perrow (1984) argued that the complexity and rate of change of technology has made it impossible to collect enough data to sort out the many possible variables and interactions that may cause accidents in complex technological systems, such as aviation and nuclear power. Yet despite the impossibility of learning from repeated experience, there is a remarkably good safety record in some of these systems (Roberts, 1990). This suggests that individuals may learn about accidents (and other organizational outcomes) through processes that do not depend on repeated experience, as in classical learning models.

Models of learning from a *single* experience of an outcome have been adapted from attribution theory. In these models, a single experience of an accident spurs the individual to induce a cause-effect rule, which, in turn, enables the individual to predict recurrences and to take preventive measures (e.g., DeJoy, 1985). Attributional models explain how individuals learn from a single, salient experience, yet they do not explain how individuals learn about events that they have not experienced, such as in the case of learning about accidents from the experience of a near miss. Attribution models predict a "lack of attributional search" and lesson drawing after near misses (DeJoy, 1990: 115).

The current research contributes a model of learning from *imagined* experiences, from counterfactual thoughts. While counterfactual thinking is like attribution in that it involves causal inference, it differs in focus. Whereas attribution is an inference about how one's actual, experienced outcome was caused, CFT is an inference about how a counterfactual outcome could have been caused. The distinctness of these psychological processes is illustrated by a recent study that interviewed individuals in various work settings about their post-outcome sensemaking (Cannon, 1999): of all the respondents interviewed, the one showing the most external pattern of attribution (an aviation pilot who attributed a minor accident to unusual wind patterns) also showed the strongest pattern of self-focused upward CFTs. This illustrates a case in which a performance-promoting lesson is drawn not from one's attribution about the actual accident but from one's imaginative simulations of how it could have been averted. While this case happens to come from the aviation setting, Cannon also observed patterns of self-focused upward CFT among respondents in a variety of industry and career settings. While learning was not measured, the ubiquity of CFTs in reactions to workplace outcomes suggests that they may play a pervasive role in individuals' process of learning to optimize their work performance.

The price of hierarchy. A second area of organizational behavior research to which the current studies contribute is the literature on the dysfunctional side-effects of hierarchical accountability. Hierarchical accountability is a basic means of ensuring coordination among organizational units, yet many field studies of organizations have suggested that hierarchical structures are associated with lower levels of innovation and learning (see Tushman and O'Reilly, 1997). Why does hierarchical accountability have this chilling effect? Past research has highlighted that accountability spurs employees to deflect blame by distorting and filtering the information they communicate to others in the organization (O'Reilly and Roberts, 1974; O'Reilly, 1978). There are several notable features of this mechanism for the chilling effect, namely, that it involves intentional, self-interested actions by individuals, which impede learning at the organizational level. The current research identifies a different way hierarchical accountability squelches learning—by inhibiting self-focused upward CFTs. This consequence of accountability operates below the level of conscious awareness (Lerner and Tetlock, 1999). Hence, this mechanism differs from the previous one in that it involves unintentional cognitive strategies, which impede learning at the individual level. It may be even more harmful to organizations than the first barrier to learning because all employees are susceptible, not solely those willing to engage in intentional spin control.

Limitations

Despite the contributions of the current research, several shortcomings are visible in hindsight; to develop suggestions for future research, it is useful to consider what we could have done better. One limitation is our operationalization of learning. There are two primary approaches to learning: a cognitive conceptualization, in which learning is measured in terms of knowledge production, and a behavioral conceptualization, in which learning is measured in terms of performance gains. Our two main studies take the cognitive approach, measuring learning in terms of the production of performance-promoting lessons. We provide only preliminary evidence (from Study 2a) that these lessons translate into improved performance. It would be useful to complement the current studies with evidence that self-focused upward CFTs lead to performance gains. The most compelling evidence would be measures of real-world success, for example, a study showing that student pilots who engage in more self-focused upward CFTs later reach higher levels of proficiency as pilots.

