Learning from Complexity: Effects of Prior Accidents and Incidents on Airlines' Learning

Pamela R. Haunschild
University of Texas at Austin
Bilian Ni Sullivan
Stanford University

Using data on accidents and incidents experienced by U.S. commercial airlines from 1983 to 1997, we investigated variation in firm learning by examining whether firms learn more from errors with heterogeneous or homogeneous causes. We measured learning by a reduction in airline accident and incident rates, while controlling for other factors related to accidents and incidents. Our results show that heterogeneity is generally better for learning, as prior heterogeneity in the causes of errors decreases subsequent accident rates, producing a deeper, broader search for causality than simple explanations like "blame the pilot." The benefits of heterogeneity, however, apply mainly to specialist airlines. Generalist airlines learn, instead, from outside factors such as the experience of others and general improvements in technology. These results suggest a theory of learning across organizational forms: complex forms benefit from simple information, and simple forms benefit from complex information. The implications of our study for learning theories and work on organizational errors are discussed.●

> *For all the scientific pizzazz [involved in airline accident investigations], unraveling the subtle, complex chain of events leading to aviation deaths is proving more elusive than ever.*
>
> —"Why more plane-crash probes end in doubt," *Wall Street Journal*, March 22, 1999

Organizations like airlines try to learn from experience, understanding what went wrong so that it won't go wrong next time. But if, as the quote above suggests, the causes are often left in doubt, such learning is likely to be difficult. Learning is also likely to vary across firms, despite industry regulation that should affect all airlines equally. Investigators of the 2000 Air France Concorde crash discovered that British Airways had recommended changes to the Concorde's water deflector in 1995 but that Air France had not made those changes (Phillips, 2000). As Donoghue (1998: 36) explained, ". . . any safety initiative has an unequal effect on the carriers and becomes an issue to be promoted or fought . . . seeking the path that best suits [the airline] individually." Other heavily regulated industries, such as nuclear power, also show substantial variance in incident rates among firms (Morris and Engelken, 1973), which indicates that firms vary in how effectively they learn from their experience. Despite much discussion and analysis of aviation errors (airline accidents and incidents), there has been little work investigating the role of organizational learning and none examining variation in learning across firms in the industry.

Learning from experience has been shown to have important effects on such varied outcomes as manufacturing plant productivity (e.g., Argote, Beckman, and Epple, 1990), service timeliness (Argote and Darr, 2000), and hotel survival (Baum and Ingram, 1998). If firms learn from experience, then the attributes of this experience are likely to affect the rate and effectiveness of learning. Some firms have heterogeneous experience in that their accidents and incidents ("errors") are caused by a large number of different factors, which are likely to interact in complex ways. Some firms have more homogeneous experience, with errors caused by a small number

0001-8392/02/4704-0609/$3.00.

●

We would like to thank Ozgecan Kocak for help with data collection and comments on earlier versions of this paper. Thank you also to Glenn Carroll, Martin Evans, James March, Daniel Stewart, Ezra Zuckerman, and seminar participants at University of Texas at Austin, University of Toronto, University of Washington, University of Wisconsin, University of California at Irvine, University of Bologna, and the 2000 Conference on Knowledge and Innovation. This paper was also significantly helped by the especially helpful and detailed comments of the *ASQ* reviewers and Christine Oliver, as well as by Linda Johanson's editing.

of similar factors. It is likely that the complexity of prior experiences, as well as characteristics of the firms themselves, affect how well airlines can learn from that experience. We investigate these issues in the context of airline accidents and incidents to explain variation in learning among firms in the airline industry.

According to the NTSB (2001) Code of Federal Regulations (49CFR830.2, p. 1195), an accident "means an occurrence . . . in which any person suffers death or serious injury, or in which the aircraft received substantial damage." An incident is "an occurrence other than an accident, which affects or could affect the safety of operations." Accidents and incidents are the error experiences from which airlines have the potential to learn.

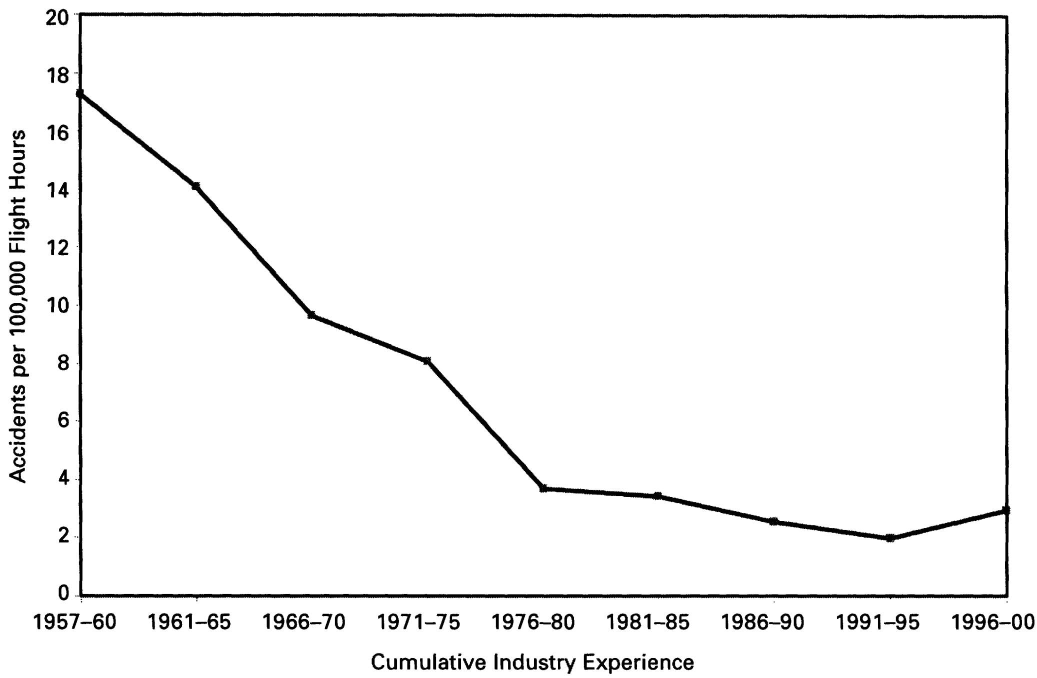
EFFECTS OF PRIOR EXPERIENCE ON LEARNING

In the literature on organizational learning there is a large body of work on the learning curve. The learning curve is an empirical finding showing that, in general, experience produces improvement. Early empirical work on the learning curve showed that the log of unit costs tends to decrease linearly with the log of cumulative production volume. So, for example, cumulative production experience tends to lower costs in shipbuilding and automotive production (Argote and Epple, 1990), nuclear power plant production (Zimmerman, 1982), and coal generation (Joskow and Rose, 1985). More recent work has moved away from a focus on cost reduction and productivity improvement to other outcomes of learning. These studies have shown that experience improves customer service and product quality (Darr, Argote, and Epple, 1995; Lapre, Mukherjee, and Van Wassenhove, 2000) and increases the survival rates of hotels (Ingram and Baum, 1997; Baum and Ingram, 1998) and banks (Kim and Miner, 2000).

In the context of airlines and their errors, it may be that airlines learn from error experience and are able to improve performance over time, reducing subsequent errors (i.e., accidents and incidents). If we look at the airline industry over long time periods, this seems to be the case. Figure 1 plots the accident rate (accidents per 100,000 hours flown) for all U.S. airlines from 1955 to 1997 and exhibits a characteristic learning curve, i.e., as experience accumulates with the passage of time, the error rate declines.

When individual airlines' accident rates are broken out, as they are in table 1 for some of the larger U.S. airlines, we see the same general decrease in accidents over time as in figure 1, but there is also a fair amount of variance across airlines. For example, from 1957 to 1986, American Airlines had an average of 10.3 accidents per million departures and US Air had 6.6. Variation in airline error rates could come from many sources. One obvious source is the characteristics of the individual airline, e.g., whether it is large or small, the age of its fleet, characteristics of its corporate culture, its management team, and its training procedures. Another possible source of variation, however, is differences in the characteristics of the accidents and incidents experienced by these different airlines. Because experience affects organizational

Figure 1. U.S. airline accident rates for all airlines by time period.*



* Source: NTSB, *FAA Flight Statistics Reports*, various years.

learning, different types of experiences are likely to produce variation in learning rates. One source of differences in experience is whether that experience has homogeneous or heterogeneous causes.

Homogeneity and heterogeneity of experience have been shown to affect learning about mergers and acquisitions (Beckman and Haunschild, 2002), and heterogeneity may also affect an organization's ability and/or motivation to learn from errors such as accidents and incidents. The distinction between heterogeneous and homogeneous causes of accidents is evident in the following examples of two recent airline accidents. Southwest Airlines had an accident in 1997 in

Table 1

Accident Rates (per Million Departures) for Selected Large Airlines*

	Period						
Airline	1957–60	1961–65	1966–70	1971–75	1976–82	1981–86	Average
Alaska	95.5	41.1	17.7	7.5	9.5	3.3	26.0
American	24.8	14.2	7.7	9.1	6.4	3.8	10.3
Braniff	7.9	10.8	9.0	4.5	4.7	4.0	6.7
Continental	13.2	20.9	7.3	4.1	8.4	7.8	10.1
Delta	10.9	11.7	11.2	11.6	4.9	1.6	8.3
Northwest	33.2	16.8	17.2	12.2	0.0	1.8	10.6
Pan Am	40.3	30.2	15.4	17.2	7.7	6.1	19.8
Southwest						2.9	2.9
TWA	19.4	21.5	8.1	15.4	2.8	1.6	10.9
United	11.5	12.6	12.6	4.6	1.6	4.5	7.7
US Air	12.5	7.7	6.7	6.1	4.7	3.6	6.6
Western	12.3	9.5	6.0	2.6	5.0	1.0	5.7

* Source: Rose (1990)

which one passenger was seriously injured and four others received minor injuries. The cause of this accident was attributed to (1) birds ingested in the left engine, which caused a compressor stall; (2) an improperly rejected takeoff by the pilot; (3–5) a flight attendant ordering a passenger evacuation without (a) informing the captain, (b) communicating with the other attendants, or (c) determining the location of a suspected (but not actual) fire; and (6) the airline failing to provide joint crew resource management (CRM) training to all crew members (NTSB report ATL96FA101). This is an accident with heterogeneous underlying causes. The causes interacted in complex ways (Perrow, 1984), resulting in an accident in which, had any one of these six causes not been present, the accident probably would not have occurred. In contrast, in a Delta Airlines accident in 1995, a flight attendant was seriously injured during an episode of severe turbulence. The cause of the accident was attributed to the fact that the flight attendant was not seated, despite notification from the pilot that everyone should be seated because of turbulence in the area (NTSB report MIA95LA055). This is an accident with a homogeneous underlying cause.