Another limitation of our empirical tests is that our evidence comes solely from the aviation setting, and evidence from other industry settings would be helpful for testing the generalizability of our hypothesis about CFTs and learning. One setting that has been studied experimentally is managerial responses to factory accidents, where we have found evidence for the converse hypothesis that other-directed upward CFTs do not foster adaptive lessons (Morris, Moore, and Sim, 1999). As for evidence from other industry settings that self-focused upward CFTs engender learning, we know of only correlational and anecdotal evidence. Baron (1999) found that successful entrepreneurs differ from novice entrepreneurs and non-entrepreneurs in that they find it easier to admit having made mistakes. Although other interpretations are possible, this correlational result may reflect that entrepreneurs are more likely to process experiences by envisioning how they might have acted otherwise to obtain greater success and that these visions enable future success. More anecdotally, another source of evidence for the facilitatory role of self-focused, upward CFTs is the success of the U.S. Army's After Action Review (AAR) program (Davenport and Prusak, 1998), in which an external observer prompts soldiers at the end of the day to review their critical actions and generate thoughts about what could have been done differently to obtain a better outcome—self-focused upward CFTs. The army's goal is to elicit "grounded lessons," insights from the perspective of the actor on the ground about feasible, incremental changes that are efficacious without requiring other simultaneous changes to other factors in the complex system. The strategies that emerge from this change process are concrete, implementable grassroots lessons, unlike the more abstracted, idealized strategies promulgated in the top-down planning of the Vietnam era. AAR procedures are increasingly advocated as useful tools in a variety of business settings where improved performance is sought (Darling, 2000). Although the aforementioned evidence about entrepreneurs and soldiers is suggestive, the causal efficacy of CFTs as a means of learning should be investigated in these settings using methods like those used in the current research.

Another boundary condition that calls out for investigation is the nature of the performance task. For what kinds of tasks will lessons spawned by self-focused upward CFTs translate into performance gains? One clear answer is that performance gains should accrue only in domains in which people's counterfactual search isolates actions that do have a causal impact on performance. The lessons drawn from self-directed upward CFTs may not translate into efficacious actions in domains in which outcomes are determined by chance, such as gambling (Gilovich and Douglas, 1986), or by complex, external factors, such as political forecasting (Tetlock, 1998). In these domains, self-focused upward CFTs may result in superstitious learning rather than adaptive learning.

In addition to these empirical limitations, there are also some conceptual limitations to the current research. These point to directions in which the reasoning could be elaborated or extended. Lessons for individual performance improvement is not the only kind of learning, and different hypotheses will be required for different kinds of learning. A key point in our argument is that the special status of self-focused upward CFTs is a matter of individual psychology rather than of logic. The advantage of this type of CFT is that it takes very little cognitive processing to get from the CFT to a lesson. Only one cognitive step is required to go from a proposition about how one could have produced a better outcome to an actionable plan for promoting better future outcomes; downward CFTs and other-focused CFTs do not facilitate lesson drawing in the same way. Yet from a purely logical standpoint, CFTs of any subject or direction have equal relevance to predictions about the future. Downward comparisons reveal which factors in one's own actions or the environment should *not* be changed, and externally focused upward comparisons reveal which factors in the environment make a positive difference. Because of their logical relevance, these other types of counterfactual thinking may play important roles in learning processes that are freed from the processing constraints of the individual human mind. Many types of counterfactual thoughts may play a useful role in learning by artificial intelligence programs (e.g., Costello and McCarthy, 1999). Also, many types of counterfactual thinking and simulation may play a useful role in organization-level learning (March, Sproull, and Tamuz, 1991). Organization-level learning in a pure form involves an agent of the system (a manager or team) who can evaluate the whole system's performance and make appropriate changes to policies and procedures. As March, Sproull, and Tamuz (1991) have described, this involves forms of simulation such as computer modeling, which is free from the constraints limiting mental simulation with imagined counterfactual scenarios.