Clearly, some aviation errors are due to more heterogeneous causes than others, with the causes often interacting in complex ways. Also, some airlines tend to experience more heterogeneous, complex accidents and incidents than other airlines. It is not obvious, however, what the role of such heterogeneity is in the management of errors and whether complexity benefits or hinders airlines in their ability to learn from these experiences. Existing literature provides conflicting accounts, with some suggestion that people learn better from heterogeneity or complexity and some that they learn better from homogeneity or simplicity. The literature on organizational learning provides some starting points for understanding these issues. Scholars have identified several key learning processes, including trial-and-error learning, vicarious learning from others, and active inferential learning (Miner and Mezias, 1996). In rare events, such as an accident or near-accident, trial-and-error learning is difficult (Weick, 1987). Airlines instead try to overcome the paucity of experience by simulating many potential accident and incident situations. Yet the number of permutations and combinations of events that could conceivably cause a problem is too large for all situations to be simulated. As a result, some scholars are pessimistic about whether organizations like airlines can learn from their accident and incident experiences (Carroll, 1995). Despite the difficulties associated with learning from rare events, however, some hazardous systems (including airlines) do seem to have remarkably good safety records (Perrow, 1984; Roberts, 1991, 1993). Also, the fact that individual airlines have significantly reduced their accident rates over time suggests that some form of learning is occurring. Perrow (1994), more optimistic about the possibility that airlines can learn, suggested that the relatively low number of serious accidents in the airline industry relative to other industries may be due to industry-specific factors such as the efficient dispersion of information on aviation errors and airline density (Perrow, 1994: 17).

While the airline industry is heavily regulated, and the dispersion of information about accidents and incidents is fairly efficient, there is still likely to be variance in firm-level learning. Figure 1 showed that accidents and incidents are relatively rare, as less than 10 occur in every 100,000 flight hours. Most of these accidents are not serious, and fatal accidents are quite rare, occurring once in every three million flights, on average. If we look at the number of accidents per year instead of the rate per flight hour, however, accidents are actually not so rare. The nine largest airlines experienced an average of 3.1 accidents and 18 serious incidents per year during the 1986–1996 period. So while errors are rare on a relative basis, they are not all that rare on an absolute basis. Airline officials, pilots, and regulators do pay attention to both accidents and incidents and generally attempt to learn from them (Reason, 1997). This is consistent with the well-known idea that organizations tend to search for solutions when results fail to meet aspirations (March and Simon, 1958; Cyert and March, 1963; Greve, 1998). And contemporary learning studies show that organizational failures, such as accidents and incidents, are an important stimulant to organizational learning and change (Sitkin, 1992; Greve, 1998; Miner et al., 1999; Kim and Miner, 2000). Of course, attempts to learn are not always successful in terms of preventing future accidents and may even be harmful (Perrow, 1984), but there are likely to be some factors that affect whether these attempts to learn are successful, among them the homogeneity or heterogeneity of the causes of prior accident and incident experiences.

Heterogeneous Causes

There is some suggestion in the literature that prior experience with heterogeneous causes will be better for reducing subsequent errors than experience with homogeneous causes. Three factors suggest that heterogeneity will be better than homogeneity: (1) variance helps focus attention on latent causes and thus leads to a deeper analysis of the problem (Reason, 1997); (2) variance forces a situational analysis, rather than a simple "blame the operator" response, and (3) heterogeneity produces constructive conflict in groups, which leads to better analyses and problem responses (Jehn, Northcraft, and Neale, 1999). The argument that heterogeneity produces better analyses of latent causes comes from work on organizational accidents. Reason (1997) argued that organizations tend to focus on the surface when attempting to learn from failures. They focus on active failures, rather than trying to dig deeper and uncover problematic latent conditions. Heterogeneity in causes, with complex interactions of multiple factors, is likely to force organizations to look harder, to get away from the proximate or simple explanations.

As part of this study, we conducted informational interviews with executives of three major and six regional airlines. One of these executives provided an example of how heterogeneity can aid in a deeper causal analysis. This airline had two different accidents occur within a one-year period. One accident was caused when a pilot decided to take off without de-icing a plane (a decision under that pilot's control), and the ice buildup caused an accident. The second accident occurred as the result of a pilot calling in maintenance personnel to check out an engine warning light. After inspecting and fixing the problem, the maintenance crew released the aircraft to the flight crew but neglected to replace the drive covers on the engines, resulting in the loss of oil from both engines during flight. On the surface, the causes of these two accidents do not seem to be related. But they both occurred in an airline whose culture promoted on-time departures as a key goal, which affected the decisions made by both the pilot (to save time by not de-icing) and the maintenance worker (in a hurry to finish quickly and, so, forgot to replace the covers). Thus, an aspect of the organization's culture contributed to both accidents (cf. Reason, 1997; Vaughan, 1999). The airline executive noted that this discovery only came about because the airline underwent an in-depth consideration of the underlying causal structure of accidents, which was prompted by the different causes present in these two accidents.

Underlying systematic causes are more likely to be noticed and corrected when errors are diverse, as homogeneous errors will not lead to a deep analysis of the underlying structure of the problem. It is difficult to look for connections among disconnected events, but to the extent that it can happen, it is likely to result in better solutions. In fact, many researchers propose that organizations look deeper, and not focus narrowly in their analysis of causes, as a prescriptive solution to accident prevention (e.g., Reason, 1987). Thus, an airline experiencing heterogeneous accident causes may be forced to dig deeper, to get away from the proximate causes (e.g., the pilot, the maintenance person) and look at latent conditions (e.g., organizational culture).

A second, related reason why heterogeneous causes may lead to more learning from errors than homogeneity is that heterogeneity may force an organization to shift attention away from a given individual as the cause of the accident. As we know from attribution theory, there is a tendency to focus on the person, not the situation, as the cause of events (Nisbett and Ross, 1980; Fiske and Taylor, 1984). This is especially true when the events have serious consequences, which results in a tendency to "blame the operator" (Perrow, 1984). In airline accidents, there is a strong tendency to blame the pilot. Attributing the causes of an accident to human error inhibits the organization's ability to learn from that accident (Sagan, 1993), in part because once the human being has been fired, transferred, or replaced, there is assumed to be no more problem in the system. Yet Perrow (1984) and Reason (1997) noted that human error is rarely the only cause of an accident or incident. Thus, heterogeneous causes are more likely to lead an organization to acknowledge and deal with the multiple underlying causes, not all of which are human error, and thus lead to better understanding and solutions to the problems, which will reduce future accidents.

The final reason why heterogeneous causes may promote more learning than homogenous ones is that many error investigations, including airline accident investigations, are done by groups. The literature on small groups offers insight into the processes set in motion by diverse information and the links between diverse information and improved performance. Diverse information stimulates constructive conflict around issues, which leads people to deliberate about appropriate action, and this deliberation tends to improve group performance, especially on complex tasks (Jehn, Northcraft, and Neale, 1999). Thus, the process of attempting to learn from experience with heterogeneous causes is likely to stir constructive conflict, generate debate, and result in more perspectives on the problem. This, in turn, should lead to a better understanding of the problem and to solutions that reduce future accidents.

The above discussion suggests that experience that has heterogeneous causes puts organizations in a better position to learn from that experience, and this will result in a greater reduction in subsequent error rates than experience that has homogeneous causes. This is because heterogeneous experience results in more in-depth analysis of latent causes, a shifting of focus away from a blame-the-operator response, and/or more constructive conflict and airing of perspectives.

Hypothesis 1 (H1): The greater the heterogeneity of prior causes of an organization's errors, the lower the subsequent error rate for that organization.

While each accident and incident has an associated level of heterogeneity in causes, each airline also has an associated level of heterogeneity in causes when the heterogeneity of its accidents and incidents is aggregated over some time period. It is not clear whether airlines will learn from heterogeneity in causes within or across accidents, or both. The argument about in-depth analysis and the example from our interviews about discovering a cultural explanation would suggest that heterogeneity in causes across accidents might be more likely to trigger this analysis process. The argument about constructive group conflict might be more applicable to one particular accident, as different groups are likely to be involved in investigating different accidents. Because of this ambiguity, we conducted and report analyses at both levels: within-accident heterogeneity and across-accident heterogeneity.

We had earlier suggested that one reason why heterogeneity in causes might result in fewer and less severe accidents and incidents is that causal heterogeneity is more likely than homogeneity to result in a reduction in the tendency to blame the operator. Therefore, if we do find that the heterogeneity of causes reduces error rates, and if the reason behind this decline is a decrease in the tendency to blame the operator (the pilot), then we should see fewer operator errors in airlines with more prior experience with heterogeneous causes. This leads to the following hypothesis:

Hypothesis 2 (H2): The greater the heterogeneity of prior causes of an organization's errors, the lower the subsequent number of operator errors for that organization.

Homogeneous Causes

Although we have argued that heterogeneity in the causes of errors might benefit organizations in their attempts to learn from their errors, there are also theoretical arguments for why homogeneity, rather than heterogeneity, might benefit learning. Three mechanisms may explain how the homogeneity of a firm's own experience facilitates learning: (1) salience and attention factors, (2) ease of understanding, and (3) as an aid to viewing errors as systematic rather than random. Each of these mechanisms may, in turn, lead to better problem fixes, which are then reflected in a reduction in subsequent error rates.

One reason why firms have difficulty learning and changing is inertia (Hannan and Freeman, 1984). Organizations find it difficult to recognize and respond to problems, which makes factors that overcome inertia important in stimulating learning and change. One such factor is salience (March, Sproull, and Tamuz, 1991), and one determinant of salience is repetition (Fiske and Taylor, 1984). Repetition of the same causes is likely to make those causes salient. Thus, the repetition of the same accident causes found with homogeneous error experience may aid in learning and thus reduce subsequent error rates.

One might think that in situations in which errors have potentially serious consequences, and in highly regulated industries, any problem that poses a potential threat to safety would be swiftly and effectively dealt with, but that does not always happen. Marcus and Nichols (1999) listed several warning signs preceding disasters like Bhopal and Three Mile Island–and how these signs were ignored. Sagan (1993) outlined several nuclear-alert activities that continue to be practiced, despite their dangers having been demonstrated during the Cuban missile crisis. Nance (1986) discussed the repeated problems with Boeing 737-200s "pitching up," which received little attention. Thus, even in the airline industry, inertia may have to be overcome. Repetition of the same accident causes found with certain types of accident experience may help overcome inertia and thus reduce subsequent error rates of that same type.

Hypothesis 3 (H3): The greater the number of an organization's prior errors of a given type, the lower the subsequent rate of errors of that same type.

Not all experience with simple causes constitutes repetition of the same events. Some organizations may just have errors with relatively simple causes, though the causes may differ from event to event. Decision-making theories suggest that such experience is easier to learn from than events with more heterogeneous causes, because simple causes are easier to understand and analyze than complex interactions of multiple factors. Since complex causes are ambiguous, they tend to produce biased interpretations, reconstructions of history to meet perceptions, and myths, fictions, and stories (March and Olsen, 1988; Sagan, 1993). The complexity involved in multiple heterogeneous causes may therefore result in poorer problem fixes than those resulting from an analysis of simpler causes.

Another reason why homogeneity in causes may result in more learning than heterogeneity comes from Reason's (1997) analyses of biases in error management. Reason suggested that one key bias in accident management is that errors tend to be viewed as random, not systematic. It is easy to imagine that the six factors that interacted to produce the Southwest Airlines accident described earlier would be highly unlikely to occur as a unit again. Viewing errors as random is likely to hinder learning by decreasing motivation and preventing an in-depth analysis of the problem situation. If homogeneity helps learning because it is easier to understand simple situations or because it helps overcome the tendency to view errors as random, and thereby do little about them, then this leads to the following counter-hypothesis to H1:

Hypothesis 4 (H4): The greater the homogeneity of prior causes of an organization's errors, the lower the subsequent error rate for that organization.

Generalists and Specialists

The above discussion assumes that all airline organizations learn equally from their errors, but this assumption may not be warranted. In the airline industry, for example, Morris and Moore (2000) found that private pilots tend to learn more from their accident experience than commercial airline pilots. They attribute this effect to the greater accountability of commercial pilots to organizational supervisors, which dampens their ability to learn from the experience. In commercial airlines, there may be analogous differences in learning between generalists and specialists.