Questions for future research about the relation between individual- and organization-level learning are raised by consideration of the fact that our Study 1 data came from the Aviation Safety Reporting System, an institution designed and used for organization-level learning about reducing aviation accidents. On the one hand, given that the ASRS creates a context that spurs individual-level learning, should organization-level learning systems be designed to prompt individual-level learning? On the other hand, might this use of the ASRS as an arena for self-evaluation foil its original purpose? The FAA hoped to gain insight about problematic aspects of airports and the air traffic control system (i.e., factors under FAA control). To the extent that pilots center their narrative reports on self-focused CFTs, the content of the ASRS may be directed away from the factors that can be effectively controlled at the organizational level. Detailed studies of how the FAA has used the data base in revising aviation policies might elucidate whether learning processes at the two levels are synergistic or antagonistic.

Practical Implications for Knowledge Management Systems

The current findings raise important issues for organizations seeking to maximize processes of adaptive learning at the individual and organizational levels. The outlook for individual learning is grim if learning cannot take place in the presence of organizational accountability—aviation organizations will always need to place pilots in positions of accountability to others. Moreover, there are several ways that systems promoting organization-level learning can create particularly threatening accountability: employees are interrogated in review procedures; they are asked to keep performance logs; and their behavior is monitored with technology, such as cockpit flight recorders in airplanes, cameras in factories, computer systems in offices, and so forth. All this information is useful to the organization in reconstructing the causes of accidents or other negative outcomes and in monitoring the effects of policy changes, but these trappings of a learning organization can come across as an Orwellian regime of surveillance. This induces a mindset of defensive bolstering, which reduces the likelihood of an employee learning. Hence, a dilemma arises between the competing ends of organization-level learning and individual-level learning.

The current research, in combination with a look to evolving practices, suggests that there are several ways to resolve this dilemma. One approach is creating a culture in which accountability to organizational superordinates feels like an opportunity to contribute rather than a threatening event. This can be done by embedding it in a more general program and emphasis on improved safety and effectiveness. The U.S. Army's AAR program is linked to an organizational-learning program, the Army's Center for Army Lessons Learned (CALL), which works to share the grounded lessons learned in one training exercise or mission with soldiers elsewhere facing a similar task. For example, lessons learned in peacekeeping missions in Somalia and Rwanda in the early 1990s in tasks such as searching houses for weapons led to improved performance of these tasks in the 1994 Haitian mission. Research by Hofmann and Stetzer (1998) has documented that all kinds of learning from accidents are fostered by a climate of emphasis on organizational safety (e.g., availability of safety equipment, managerial safety instruction, and safety incentives).

A second approach to resolving the dilemma is creating accountability to an outside party rather than to an employee's superordinates. This accountability is nonthreatening because the outside parties lack direct control over the employees. As in the case of the ASRS, the outside party can convey summary information to regulators and others who make policy in the industry, as well as to the organizations involved. Other industries are developing similar institutions, such as the Center for Quality Improvement and Patient Safety, which aims to reduce hospital errors by half over the next five years (*American Medical News,* 1999).

A third approach, which fits naturally in industrywide approaches, is ensuring anonymity to employees who report near accidents. This is an important element in the ASRS, and it is a very important element in the medical setting. The American Medical Association has said that to reduce hospital errors that result in injury or death of patients, the data reporting and collection process must have the confidence of doctors and nurses. In part, this means the system needs to be a "non-punitive, evidence-based error reporting system that provides strong legal protections [e.g., confidentiality and protection from discovery in legal proceedings] for participants in safety programs" (CNN, 2000). Overall, it may be that some of the same features of knowledge management programs that foster organization-level learning are helpful to individual-level learning in that they reduce the pressure of threatening accountability.

REFERENCES

American Medical News
1999 "Learn from mistakes." February 6: 1.

Anderson, P. A.
1981 "Justifications and precedents as constraints in foreign policy decision-making." American Journal of Political Science, 738–761.

Antonioni, D.
1994 "The effects of feedback accountability on upward appraisal ratings." Personnel Psychology, 47: 349–356.

Baron, R. A.
1999 "Counterfactual thinking and venture formation: The potential effects of thinking about what might have been." Journal of Business Venturing, 15: 79–91.

Baron, R. M., and D. A. Kenny
1986 "The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations." Journal of Personality and Social Psychology, 51: 1173–1182.