There is much work in organizational ecology investigating the differential life chances of generalists and specialists over time (e.g., Carroll, 1985). Specialist organizational forms depend on a narrow range of environmental resources for survival (Freeman and Hannan, 1983; Carroll, 1985). Their formal structure, patterns of activity, and normative order are different from generalist organizations, which depend on a wide range of environmental resources. Carroll (1985) outlined the dynamics of markets composed of generalists and specialists in his resource-partitioning theory, and empirical support for resource-partitioning predictions has been found by researchers using data from several different industries, including beer brewing (e.g., Carroll and Swaminathan, 1992), wine production (Swaminathan, 1995), automobile manufacturing (Torres, 1995), and airlines (Seidel, 1997). These studies have made advances in our understanding of the life chances of specialists and generalists over time, but only a small number of studies have investigated the differences between generalists and specialists in their learning processes. These studies include Barnett, Greve, and Park (1994), who found that specialist banks have higher returns to experience (measured as a return on assets) than generalists; Ingram and Baum (1997) who found that generalists benefit less from their own operating experience and more from industry experience than specialists; and the above-mentioned study by Morris and Moore (2000). This work suggests that generalists and specialists may also differ in their ability to learn from prior experience with heterogeneous causes.

Generalists and specialists differ in the complexity of their organizational form. Generalists are more complex because their larger resource space and broader range of products for customers mean that they have to deal with a wider range of issues than specialists do. Complexity magnifies the problems organizations have in controlling and coordinating behavior. Complexity may also dampen learning from experience, especially heterogeneous experience. Complex organizations are likely to be more political, and politicized organizations investigate accidents in ways that do not necessarily promote accurate learning but, rather, protect the interests of the powerful actors in the organization (Perrow, 1984; Sagan, 1993). Complex organizations also tend to have more hierarchy and more compartmentalization, which makes it difficult to get accurate, complete feedback from the operators involved with an accident or incident. Actors in one part of the organization may not know what those in other parts are doing, let alone learn from their experiences (Sagan, 1993). Because errors with heterogeneous underlying causes are inherently complex, the information from such errors is less likely to be utilized effectively in complex systems, as the system is already overloaded with diverse information. It is more likely to be used effectively in simpler organizations. This suggests that specialist organizations will learn more from heterogeneous prior error experience than generalists.

Hypothesis 5 (H5): Specialist organizations will learn more from errors with heterogeneous causes than will generalist organizations.

METHOD

Sample and Data

The initial sample for the study is all U.S.-based commercial airlines that existed during the 1983–1997 period. The U.S. National Transportation Safety Board (NTSB) tracks the accidents and serious incidents for these airlines and publishes this information in the NTSB Accident Database. Private pilots are not included in the database. There were 310 airlines in existence during the 1983–1987 period, though there was also a lot of change, with a number of new entrants, failures, and mergers. The airlines in the dataset vary substantially in size. The NTSB classifies airlines into the following size categories: large majors (e.g., American Airlines), nationals (e.g., Alaska Air), large regionals (e.g., Midwest Express), and medium regionals (e.g., Independent Air).

We collected data from the NTSB database on all the accidents and incidents experienced by these airlines during the 1983–1997 period. We used 1983 as our starting year because accident reporting went through some changes that year, so data prior to 1983 are not completely comparable with data after 1983. As noted earlier, the NTSB defines any event that led to human injury, death, or serious equipment damage as an accident. An incident is an event not classified as an accident in which a hazard or potential hazard to safety was involved (NTSB, 2001). Because accidents are difficult to conceal, the reporting of them is quite accurate (Rose, 1990). Incidents were more inconsistently reported, but the 1984 implementation of a computerized surveillance system by the Federal Aviation Administration (FAA) substantially increased the number of incidents reported (Tamuz, 1987). To check for effects of possible bias in incident reporting, we analyzed accidents and incidents both separately and together. We also combined both fatal and nonfatal accidents in our analyses for two reasons: (1) the incidence of fatal accidents is extremely rare, making the estimation of fatal-accident models very difficult, and (2) fatal and nonfatal accidents seem to be affected by the same factors (Rose, 1990). The 310 commercial airlines in our sample experienced a total of 1,346 accidents and incidents during the 1983–1997 period. There were 12 large majors in the dataset, and they experienced a total of 531 accidents and incidents over the study period. There were also 34 nationals with 220 accidents and incidents, 90 large regionals with 224 accidents and incidents, and 174 medium regionals with 371 accidents and incidents. Calculating the event rates for these airlines reveals that while the large majors had larger numbers of accidents, their rate of accidents/incidents per 100,000 departures is only 1.18, while the nationals had an average rate of 2.48, the large regionals a rate of 5.99, and the medium regionals a rate of 8.78. Thus the smaller airlines, while they experience fewer accidents and incidents on an absolute basis, actually have higher rates than the large majors.

Dependent Variable

Our primary dependent variable is the accident and incident (event) rate for each airline, which we calculated as the number of accidents plus incidents per 100,000 departures. Using the rate per departure is a common way to measure accident rates, as most accidents occur on takeoff and landing, thus making the rate per departure more relevant than the rate per miles flown (Fromm, 1968). This variable was calculated for each airline and updated annually. We conducted analyses of accident rates and incident rates both separately and together as event rates.

Independent Variables

Prior accident heterogeneity. We constructed the prior-accident heterogeneity measure as follows. We first used the NTSB accident/incident reports to construct a set of 23 accident and incident cause categories. The NTSB reports contain narrative descriptions of accidents and incidents and, in some cases, a listing of NTSB-generated codes for the accident or incident's observed cause(s). Because not all reports included NTSB-generated codes, we coded causes from the narratives and then checked our coding against the NTSB-generated codes. A sample accident report and a listing of our cause codes are shown in Appendix A and Appendix B. Two coders reviewed each NTSB accident or incident report and classified the accident/incident causes into the 23 cause categories listed in Appendix A. After a training period involving the coding of 50 reports, we calculated interrater reliability. Cohen's Kappa (k) (Cohen, 1960) for all accident causes was .87, which is deemed good. After the training period, the coders worked alone. All discrepancies were resolved through discussion.

To construct an accident-cause heterogeneity measure, we used an entropy-based index, which is typical in measuring diversity or heterogeneity with categorical variables (Teachman, 1980; Ancona and Caldwell, 1992; Jehn, Northcraft and Neale, 1999). The index is calculated as follows:

$$\text{Diversity} = -[P_i(\ln P_i)], \text{ for } i = 1 \ldots 23$$

where P is the proportion of causes in each category i, and i represents the 23 possible categories of causes. Higher scores on the index mean the accident had more heterogeneous underlying causes: the larger the number of causes and the more equal the proportion of each, the greater the heterogeneity. The index is not the same as a simple count of the number of causes, as an accident may have multiple causes of a single type. For example, an accident that was caused by two mechanical errors and one pilot error has less heterogeneous causes than an accident caused by one mechanical error, one pilot error, and one maintenance error. We also conducted analyses using a simple count measure instead of the heterogeneity measure. The results of these analyses are similar to those using the heterogeneity index, but they are also weaker in effect, suggesting that there is additional information in the heterogeneity measure that is not captured in a simple count of accident causes. We therefore report analyses using the heterogeneity measure (results of the count analyses are available from the authors).

Each accident and incident has an associated heterogeneity measure. Each airline also has an associated heterogeneity measure when the heterogeneity of its accidents and incidents is aggregated over a period of time. Since it is not clear whether airlines learn from heterogeneity or homogeneity within accidents or across accidents (or both), we conducted and report analyses at both levels. For the within-accident heterogeneity measure, we looked at the causes of each accident and constructed the measure at the accident level. For the across-accident heterogeneity measure, we looked at all prior accident causes and constructed the measure at the organization level.

It is also unclear how much prior error experience is relevant and whether the effects of prior experience decay over time. Studies in other contexts have found that the value of experience does tend to decay over time (e.g., Argote, 1999; Baum and Ingram, 1998), but they were done in contexts in which events are repeated frequently, such as product production or years of operating experience. We were dealing with relatively rare but salient events. The value of these events was not likely to depreciate as rapidly as that of the events in these other studies. Yet it is still likely that accidents that occurred in 1983 were more relevant in 1984 than they would have been ten years later, in 1994.

We therefore used as our starting assumption that learning can occur within or across prior events but that the value of those events depreciates with time. We tested this assumption by running analyses using (1) only prior year's events,

(2) a rolling window of the prior three years' events, and (3) all prior events, back to the beginning of the sample. We also used various methods of discounting prior experience, to account for depreciation in the impact of events. As it is not clear a priori which discounting method is appropriate, we followed Baum and Ingram (1998) and ran analyses using four types of discounting. The first is no discount (e.g., each event is weighted by one). The second is discounting by the age of experience (e.g., for the rolling three-year window analyses, events the prior year are weighted by 3/3, the year before by 2/3, and the year before that by 1/3). The third is discounting by the square root of the age of experience, to account for the possibility that events depreciate more slowly than a straight age discount. Finally, the fourth is discounting by the square of the age of experience, to account for the possibility that events depreciate faster than a straight age discount.

We generally found few differences in hypothesized effects across these various accumulation and discounting possibilities. The models using a prior three-year window are fairly stable and tend to fit better than those using only the prior year's experience and those using experience back to the start of the sample. In addition, discounting the prior three years' experience linearly (i.e., weighting last year's event by 3/3, the year before by 2/3, and the year before by 1/3) produces a better model fit than other discounting specifications (e.g., the square root or no discount). Therefore, we report analyses using a prior three-year window of experience, discounted linearly to account for depreciation of knowledge. As our sample starts in 1983, observations for 1984 and 1985 were based on a one-year and two-year rolling window instead of three years. The observations were still discounted.[1]

Because experience was aggregated across three years and then discounted, the within-event heterogeneity measure was constructed as follows: we measured heterogeneity for each airline's accident or incident, discounted each year's measure appropriately, then averaged the prior three years' measures.

$$\begin{gathered}\text{Within-event heterogeneity, Airline X Year}_y = \\ \text{mean}\ \{(\text{H A1}_{y-1}, \text{H A2}_{y-1} \ldots \text{H An}_{y-1})*1 + \\ [(\text{H A1}_{y-2}, \text{H A2}_{y-2} \ldots \text{H An}_{y-2})*.67] + \\ [(\text{H A1}_{y-3}, \text{H A2}_{y-3} \ldots \text{H An}_{y-3})*.33]\}\end{gathered}$$

where y = year; H = heterogeneity; A1, A2 . . . An = accident or incident 1 . . . n for a given airline.

The across-event heterogeneity measure was constructed as follows: we measured heterogeneity across all an airline's accidents and incidents for a given year, discounted appropriately, and summed the prior three years' measures. We summed instead of averaging on this measure because across-accident heterogeneity means that the prior three years' experience essentially should count as one accident with multiple causes.

1
Distributional diagnostic tests performed on the heterogeneity variables show the distribution of the heterogeneity variable is actually fairly smooth and normal. There is a clustering of values near zero and a few outliers at the very highest levels. We ran a series of analyses to see if our results are sensitive to the deletion of the near-zero values or the outliers (which occur in the top 1 percent of the values). Results show that the deletion of these cases does not change the significance of our hypothesized results.

$$\text{Across-event heterogeneity, Airline X Year}_y = (\text{H A1} \ldots n_{y-1}) + [(\text{H A1} \ldots n_{y-2})*.67] + [(\text{H A1} \ldots n_{y-3})*.33]$$

Table 2 presents some sample heterogeneity calculations for 1986, within and across accidents and incidents, and discounted and undiscounted. As can be seen from this table, the highest levels of heterogeneity are for those airlines whose accidents and incidents have many causes (denoted by C) distributed across the various cause categories (e.g., Airline F in table 2).

Generalist or specialist. We used the *Official Airline Guide (OAG),* U.S. edition, and data from the Federal Aviation Administration and NTSB to classify airlines into specialists and generalists. The measure was designed to capture organizational complexity. Organizational ecology studies have been moving away from using measures of generalism and

Table 2

Sample Heterogeneity Calculations for Accident and Incident Causes (C) for 1986

Year	Airline A	Airline B	Airline C	Airline D	Airline E	Airline F
1983	C1	C1,C2,C3,C4	C1	C1	C1,C2,C3,C4	C1,C2,C3,C4
1983			C1	C2	C1,C2,C3,C4	C5,C6,C7,C8
1984			C1	C3	C1,C2,C3,C4	C9,C10,C11,C12
1985			C1	C4	C1,C2,C3,C4	C13,C14,C15,C16

Heterogeneity Calculations		Het = – {Sum[(Pi)*(lnPi)]} for i = 1...23	
Airline A	Within-accident heterogeneity	–[(1/1)*ln(1/1)]	0.00
	Within-accident heterogeneity (disctd.)	–[(1/1)*ln(1/1)]*.33	0.00
	Across-accident heterogeneity	Same as within	0.00
	Across-accident heterogeneity (disctd.)	Same as within	0.00
Airline B	Within-accident heterogeneity	–{4*[(1/4)*ln(1/4)]}	1.39
	Within-accident heterogeneity (disctd.)	–{4*[(1/4)*ln(1/4)]}*.33	0.46
	Across-accident heterogeneity	Same as within	1.39
	Across-accident heterogeneity (disctd.)	Same as within	0.46
Airline C	Within-accident heterogeneity	–mean {[(1/1)*ln(1/1)] + [(1/1)*ln(1/1)] + [(1/1)*ln(1/1)] + [(1/1)*ln(1/1)]}	0.00
	Within-accident heterogeneity (disctd.)	–mean {[(1/1)*ln(1/1)]*.33 + [(1/1)*ln(1/1)]*.33 + [(1/1)*ln(1/1)]*.67 + [(1/1)*ln(1/1)]*1}	0.00
	Across-accident heterogeneity	–[(4/4)*ln(4/4)]	0.00
	Across-accident heterogeneity (disctd.)	–{[(2/4)*ln(2/4)]*.33 + [(1/4)*ln(1/4)]*.67 + [(1/4)*ln(1/4)]*1}	0.00
Airline D	Within-accident heterogeneity	–mean {[(1/1)*ln(1/1)] + [(1/1)*ln(1/1)] + [(1/1)*ln(1/1)] + [(1/1)*ln(1/1)]}	0.00
	Within-accident heterogeneity (disctd.)	–mean {[(1/1)*ln(1/1)]*.33 + [(1/1)*ln(1/1)]*.33 + [(1/1)*ln(1/1)]*.67 + [(1/1)*ln(1/1)]*1}	0.00
	Across-accident heterogeneity	–{4*[(1/4)*ln(1/4)]}	1.39
	Across-accident heterogeneity (disctd.)	–{[(1/4)*ln(1/4)]*.33 + [(1/4)*ln(1/4)]*.33 + [(1/4)*ln(1/4]*.67 + [(1/4)*ln(1/4)]*1}	0.81
Airline E	Within-accident heterogeneity	–mean {4*[(1/4)*ln(1/4)] + 4*[(1/4)*ln(1/4)] + 4*[(1/4)*ln(1/4)] + 4*[(1/4)*ln(1/4)]}	1.39
	Within-accident heterogeneity (disctd.)	–mean {4*[(1/4)*ln(1/4)]*.33 + [4*(1/4)*ln(1/4)]*.33 + 4*[(1/4)*ln(1/4)]*.67 + 4*[(1/4)*ln(1/4)]*1}	0.81
	Across-accident heterogeneity	–{4*[(4/16)*ln(4/16)]}	1.39
	Across-accident heterogeneity (disctd.)	–{[4*(2/16)*ln(2/16)]*.33 + [4*(1/16)*ln(1/16)]*.67 + [4*(1/16)*ln(1/16)]*1}	0.56
Airline F	Within-accident heterogeneity	–mean {4*[(1/4)*ln(1/4)] + [4*(1/4)*ln(1/4)] + 4*[(1/4)*ln(1/4)] + 4*[(1/4)*ln(1/4)]}	1.39
	Within-accident heterogeneity (disctd.)	–mean {4*[(1/4)*ln(1/4)]*.33 + [4*(1/4)*ln(1/4)]*.33 + 4*[(1/4)*ln(1/4)]*.67 + 4*[(1/4)*ln(1/4)]*1}	0.81
	Across-accident heterogeneity	–{16*[(1/16)*ln(1/16)]}	2.77
	Across-accident heterogeneity (disctd.)	–{8*[(1/16)*ln(1/16)]*.33 + 4*[(1/16)*ln(1/16)]*.67 + 4*[(1/16)*ln(1/16)]*1}	1.62

specialism that are highly correlated with size to those that more directly capture organizational complexity. These measures tend to vary substantially from study to study, as they tend to be rooted in the specifics of the industry being studied. So, for example, Carroll (1985) defined generalist newspaper organizations as those that publish one or more general-interest papers (Carroll, 1985). Dobrev, Kim, and Hannan (2001) used the spread of engine capacity over all models produced by an automotive firm as a measure of generalism. We used a commonly accepted airline industry standard for measuring level of specialization: fleet diversity (cf. Seidel, 1997). Fleet diversity was calculated using the same formula as causal heterogeneity, Diversity = $-[P_i(\ln P_i)]$, for i = 1 . . . n, where P is the proportion of planes in each type i, and i represents the possible types of planes (e.g., 727s, 737s, MD80s). Higher scores on the index mean the airline has a more heterogeneous fleet. An airline with only one type of plane in its fleet will have a fleet diversity score of zero (i.e., is highly specialized).

We wanted to be sure that our generalist and specialist codings were not simply size categories. Although generalists are usually large and specialists small, size and generalism are really separate concepts. There are small generalists and large specialists. Recent work in organizational ecology has noted the importance of distinguishing size effects from generalism effects (Dobrev, Kim, and Hannan, 2001). We therefore included separate measures of airline size. Since we controlled for size in all models, we should be able to detect the independent effect of generalism or specialism in our results. In our data, size and generalism are correlated at .594. This indicates that while size and form do covary, there is still a fair amount of independent information in the generalism/specialism variable. Some airlines in our sample are large, in the sense that they have many departures, yet they also have homogeneous fleets. Mesa West is one example, and Southwest Airlines, which still flies with only one type of airplane in its fleet, is another. America West, in contrast, is a relatively small generalist.

Control Variables

Prior studies suggest several control variables that are important for these analyses. One control is airline size, as small carriers tend to have higher accident and incident rates than large carriers. We used a commonly accepted measure of airline size: Revenue passenger miles (in thousands) flown by each airline during the prior year. These data are available from the *Air Traffic Statistics* published by the U.S. Department of Transportation, Bureau of Transportation Statistics. As this variable is highly skewed, we logged it before entering it into the analysis.

Another important control variable is the airline's experience with accidents and incidents. We wanted to be sure that the main independent variable, accident heterogeneity, was independent of accident volume. Additionally, as noted earlier, there is substantial research showing that organizations tend to improve with experience (see Argote, 1999, for a review). While this effect, to our knowledge, has not been shown for error reduction as a dependent variable, there is no reason to think that it might not apply. That is, organizations may tend to reduce their error rates as a function of the volume of error experience (rather than, or in addition to, the heterogeneity of experience). Thus, we used the cumulative number of prior accidents and incidents for each airline, using the same rolling three-year window as our heterogeneity measure. We also discounted this experience variable in the same way as the heterogeneity variable, using the age of experience. Thus, for accidents that occurred in 1990, we used the cumulative number of accidents experienced by that airline from 1987 to 1989 (discounted so that 1987 accidents carry 1/3 the weight of 1989 accidents).

As our data are left truncated at 1983, we included organizational age as a fifth control variable. This controls for the left-truncation on variables like the number of miles flown and the number of prior accidents for those airlines that existed prior to 1983. We also obtained data on airline profitability, measured as return on assets, to control for the possibility that less profitable airlines may cut corners on safety and thus have more accidents than more profitable airlines, an effect found by Rose (1990). We also controlled for the severity of each accident or incident, as it may be that more severe accidents have more causes attributed to them. Severity data are available from the NTSB reports, where severity is reported on a scale of 1–4 (1 = not at all severe, e.g., a minor injury occurred, to 4 = severe, e.g., one or more deaths occurred).

Finally, we also included variables that measure industry-level cumulative experience, which increases with time, to control for effects that are associated with a change in airlines' accident and incident rates over time as well as the influence of experience outside the organization. Changes over time may be due to things like general technological improvements, increased regulatory surveillance or other regulatory changes, or other factors outside the airline (as within-airline improvements are captured in airline-level variables). Outside experience (the experience of other airlines) may also affect airlines' accident rates directly, as some studies have shown that firms can learn from the production and operating experience of their competitors (Zimmerman, 1982; Irwin and Klenow, 1994; Ingram and Baum, 1997). We controlled for industry experience with a variable measuring cumulative revenue miles for all airlines (in millions).[2] This variable was lagged by one year. As this variable was highly skewed, we logged it prior to entering it into the analysis. We also conducted analyses using dummy variables for each year (1983–1997) and a measure of elapsed time instead of these industry-level measures. The results of our main hypothesized effects did not differ in these analyses.[3]

Analysis

The data consist of a panel of observations on organization-years. The full sample included 2,710 firm-years of observations. But 929 of these observations occurred in firm-years in which the airline in question had no accidents or incidents during the prior three years, so the heterogeneity

2
We also ran models in which we substituted cumulative departures as a measure of industry experience, with no change in the hypothesized results.

3
When exact data on departures, miles flown, and airline profitability were missing, we used the size distributions of firms with observed values for these variables to impute values for the missing cases. We coded airlines into small, medium, and large and randomly assigned a value from a uniform distribution of the variables for other firms in the relevant size category for the relevant year. Almost all of the missing data were for very small airlines. We also ran analyses excluding these airlines from the analyses, with no difference in the hypothesized results.

variable could not be calculated for these airlines during these years. To be sure that excluding these observations did not bias our results, we ran analyses using the Heckman (1979) selection bias model to estimate the dependent variable (accident and incident rates) in the full sample. We then used the parameter estimates (the inverse Mills ratio scores) from that model in a second stage model to predict the effects of heterogeneity on accident and incident rates for those firms that had at least one accident or incident during the prior three years. The hazard rate from the selection model is labeled "hazard" in the subsequent analyses and effectively controls for the likelihood of an observation being included in the subsample (i.e., the likelihood of having one or more accidents or incidents in the prior three years).