Cannon, D. R.
1999 "Cause or control?" The temporal dimension in failure sense-making." Applied Behavioral Science, 35: 416–438.

Carroll, J.
1998 "Organizational learning activities in high-hazard industries: The logics underlying self-analysis." Journal of Management Studies, 35: 699–717.

Cheng, P. W., and L. R. Novick
1992 "Covariation in natural causal induction." Psychological Review, 99: 365–382.

Costello, T., and J. McCarthy
1999 "Useful counterfactuals." Linkoping University Electronic Press, 3. http://www.ep.liu.se/ea/cis.

Crossan, M. M., H. W. Lane, and R. E. White
1999 "An organizational learning framework: From intuition to institution." Academy of Management Review, 24: 522–537.

CNN
2000 "AMA votes to support tracking of medical errors." http://www.cnn.com/2000/HEALTH/06/15/ama.errors/index.html, 6/15/2000.

Darling, M.
2000 "Army tactics for business: Answering an After Action Review may solve company problems." http://www.cnn.com/2000/cnnfn/news, 3/02/2000.

Davenport, T. H., and L. Prusak
1998 Working Knowledge: How Organizations Manage What They Know. Boston: Harvard Business School Press.

DeJoy, D. M.
1985 "Attributional processes and hazard control management in industry." Journal of Safety Research, 16: 61–71.
1990 "Spontaneous attributional thinking following near-miss and loss producing traffic accidents." Journal of Safety Research, 21: 115–124.

Festinger, L.
1954 "A theory of social comparison processes." Human Relations, 7: 117–140.

Fox, F. V., and B. M. Staw
1979 "The trapped administrator: The effects of job insecurity and policy resistance upon commitment to a course of action." Administrative Science Quarterly, 24: 449–471.

Gilovich, T., and C. Douglas
1986 "Biased evaluations of randomly determined gambling outcomes." Journal of Experimental Social Psychology, 22: 228–241.

Hendee, W. R.
1995 "Accountability in the acquisition and use of medical technologies." International Journal of Technology Management, S (II): 38–47.

Hofmann, D., and A. Stetzer
1998 "The role of safety climate and communication in accident interpretation: Implications for learning from negative events." Academy of Management Journal, 41: 644–657.

Kahneman, D., and A. Tversky
1982 "The simulation heuristic." In D. Kahneman, P. Slovic, and A. Tversky (eds.), Judgment under Uncertainty: Heuristics and Biases: 201–208. New York: Cambridge University Press.

Lerner, J., and P. Tetlock
1999 "Accounting for the effects of accountability." Psychological Bulletin, 125: 255–275.

Liberman, N., L. C. Idson, C. J. Camacho, and E. T. Higgins
1999 "Promotion and prevention choices between stability and change." Journal of Personality and Social Psychology, 77: 1135–1145.

Luthans, F., and R. Kreitner
1985 Organizational Behavior Modification and Beyond. Glenview, IL: Scott, Foresman.

March, J.
1994 A Primer on Decision Making: How Decisions Happen. New York: Maxwell Macmillan International.

March, J. G., L. S. Sproull, and M. Tamuz
1991 "Learning from samples of one or fewer." Organization Science, 2: 1–13.

Medvec, V. H., S. F. Madey, and T. Gilovich
1995 "When less is more: Counterfactual thinking and satisfaction among Olympic medallists." Journal of Personality and Social Psychology, 69: 603–610.

Miller, D. T., W. Turnbull, and C. McFarland
1990 "Counterfactual thinking and social perception: Thinking about what might have been." In M. P. Zanna (ed.), Advances in Experimental Social Psychology, 23: 305–331. New York: Academic Press.

Morris, M. W., P. C. Moore, and D. L. Sim
1999 "Choosing remedies after accidents: Counterfactual thoughts and the focus on fixing human error." Psychonomic Bulletin and Review, 6: 579–585.

O'Reilly, C. A.
1978 "The intentional distortion of information in organizational communication: A laboratory and field investigation." Human Relations, 31: 173–193.