Our main analyses used annual accident and incident rates as the dependent variable. As we had multiple observations for each airline, and airlines may have had more than one accident or incident in a given year, we used random-effects regression to correct for firm-specific autocorrelation. Airlines vary in characteristics that might affect their accident rates, e.g., they have older fleets, they operate in areas with harsh weather conditions, etc. These firm-specific characteristics affect rates across airlines and also mean that the longitudinally clustered data violate the underlying assumption of independence. One way to account for unobserved heterogeneity with these types of data is to estimate random-effects models that account simultaneously for within- and between-effects information. The random-effects estimator is the equivalent of estimating $(y_{it} - \theta y_i) = (1 - \theta)\alpha + (\mathbf{x}_{it} - \theta \mathbf{x}_i)\beta + [(1 - \theta)\nu_i + (\varepsilon_{it} - \theta\varepsilon_i)]$ where ν_i is the firm-specific residual, ε_{it} is the error term, and θ is a function of σ_ν^2 and σ_ε^2. We also ran models using a fixed-effects estimator, and all hypothesized results remained unchanged in these models (results of these analyses are available from the authors).

RESULTS

Table 3 presents descriptive statistics and correlations for key study variables. There are some fairly high correlations among the independent variables. Many of the largest are among variables that are not run in the same model (e.g., within-event heterogeneity and across-event heterogeneity are correlated at .71). Own prior events (our experience measure) and specialism are fairly highly correlated with some of the heterogeneity measures. Our large sample size, however, produces a high level of statistical power, which can overcome even extremely high correlations among variables (e.g., $r > .95$) (Mason and Perreault, 1991), but we can also assess results independent of collinearity concerns by using nested models. This allows us to compare model fit across models, thus overcoming any problems of interpretation that might be caused by multicollinearity among the independent variables, as multicollinearity does not affect model fit. We therefore used nested models for our primary analyses and report the significance of changes in model fit using likelihood ratio tests.

Table 3

Descriptive Statistics and Correlations for Key Study Variables

Variable	Mean	S.D.	Min.	Max.	1	2	3	4
1. Event rate (acc./inc. per 100k departures)	4.07	10.77	.00	117.44				
2. Accident rate	2.42	8.46	.00	110.00	.90			
3. Incident rate	1.72	6.17	.00	117.44	.74	.49		
4. Within-event heterogeneity (disc.)	.44	.23	.02	1.35	–.06	–.05	–.05	
5. Across-event heterogeneity (disc.)	2.19	1.99	.00	6.78	–.15	–.14	–.09	.71
6. Within-accident heterogeneity (disc.)	.42	.21	.00	1.21	–.06	–.07	–.04	.88
7. Across-accident heterogeneity (disc.)	.99	.74	.00	3.94	–.12	–.08	–.08	.73
8. Within-incident heterogeneity (disc.)	.31	.19	.00	.99	–.03	–.02	–.02	.93
9. Across-incident heterogeneity (disc.)	2.04	1.47	.00	4.29	–.09	–.02	–.06	.61
10. Level of specialism (fleet diversity)	.68	.82	.00	2.41	.14	.10	.10	–.31
11. Own prior events (disc.)	2.89	4.82	.00	26.00	–.15	–.15	–.09	.58
12. Industry experience	9.93	.96	8.12	11.07	–.03	.01	–.03	.13
13. Airline size	10.29	3.30	.00	13.73	–.20	–.18	–.15	.44
14. Airline age	16.99	17.22	.00	76.00	–.10	–.15	–.08	.42
15. Airline profitability	.36	4.04	–19.56	9.40	.03	.04	.01	–.09
16. Accident/incident severity	1.36	.65	1.00	4.00	.13	.11	.10	.20

Variable	5	6	7	8	9	10	11	12	13	14	15
6. Within-accident heterogeneity (disc.)	.72										
7. Across-accident heterogeneity (disc.)	.94	.78									
8. Within-incident heterogeneity (disc.)	.69	.54	.39								
9. Across-incident heterogeneity (disc.)	.71	.49	.42	.84							
10. Level of specialism (fleet diversity)	–.63	–.29	–.62	–.18	–.54						
11. Own prior events (disc.)	.82	.56	.86	.59	.81	–.63					
12. Industry experience	.19	.18	.13	.16	.18	–.08	–.01				
13. Airline size	.53	.41	.40	.38	.32	–.59	.50	–.08			
14. Airline age	.46	.40	.44	.47	.44	–.74	.75	.11	.54		
15. Airline profitability	–.19	–.05	–.08	–.04	–.06	.13	–.14	–.04	–.09	.05	
16. Accident/incident severity	.32	.20	.20	.18	.18	–.23	.52	–.02	.24	.23	–.02

Accident and Incident Rates

Table 4 presents the results of a random-effects regression in which the dependent variable is the number of accidents and incidents per 100,000 departures for each airline. Model 1 of table 4 includes only the control variables, and results show that older airlines tend to have fewer accidents and incidents than newer airlines. Our experience effect (own prior events, which are prior accidents and incidents) is negative and significant, indicating a basic learning effect: experience with accidents and incidents decreases subsequent accident/incident rates. Average prior event severity is positively related to subsequent event rates. Airline size, profitability, and industry experience do not appear to affect accident and incident rates in this model.

To test H1 and H4, whether heterogeneity or homogeneity reduces subsequent accident rates, we first added a measure of discounted within-event heterogeneity in model 2. As model 2 shows, the effect of within-event heterogeneity is negative and significant, supporting H1. The more heterogeneous the airline's prior accident and incident experience, the lower its subsequent accident and incident rate. Model fit improves significantly with the addition of prior experience heterogeneity. Thus, across all airlines, it appears that airlines learn more from heterogeneous than homogeneous accident and incident experiences. The coefficient on within-event heterogeneity is estimated at –3.840 (standard error = 1.15). Thus, an increase of .23 in discounted within-event

Table 4

Random Variable Effects Estimates of Carrier Accident and Incident Rates (Events per 100,000 Departures)

Variable	1	2	3	4*	5
(Log) Airline size	0.125	0.160	0.143	0.157	0.159
	(0.127)	(0.127)	(0.127)	(0.087)	(0.089)
Airline age	–0.215••	–0.226••	–0.214••	–0.877••	–0.836••
	(0.078)	(0.079)	(0.079)	(0.129)	(0.133)
Own prior events	–0.041•	–0.045•	–0.022	–0.244••	–0.114
	(0.020)	(0.019)	(0.020)	(0.094)	(0.094)
Airline profitability (ROA)	–0.004	–0.003	–0.020	0.011	0.007
	(0.047)	(0.047)	(0.048)	(0.055)	(0.056)
(Log) Industry experience	1.390	1.604	1.703	1.541	1.499
	(0.871)	(0.876)	(0.883)	(0.813)	(0.814)
Event severity	2.239••	2.238••	2.339••	2.134••	2.136••
	(0.378)	(0.376)	(0.376)	(0.334)	(0.338)
Within-event heterogeneity		–3.840••		–5.984••	
		(1.150)		(1.667)	
Across-event heterogeneity			–0.834••		–1.814••
			(0.258)		(0.337)
Hazard	0.011	0.013	0.015	0.009	0.006
	(0.009)	(0.011)	(0.010)	(0.008)	(0.008)
Constant	–6.372	–11.250•	–12.944••	–12.454	–12.872
	(4.666)	(4.679)	(4.752)	(8.459)	(8.460)
N (total firm-years)	1781	1781	1781	1265	1265
Log-likelihood	–5167.48	–5148.61	–5162.26	–3170.42	–3169.84
Likelihood ratio		37.74••	10.44••		
D.f. (vs. model no.)		1 (M1)	1 (M1)		

Variable	6†	7†	8‡	9‡
(Log) Airline size	–0.118	–0.103	0.184	0.192
	(0.114)	(0.115)	(0.077)	(0.084)
Airline age	–0.238••	–0.208••	–0.222••	–0.208•
	(0.064)	(0.069)	(0.061)	(0.062)
Own prior events	–0.070•	–0.135•	–0.091•	–0.037
	(0.030)	(0.068)	(0.044)	(0.048)
Airline profitability (ROA)	0.006	–0.001	–0.003	–0.005
	(0.037)	(0.040)	(0.028)	(0.029)
(Log) Industry experience	0.988	0.989	1.263	1.244
	(0.584)	(0.592)	(0.648)	(0.622)
Event severity	1.177•	1.142•	1.845••	1.827••
	(0.408)	(0.412)	(0.312)	(0.313)
Within-event heterogeneity	–2.384•		–0.130	
	(1.102)		(0.738)	
Across-event heterogeneity		–0.584•		–0.525•
		(0.268)		(0.260)
Hazard	0.009	0.009	0.010	0.009
	(0.007)	(0.009)	(0.008)	(0.010)
Constant	–14.874•	–14.888•	–8.872•	–7.926
	(6.824)	(6.439)	(4.872)	(4.826)
N (total firm-years)	1781	1781	1781	1781
Log-likelihood	–4849.23	–4852.34	–5166.32	–5164.82

•• $p < .01$; • $p < .05$; two-tailed tests.

* Models 4 and 5 eliminate accident causes over which airlines have relatively little control.

† The dependent variable for this model is accidents per 100,000 departures; heterogeneity is calculated for accidents alone.

‡ The dependent variable for this model is incidents per 100,000 departures; heterogeneity is calculated for incidents alone.

heterogeneity (one standard deviation) will reduce the expected accident and incident rate by approximately .88 accidents and incidents per 100,000 departures.

In model 3, we substituted a measure of discounted across-event heterogeneity, and the effect of across-event heterogeneity is also negative and significant, again supporting H1 and not H4. Thus, both within- and across-event heterogeneity benefit learning. The coefficient on across-event heterogeneity is estimated at –.834 (standard error = .258). Thus, an increase of 1.99 in discounted across-event heterogeneity reduces the expected accident and incident rate by approximately 1.66 accidents and incidents per 100,000 departures. The across-event heterogeneity effect appears to be slightly stronger than the within-event heterogeneity effect, which means that airlines are able to utilize the additional information inherent in heterogeneity across different accidents and incidents.

As our theory is at the airline level (airlines learn to reduce accidents and incidents), it would be best to conduct analyses on only those causes that are under the direct control of the airline, but it is not clear what is under airline control. Even factors that would normally appear not to be under airline control can be controlled in some situations. For example, weather problems can sometimes be avoided by requesting different flight paths or altitude changes, and passenger-behavior problems are better controlled by some airlines than others. Yet there is probably a continuum of influence and control over various causal factors. To see whether results varied when we excluded accidents caused by factors less under airline control, we did the following. First, we interviewed airline executives to get their input about which factors are more under their control and which are less so. We conducted two sets of analyses in which we eliminated the eleven factors generally perceived to be less under their control (see description of these factors in Appendix A). In the first set of analyses, we eliminated these factors from both the independent and the dependent variables. In the second, we eliminated them from the dependent variable alone, assuming that these factors can be learned from, even though they are not under airlines' direct control. We reanalyzed the data. We found that eliminating factors from the heterogeneity variable not under airlines' direct control does not change the significance of the results of models 2 and 3 of table 4. We also found that eliminating factors from the dependent variable not under direct control of the airline (e.g., eliminating accidents caused by only those factors) does not change the significance of the heterogeneity result and, in fact, seems to strengthen it to some degree. This latter analysis is reported in model 4, for within-event heterogeneity, and model 5, for across-event heterogeneity. For these models, we excluded those 516 accidents and incidents whose causes were relatively less under airline control. As shown in models 4 and 5, the heterogeneity effects are stronger, and model fit improves with the deletion of these cases. Because including all accidents and incidents presents a more conservative test, we took the more conservative approach and used the full sample of accidents in all further analyses.

In models 6 through 9, we analyzed accidents and incidents separately. Model 6 includes accidents alone as the dependent variable and within-accident heterogeneity as an independent variable. Model 7 repeats model 6, but across-accident heterogeneity is substituted for within-accident heterogeneity. Model 8 includes incidents alone as the dependent variable and within-incident heterogeneity as an independent variable. Model 9 repeats model 8, but across-accident heterogeneity is substituted for within-accident heterogeneity. For these analyses, we used discounted number of prior accidents (models 6 and 7) and discounted number of prior incidents (models 8 and 9) as experience controls.

Models 6–9 show that the effect of prior experience heterogeneity on the reduction in subsequent errors applies to both accidents and incidents. More heterogeneous across-accident experience improves subsequent accident rates, and more heterogeneous across-incident experience improves subsequent incident rates. These results apply to both within- and across-accident heterogeneity, though the effect of within-accident heterogeneity on incident rates is only marginally significant ($p < .10$). Overall, then, it seems that heterogeneity benefits airlines, as prior heterogeneous causes result in a reduction in both subsequent accident and incident rates.

Pilot Errors

We conducted several analyses designed to delve more deeply into the question of why heterogeneity leads to a reduction in accident and incident rates. We first ran models testing hypothesis 2, whether heterogeneity leads to a reduction in operator errors. Table 5 reports these analyses. The dependent variable is the number of pilot errors reported for each airline for each year. We used a random-effects

Table 5

Random Effects Negative Binomial Estimates of Number of Errors

Variable	1 Pilot errors	2 Pilot errors	3 Errors type X*	4 Errors type X†
(Log) Airline size	–0.190••	–0.131••	0.011	0.024
	(0.020)	(0.018)	(0.010)	(0.020)
(Log) Industry experience	–2.419••	–2.367••		
	(0.570)	(0.598)		
Number of prior pilot errors	0.022••	0.024••		
	(0.009)	(0.010)		
Number of prior errors of type X			0.007	–0.176••
			(0.005)	(0.021)
Across-event heterogeneity		–1.247••		
		(0.437)		
Constant	3.487••	3.295••	–23.990	1.194•
	(0.618)	(0.621)	(20.788)	(0.542)
N	1781	1781	23467	7382
Wald chi square	12.83	18.94	5.92	72.08
Log likelihood	–2949.37	–2836.29	–6234.33	–185.43

•• $p < .01$; • $p < .05$; two-tailed tests.
* Sample is all accidents and incidents.
† Sample is only single-cause accidents and incidents.

negative binomial model, as the dependent variable is a count variable, and the random-effects specification controls for nonindependence among observations. Model 1 of table 5 presents the results of a set of control variables that might explain the number of pilot errors reported. As model 1 shows, the number of pilot errors decreased significantly with increasing industry experience. Large airlines have fewer pilot errors than small airlines. In model 2, we added the effect of discounted across-accident-and-incident heterogeneity on current pilot errors, and the effect is negative and significant. This shows that airlines with heterogeneous accident and incident experience tend to have fewer pilot errors than airlines with homogeneous accident and incident experiences. In other results (available from the authors), we also found the same effect for within-accident/incident heterogeneity, though the effect is weaker than across-accident heterogeneity.

We had suggested that airlines with heterogeneous experiences may be learning better than those with homogeneous experiences because they are delving deeper into causes and not using a simple blame-the-operator response, and our evidence is consistent with this explanation. An alternative explanation for this result is that pilots themselves are actually learning more from heterogeneous accident and incident experiences and thus make fewer subsequent errors. This argument is a variant on the idea that variance is necessary for learning to occur. We cannot disentangle these two process explanations with the current data structure, though the result showing that across-accident heterogeneity is stronger than within-accident heterogeneity suggests that the process is more likely to be occurring at the airline level, as different pilots are likely to be involved in different accidents and incidents. But whichever is the source of learning—airline or pilot or both—the results support the idea that heterogeneity aids learning.

We conducted additional analyses to see which other types of errors are reduced as a function of prior heterogeneity. Results indicate that some other forms of human error are reduced, including ground-crew error, flight-attendant error, and maintenance errors. In all these cases, prior error heterogeneity tends to decrease subsequent errors of these types but has no effect on other types of human error, such as Air Traffic Control errors or passenger errors. Prior error heterogeneity also reduces subsequent errors that are due to supervisory factors but has no effect on errors due to external factors or faulty equipment (details of all these results are available from the authors). In general, these results show that learning related to human factors that are most associated with the aircraft itself (pilot, ground crew, flight attendant, and maintenance personnel) benefits most from heterogeneity. This suggests that either heterogeneity reduces the tendency to blame human beings, and/or human beings learn from the diverse experiences provided by heterogeneity.

Homogeneity in Causes

We have already shown that heterogeneity, not homogeneity, is beneficial in reducing accident and incident rates, as we found support for H1 and not H4. But we also wanted to directly test hypothesis 3, whether the number of prior causes of a certain type reduces subsequent causes of that type. There are actually many different ways of testing this relationship, and we conducted multiple tests using several different ways of aggregating causes. As a general test, we analyzed the number of each of the 23 causes each airline experienced in a given year, as a function of the number of prior causes of each type experienced by that airline. Thus we were modeling the following:

$$C_xA_yY_z = \text{sum}(C_xA_yY_{z-1\,...\,z-n})$$

where C_x = Cause 1 . . . 23, A_y = Airline 1 . . . 310, and Y_z = Year 1984 . . . 1997.

Model 3 of table 5 runs this analysis on the full sample of accidents and incidents. We used a random-effects negative binomial model, as the dependent variable is a count variable, and the random-effects specification controls for nonindependence among observations. Model 3 shows the number of prior causes of a given type does not reduce the number of subsequent causes of that same type. When we ran this analysis on only single-cause accidents (model 4), we found that the number of prior causes of a given type does reduce the number of subsequent causes of that same type, but this effect is restricted to one-cause accidents. In other models, we found that subsequent accidents with multiple causes are not reduced. Also, double-cause and triple-cause accidents do not reduce the rate of subsequent accidents with those same double or triple causes. Thus, only the most extreme form of homogeneity, e.g., a single cause, is beneficial for learning, which provides support for H3 only in limited cases. An alternative explanation is that a relationship between more accidents of a given type in the current period and fewer accidents of that type in future periods is prone to regression to the mean, but we did not find any support for H3 except in the narrow case of single-cause accidents. If regression to the mean were operating, there is no reason why it would operate only on single-cause accidents. Therefore, we think regression to the mean is not a likely alternative explanation for our rather narrow support for this hypothesis.

Generalism/Specialism and Learning Results

The analyses in table 6 assess whether generalists and specialists learn differently from heterogeneous accident experiences. The dependent variable is the accident and incident rate per 100,000 departures. Model 1 is a base model, including all the control variables, within-event heterogeneity, and level of specialism. In model 2, we added the interaction of discounted within-event heterogeneity with the variable indicating level of specialism, but the interaction of specialism and within-event heterogeneity is not significant. In model 3, we used discounted across-accident heterogeneity and, in model 4, the interaction of specialism and across-event heterogeneity. The interaction of specialism and across-event

Table 6

Random Variable Effects Estimates of Carrier Accident and Incident Rates (Events per 100,000 Departures)*

	1	2	3	4
(Log) Airline size	0.131	0.134	0.095	0.077
	(0.128)	(0.128)	(0.128)	(0.128)
Airline age	–0.263••	–0.261••	–0.276••	–0.232••
	(0.082)	(0.082)	(0.083)	(0.084)
Own prior events	–0.044•	0.043•	–0.013	–0.018
	(0.019)	(0.020)	(0.020)	(0.020)
Airline profitability (ROA)	0.005	0.004	–0.017	–0.020
	(0.047)	(0.047)	(0.047)	(0.047)
(Log) Industry experience	1.170	1.230	1.200	1.765•
	(0.648)	(0.087)	(0.691)	(0.691)
Event severity	2.220••	2.230••	2.210••	2.180••
	(0.375)	(0.375)	(0.375)	(0.375)
Within-event heterogeneity	–4.176••	–4.640••		
	(1.173)	(1.689)		
Across-event heterogeneity			–1.130••	–2.280••
			(0.281)	(0.557)
Level of specialism	1.430	1.170	2.452••	0.986
	(0.878)	(1.120)	(0.949)	(1.130)
Specialism x within-event het.		–0.562		
		(1.480)		
Specialism x across-event het.				–0.731•
				(0.306)
Hazard rate	0.009	0.009	0.010	0.011
	(0.010)	(0.008)	(0.009)	(0.012)
Constant	–11.808•	–11.342•	–14.756••	–12.640•
	(4.680)	(4.840)	(4.800)	(4.860)
N (total firm-years)	1781	1781	1781	1781
Log likelihood	–5147.28	–5147.20	–5158.92	–5154.07
Likelihood ratio		0.16		9.70••
D.f. (vs. model no.)		1 (M1)		1 (M3)

	5	6	7	8
(Log) Airline size	–0.671•	–0.762••	–0.117••	–0.076
	(0.277)	(0.289)	(0.162)	(0.191)
Airline age	–0.081	–0.687••	0.021	–0.903••
	(0.116)	(0.245)	(0.105)	(0.338)
Own prior events	–0.010	0.023	–0.911••	–0.989••
	(0.021)	(0.025)	(0.187)	(0.210)
Airline profitability (ROA)	–0.084	0.024	0.234	0.398
	(0.032)	(0.033)	(0.591)	(0.672)
(Log) Industry experience	1.290	1.920	–2.850••	–7.450••
	(0.694)	(1.180)	(0.678)	(1.400)
Event severity	–0.878•	–1.050••	3.520••	2.810••
	(0.428)	(0.432)	(0.549)	(0.559)
Within-event heterogeneity				
Across-event heterogeneity	–0.469•	–0.872•	–0.192	–0.169
	(0.230)	(0.278)	(0.781)	(0.791)
Level of specialism				
Specialism x within-event het.				
Specialism x across-event het.				
Hazard rate	0.011	0.013	0.014	0.008
	(0.010)	(0.009)	(0.008)	(0.010)
Constant	2.240	2.340	–25.190••	–57.310••
	(6.820)	(6.450)	(6.790)	(11.000)
N (total firm-years)	855	855	926	926
Log likelihood	–2079.42	0.08†	–2915.85	0.17†

•• $p < .01$; • $p < .05$; two-tailed tests.

* Models 5 and 6 analyze only specialists, and models 7 and 8 analyze only generalists. Models 6 and 8 are fixed-effects models.

† R-square (overall).

heterogeneity in model 4 is negative and significant. This indicates that specialists with a high degree of heterogeneity in prior experience learn more (i.e., their accident and incident rate is reduced), relative to other firms, but this effect only holds for across-event heterogeneity. Thus, H5, which predicted that specialists would learn more from heterogeneous experience than generalists, is supported, but only for across-event heterogeneity.

To explore this effect in more detail, we split the sample into generalist and specialist airlines, initially using a median split as a cutoff. Using the median split, the generalist airlines had a total of 734 accidents and incidents in our sample period, with an average rate of 2.82 events per 100,000 departures. The specialist airlines had a total of 612 accidents and incidents, with an average rate of 7.42 events per 100,000 departures. The generalist airlines had more heterogeneous accidents—their across-event heterogeneity averaged 3.63, while specialists were only .692. This difference in heterogeneity, however, was effectively controlled for in model 4 by including the main effect of specialism, so the interaction effect in model 4 was found independent of any difference in accident heterogeneity across generalists and specialists.