O'Reilly, C. A., and K. H. Roberts
1974 "Information filtration in organizations: Three experiments." Organizational Behavior and Human Decision Processes, 11: 253–265.

Peecher, M. E., and D. N. Kleinmuntz
1991 "Experimental evidence on the effects of accountability on auditor judgments." Auditing, 10: 108–113.

Perrow, C.
1984 Normal Accidents. New York: Basic Books.

Roberts, K.
1990 "Some characteristics of one type of high reliability organization." Organization Science, 1: 160–176.

Roese, N. J.
1994 "The functional basis of counterfactual thinking." Journal of Personality and Social Psychology, 66: 805–818.
1997 "Counterfactual thinking." Psychological Bulletin, 121: 133–148.

Roese, N. J., and J. M. Olson
1995a "Counterfactual thinking: A critical overview." In N. J. Roese and J. M. Olson (eds.), What Might Have Been: The Social Psychology of Counterfactual Thinking: 1–55. Mahwah, NJ: Lawrence Erlbaum.
1995b "Functions of counterfactual thinking." In N. J. Roese and J. M. Olson (eds.), What Might Have Been: The Social Psychology of Counterfactual Thinking: 169–197. Mahwah, NJ: Lawrence Erlbaum.
1997 "Counterfactual thinking: The intersection of affect and function." In M. P. Zanna (ed.), Advances in Experimental Social Psychology, 29: 1–59. New York: Academic Press.

Schlenker, B.
1986 "Personal accountability: Challenges and impediments in the quest for excellence." Contract No. DAAG29-81-D-0100, Navy Personnel Research and Development Center.

Skitka, L. J., K. Mosier, and M. D. Burdick
1996 "Accountability and automation bias." Unpublished manuscript, Department of Psychology, University of Illinois at Chicago.

Tetlock, P. E.
1983 "Accountability and complexity of thought." Journal of Personality and Social Psychology, 45: 74–83.
1992 "The impact of accountability on judgment and choice: Toward a social contingency model." Advances in Experimental Social Psychology, 25: 331–376.
1998 "Close-call counterfactuals and belief-system defenses: I was not almost wrong but I was almost right." Journal of Personality and Social Psychology, 75: 639–652.

Tetlock, P. E., L. Skitka, and R. Boettger
1989 "Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering." Journal of Personality and Social Psychology, 57: 632–640.

Tushman, M., and C. A. O'Reilly
1997 Working Through Innovation: A Practical Guide to Leading Organizational Change and Renewal. Cambridge, MA: Harvard University Press.

Weick, K. E.
1995 Sensemaking in Organizations. Thousand Oaks, CA: Sage.

Williams, C.W., P. R. Lees-Haley, and J. R. Price
1996 "The role of counterfactual thinking and causal attribution in accident-related judgments." Journal of Applied Social Psychology, 26: 2100–2112.

Wood, R., and A. Bandura
1989 "Social cognitive theory of organizational management." Academy of Management Review, 14: 361–384.

Zucker, L. G.
1977 "The role of institutionalization in cultural persistence. " American Sociological Review, 42: 726–743.

APPENDIX: Coding Instructions for Dependent Learning Measures in Study 2

Number of personal lessons: Number of statements written as a result of the flight that indicate the pilot will act or would like to have acted in a particular way, either the same or different from his/her behavior in actual event.

Number of lessons: Number of statements written as a result of the flight that indicate the pilot would like or would have liked his/her actions, the actions of others, or conditions in the environment (e.g., the airplane, airport, and weather) to be a particular way, either the same or different from the actual event.

Specificity of personal lessons: On the following 4-point scale, indicate the overall specificity (i.e., level of detail) of the personal lessons written in the statement: 0 = None—Statement has no personal lessons; 1 = Minimal detail—Less detail and lesson is gone; 2 = Moderate detail; or 3 = Extensive detail—Lesson is very complete.

Overall rating of personal lessons: On the following 8-point scale, indicate the amount of personal learning in the statement. [On the learning scale, 0 = none; 1 = little; 4 = moderate; and 7 = extensive].