For illustrative purposes, we ran separate models on the generalist and specialist subsamples. The sample for model 5 was used on specialists alone, and the results show that prior error heterogeneity reduces the accident and incident rate for specialist airlines. Model 6 is the same as model 5, except that we used a fixed-effects estimator as a more stringent test of whether unobserved heterogeneity was affecting our results. Model 6 shows that the effect of prior heterogeneity on specialists remains, and is in fact stronger, with the fixed-effect specification. Thus, we are confident that specialists are learning from across-event heterogeneity. It also appears that within the specialist categories, those specialists that are larger have lower accident rates than those that are smaller, as the effect of size is negative and significant. It also appears that specialists with more severe prior events tend to have fewer subsequent events, an effect opposite that of the generalists.

Model 7, which includes only generalists, shows no effect of prior experience heterogeneity for generalists. But model 7 also shows that the effect of industry experience (cumulative miles flown) is significant, indicating that generalists are learning from outside factors that change over time, in a way that specialists are not (models 5 and 6). Model 8 is the same as model 7, except that we use the fixed-effects specification. As with the specialist results, these generalist effects remain when we use the more stringent fixed-effects test.

We also tried several different methods of splitting the generalism/specialism variable to see when the learning effects hold. Results showed that the effect of prior across-event heterogeneity in causes applies to all airlines except the large major carriers, which are the most extreme generalists. The large major carriers are the only airlines that do not learn from prior experience heterogeneity and who appear to learn from alternate sources (e.g., industry experience). (Details of these analyses are available from the authors.)

Thus, it appears that the factors affecting learning (reduction in accident and incident rates) are different for generalists and specialists. Specialists appear to learn from heterogeneity in prior accidents and incidents. Generalists, in contrast, do not learn from prior experience heterogeneity, though they do learn from their own accumulated experience. Instead, they learn from outside factors. These outside factors are captured in the industry-miles variable and are likely to include such things as general technological improvements, learning by the industry as a whole, and changes in regulation that result in reduced accident and incident rates.

DISCUSSION

The results of this study contribute to our understanding of variation in how organizations learn from their prior error experience. It appears that, in general, experiences with heterogeneous causes produces more learning than experience with homogeneous causes. This is likely to be due to heterogeneity producing a deeper, broader search for causality that avoids simple explanations like blame the pilot. There is one specialized case in which homogeneity benefits learning: in simple one-cause accidents. In anything more complex, the greater the level of complexity, the greater the learning. In addition, generalist airlines appear to learn from different sources of experience than specialists do. These results have interesting implications for theories of learning and organizational form and for work on organizational errors, as well as for understanding the increasing complexity of the causes of accidents.

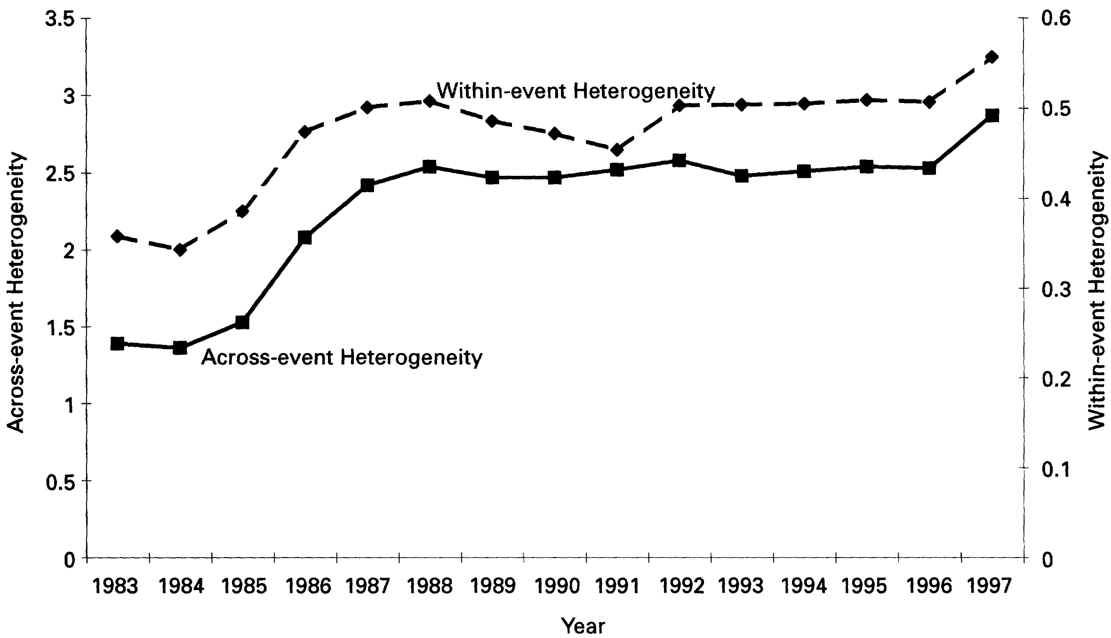
Learning Theories

Organizational learning theorists represent learning as history-dependent, routine-based systems that adapt incrementally to past experience. Consistent with this, we found that airlines learn incrementally from their own accident experience and from the heterogeneity of causes in their own prior accident and incident experiences. We also found that the benefits of learning from heterogeneity seem to apply primarily to specialist organizations, as only their accident and incident rates are reduced with heterogeneity in prior causes. Our results also show that only generalists learn from accumulated industry experience. The accident and incident rate of generalist airlines (but not specialists) decreases over time and with the accumulation of others' experience. Thus, generalists seem to learn more from the experience of the industry as a whole than specialists do. There are several possible reasons why this might be so.

First, it may be that generalists experience more pressure to respond to external experience than specialists do because regulatory agencies put more pressure on generalists than they do on specialists. This response, evidenced by a reduction in accident rates, comes as the result of external pressures (the FAA) rather than internal experience. It may also be that generalist carriers are generally more sensitive to the accident and incident experiences and responses to them experienced by the industry as a whole than the specialist airlines are. They probably have better access to industry-level information than specialist airlines do and are also more likely to have specialized internal departments charged with analyzing industry data than specialists are.

Second, an additional factor associated with industry-level experience is an increase in the heterogeneity of accident experience over time. Figure 2 graphs accident and incident heterogeneity for each year covered by our sample and

Figure 2. Heterogeneity in causes of accidents and incidents over time.



shows that accident heterogeneity has been increasing over time.[4] Thus, it may be that generalists, in addition to paying attention to information on accidents by the system as a whole, learn more from system-level heterogeneity than from their own heterogeneity in experience. Future research designed to test these different explanations for the differences between generalist and specialist airlines would be useful.

There now seems to be a fair amount of convergence across studies on the idea that the organizational form affects learning. As noted earlier, Barnett, Greve, and Park (1994) and Ingram and Baum (1997) showed that generalist organizations learn differently from specialist organizations. In general, what these studies show is that generalist firms do not learn as well from their own experience as specialists do, but these studies did not separate the quantity of experience from the characteristics of that experience. We separated quantity (e.g., number of prior accidents) from characteristics (heterogeneity of causes of prior accidents) and, in doing so, found that while both generalists and specialists tend to learn from the amount of experience, they learn differently from

4
A formal analysis of this effect shows that year is positively related to heterogeneity [(F1,1780) = 33.49, $p < .01$].

the characteristics of that experience. Thus, our findings, along with those of others who have studied the link between generalism and specialism and learning, are promising in that they illuminate the tensions between strategic choice and learning, in which the choice of form affects an organization's ability to learn from certain types of experience.

In addition to providing evidence on the relationship between learning and organizational form, our study contributes to learning theories by developing and showing the effects of a variable that is not yet well studied: the heterogeneity or homogeneity of prior experience. Beckman and Haunschild (2002) showed that experience heterogeneity affects corporate acquisitions, but their study looked at experience of others, not a firm's own heterogeneity or homogeneity of experience. Our results show that the characteristics of a firm's own past experience are important, at least for specialist firms. Thus, our results contribute to one goal in the learning literature: that of understanding the sources of variation in learning rates across different organizations (Argote, 1999). Our results show that organizations with experience with heterogeneous causes will learn more than organizations with experience with homogeneous causes. In addition, unlike Beckman and Haunschild's study, in which heterogeneity was created by variance in outcomes (i.e., successful and unsuccessful acquisitions), in this study, airlines are learning without being able to directly compare the causes of accidents (failures) and the causes of no-accidents (successes). That is, it is quite difficult for anyone to observe the presence or absence of causes present in those flights that did not have an accident. Instead, learning in this case is driven by the salience of failures, which is a powerful driver of learning (Sitkin, 1992), and not a comparison of success and failure.

Finally, we also contribute to the small but growing body of work on how and when organizations can learn from failure (Sitkin, 1992; Kim and Miner, 2000). Accidents and incidents are a form of failure, and we showed that failure can be learned from, especially when it has heterogeneous causes. One issue with all research on failure is that failures are generally rare events. In this study, they are especially rare (on an absolute basis) for the specialist airlines. The specialist airlines have fewer absolute numbers of accidents and incidents than the generalist airlines (though, as noted earlier, their failure rate per departure is actually higher). Yet our evidence shows that such failures can lead to learning. This is in keeping with studies showing that failure is an important stimulant to organizational learning, in part because failures are highly salient and difficult to hide (Sitkin, 1992; Greve, 1998; Miner et al., 1999; Kim and Miner, 2000). What is interesting is that, in our study, the specialists are learning from their own failures, but the generalists seem to be learning from the failures of others, through the accumulated experience with heterogeneity in causes in their industry. Thus, both are learning from relatively rare events, but the source of these events is different. Future research into exactly why generalists and specialists learn differently from these different sources of failure information would be useful. Research could also profit from contrasting learning from rare, salient events, like failures, and learning from common, nonsalient events, such as cumulative production experience. Much of the learning curve research to date has focused on common, nonsalient events, and yet this study showed that rare events also promote learning. Finally, research could profit from investigating whether all failures affect learning, possibly by comparing across different industries and/or different types of failure events (e.g., accidents, financial crises, organizational failures). As noted earlier, it may be that the airline industry has industry-specific features that promote learning from failure, features not present in other industries (Perrow, 1984). Such differences are a fruitful place to begin to deepen our understanding of the conditions under which failures can be a source of learning (cf. Ocasio, 1995).

One strength of our study is that we analyzed a process-oriented explanation for the results, in the context of a large-scale study looking at performance improvement. Much of the work in the area of organizational learning can be classified according to whether its primary focus is on learning as an outcome or learning as a process. Some people study learning as an outcome, one that can be seen in performance improvement or a change in organizational routines (e.g., Ingram and Baum, 1997; Baum and Ingram, 1998; Kim and Miner, 2000; also see Argote, 1999, for a summary of earlier work). Some people study learning as a process, not tied to a particular outcome. For example, there is much work on learning problems, such as that organizations tend to search locally rather than broadly (Levinthal and March, 1993), or how improvisation activities affect knowledge acquisition in organizations (Miner, Bassoff, and Moorman, 2001). Our study is primarily outcome-oriented in that we studied performance improvements (reduced error rates) but also theorized and tested some process-level explanations for results (e.g., the pilot-error analyses). Combining outcome and process in a single study is rarely done but is important for supporting the idea that learning processes are at work and for contributing to a more detailed contextual understanding of organizational learning.

Like all studies, ours has some limitations that suggest directions for future research in this area. One limitation is that, while we made some efforts to explore the processes underlying the heterogeneity effects, process factors could only be explored in a limited way. Questions about exactly how heterogeneous accident causes reduce subsequent accidents are still relatively unexplored, and thus future research investigating the process underlying this effect would be valuable. For example, airlines' responses to their accidents could be tracked over time, and detailed data on responses to heterogeneous and homogeneous accident causes could be collected. In addition, our primary data sources, airline accident reports, are subject to errors and biases of their own (Tamuz, 1987). We don't always know, for example, whether there are changes in accident causes or just changes in accident reporting. Our data show a decline in pilot errors over time, but is this due to an actual decline in such errors or to a decline in the tendency to blame the pilot in a situation in which causes are ambiguous? While such changes in reporting are not likely to bias our conclusions (i.e., it is not likely that investigators will report more heterogeneous causes for the accidents and incidents of those airlines that subsequently experience a decrease in accident and incident rates), it is still important to try to investigate the possible reporting biases in aviation errors.

Organizational Errors

The findings of this study take on added importance because they occur in the context of events of great consequence. Accidents and incidents (as well as other disasters) need to be investigated because they have a significant impact on organizations and society (Perrow, 1984). While many important studies of accidents and incidents have been done, they tend to be case studies of individual accidents or disasters. Accidents and incidents have also served as a foundation for the development of interesting theories, such as Perrow's (1984) normal accidents theory and Reason's (1997) discussions of the latent conditions underlying many accidents. There are, however, very few large-scale empirical tests of theories of organizational accidents and errors. Seeking to address this gap, we found that such errors are affected by the characteristics of prior experience with errors, specifically, the heterogeneity of their causes.

Our results are consistent with Weick's (1987) discussion of the benefits of requisite variety for the reduction of aviation errors. Weick noted the benefits of increasing system variety for reducing errors and proposed that variety can be increased through such mechanisms as promoting individual diversity, a focus on face-to-face interaction styles, reducing rigid bureaucratic controls, and promoting individual discretion over decisions. Our results are consistent with Weick (1987) in that we found that increasing system diversity (through the diverse information available from experience with heterogeneous causes) is valuable and showed the enhanced benefits of diverse information in specialist organizations, which are likely to have fewer rigid bureaucratic controls and more individual discretion than generalists.

More generally, our results are also important for the study of other organizational errors and their causes. Some work in economics has investigated the effects on errors of such factors as organizational profitability (Rose, 1990) and deregulation, but there has been little work in organizational theory on the causes or consequences of errors. Our study contributes to this literature by showing the important effects of organizational processes such as learning efforts and system constraints on the management of errors. Empirical work on the learning curve has focused on productivity and firm survival, not reduction in errors. Yet examining reduced error frequency and/or severity is an important way to study learning, because the organizational factors that result in error reduction are as yet relatively unexplored. Also, knowing more about how errors are reduced has important practical implications for both organizations and society.

Increased Complexity in Accident Causes over Time

Our results also have interesting implications for the impact of the heterogeneity of experience over time. We found that heterogeneity in the causes of accidents has been increasing over time, which may contribute to explaining the increase in the number of airline accidents whose causes remain a mystery (Paztor and Matthews, 1999). One possibility is that the simple factors that caused accidents have been fixed, so what is left in the system are the more complex interactions. Thus, pilots no longer fly into thunderstorms. Turbulence is now much more predictable through improvements in weather technology. So, when an accident occurs now, it occurs as the result of several factors whose interaction is quite difficult to predict.

A second reason why the heterogeneity in accident causes may be increasing over time is that the underlying technology has become more complex, so the causes of errors become more complex. Thus, there is isomorphism between error complexity and technical complexity. A third possible reason for the increase might be the complexity of rules. The idea here is that errors create new rules. The U.S. airline system is heavily regulated, and accidents tend to generate a lot of scrutiny. New rules are often put in place in an attempt to prevent future accidents, with the result that, over time, rules proliferate. The problem with a proliferation of rules is that they can lead to problems and may actually increase accident rates. As Reason (1997) noted, rules designed to reduce the opportunities for one kind of error can relocate errors to other parts of the system. These parts are likely to be more opaque in the sense that they will contribute to more errors based on the complex interaction of multiple underlying factors. The complexity of rules themselves may lead to a corresponding complexity in errors.

A fourth possible cause for the increase in the complexity of accident causes over time might be that reporting complexity has increased. It could be that accidents are not inherently more complex now than previously but, rather, that investigators are more open to looking for multiple causes and complex interactions of factors than they used to be. The dissemination of knowledge about how accidents can have multiple, complex, underlying factors (e.g., Perrow, 1984; Reason, 1997; Vaughan, 1999) may be affecting the accident investigators themselves. Future research testing these different causes is needed and can contribute to our understanding of the evolution of knowledge management in technical systems.

Our findings about how airlines learn from complexity suggest important features of the learning process. Beyond airlines, the research presented here points to the importance of investigating learning processes and the interaction of learning and organizational form in other areas. Organizational learning from critical events other than airline accidents, such as industrial accidents, financial crises, and medical errors might similarly be affected by prior experience heterogeneity. Or learning from less drastic errors such as product recalls may be similarly affected by heterogeneity and organizational form. The systemwide complexity of underlying causes may also increase over time for these types of events. The underlying processes of learning and knowledge management in these systems are clearly important and deserving of further study.

REFERENCES

Ancona, D., and D. Caldwell
1992 "Demography and design: Predictors of new product team performance." Organization Science, 3: 321–341.

Argote, L.
1999 Organizational Learning. London: Kluwer Academic.

Argote, L., S. Beckman, and D. Epple
1990 "The persistence and transfer of learning in industrial settings." Management Science, 36: 140–154.

Argote, L., and E. Darr
2000 "Repositories of knowledge in franchise organizations: Individual, structural and technological." In G. Dosi, R. Nelson, and S. Winter (eds.), Nature and Dynamics of Organizational Capabilities: 68–94. Oxford: Oxford University Press.

Argote, L., and D. Epple
1990 "Learning curves in manufacturing." Science, 247: 920–924.

Barnett, W., H. Greve, and D. Park
1994 "An evolutionary model of organizational performance." Strategic Management Journal, 15: 11–28.

Baum, J., and P. Ingram
1998 "Survival-enhancing learning in the Manhattan hotel industry, 1898–1980." Management Science, 44: 996–1016.

Beckman, C., and P. Haunschild
2002 "Network learning: The effects of partners' heterogeneity of experience on corporate acquisitions." Administrative Science Quarterly, 47: 92–124.

Carroll, G. R.
1985 "Concentration and specialization: Dynamics of niche width in populations of organizations." American Journal of Sociology, 94: 723–753.

Carroll, G. R., and A. Swaminathan
1992 "The organizational ecology of strategic groups in the American brewing industry from 1975 to 1990." Industrial and Corporate Change, 1: 65–97.

Carroll, J.
1995 "Incident reviews in high-hazard industries: Sensemaking and learning." Industrial and Environmental Crisis Quarterly, 9: 175–198.

Cohen, J.
1960 "A coefficient of agreement for nominal scales." Educational and Psychological Measurement, 20: 37–46.

Cyert, R. M., and J. G. March
1963 A Behavioral Theory of the Firm. Englewood Cliffs, NJ: Prentice-Hall.

Darr, E., L. Argote, and D. Epple
1995 "The acquisition, transfer and depreciation of knowledge in service organizations: Productivity in franchises." Management Science, 42: 1750–1762.

Dobrev, S., T. Kim, and M. Hannan
2001 "Dynamics of niche width and resource partitioning." American Journal of Sociology, 106: 1299–1337.

Donoghue, J. A.
1998 "Changing the future: Political winds within and outside the airline industry complicate efforts to avoid an accident-filled future." Air Transport World, 35 (5): 35–46.

Fiske, S., and S. Taylor
1984 Social Cognition. Reading, MA: Addison-Wesley.

Freeman, J., and M. Hannan
1983 "Niche width and the dynamics of organizational populations." American Journal of Sociology, 88: 1116–1145.

Fromm, G.
1968 "Aviation safety." Law and Contemporary Problems, 33: 590–618.

Greve, H. R.
1998 "Performance, aspirations and risky organizational change." Administrative Science Quarterly, 43: 58–68.

Hannan, M., and J. Freeman
1984 "Structural inertia and organizational change." American Sociological Review, 49: 149–164.

Heckman, J.
1979 "Sample selection bias as a specification error." Econometrica, 47: 153–161.

Ingram, P., and J. Baum
1997 "Opportunity and constraint: Organizations' learning from the operating and competitive experience of industries." Strategic Management Journal, 18: 75–98.

Irwin, D. A., and P. J. Klenow
1994 "Learning-by-doing spillovers in the semiconductor industry." Journal of Political Economy, 102: 1200–1227.

Jehn, K. A., G. B. Northcraft, and M. A. Neale
1999 "Why differences make a difference: A field study of diversity, conflict, and performance in workgroups." Administrative Science Quarterly, 44: 741–763.

Joskow, P. L., and N. L. Rose
1985 "The effects of technological change, experience, and environmental regulation on the construction cost of coal-burning generating units." RAND Journal of Economics, 16: 1–27.

Kim, J., and A. S. Miner
2000 "Crash test without dummies: A longitudinal study of interorganizational learning from failure experience in the U.S. commercial banking industry, 1984–1998." CD ROM, Best Paper Proceedings, Academy of Management.

Lapre, M. A., A. S. Mukherjee, and L. Van Wassenhove
2000 "Behind the learning curve: Linking learning activities to waste reduction." Management Science, 46: 597–611.

Levinthal, D., and J. G. March
1993 "The myopia of learning." Strategic Management Journal, 14: 95–112.

March, J. G., and J. P. Olsen
1988 "The uncertainty of the past: Organizational learning under ambiguity." In J. G. March (ed.), Decisions and Organizations: 335–358. Oxford, England: Basil Blackwell.

March, J. G., and H. Simon
1958 Organizations. New York: Wiley.

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

Marcus, A., and M. Nichols
1999 "On the edge: Heeding the warnings of unusual events." Organization Science, 10: 482–499.

Mason, C. H., and W. D. Perreault
1991 "Collinearity, power and interpretation of multiple regression analysis." Journal of Marketing Research, 28: 268–280.

Miner, A., P. Bassoff, and C. Moorman
2001 "Organizational improvisation and learning: A field study." Administrative Science Quarterly, 46: 304–337.

Miner, A., and S. Mezias
1996 "Ugly duckling no more: Pasts and futures of organizational learning research." Organization Science, 7: 88–99.

Miner, A. S., I. Holzinger, J. Kim, and P. R. Haunschild
1999 "Fruits of failure: Organizational failure and population level learning." In A. S. Miner and P. Anderson (eds.), Advances in Strategic Management, 16: 187–220. Stamford, CT: JAI Press.

Morris, M. W., and P. C. Moore
2000 "The lessons we (don't) learn: Counterfactual thinking and organizational accountability after a close call." Administrative Science Quarterly, 45: 737–765.

Morris, P. A., and R. H. Engelken
1973 "Safety experience in the operation of nuclear power plants." In International Atomic Energy Agency Principles and Standards of Reactor Safety: IAEA Symposium Proceedings, Vienna: 429–446.

Nance, J. J.
1986 Blind Trust. New York: William Morrow.

Nisbett, R. E., and L. Ross
1980 Human Inference: Strategies and Shortcomings of Social Judgment. Englewood Cliffs, NJ: Prentice-Hall.

NTSB (National Transportation Safety Board)
2001 Code of Federal Regulations, Title 49, vol. 5, chap. 8, part 830, sect. 830.2 pp. 1195–1196.

Ocasio, W.
1995 "The enactment of economic adversity: A reconciliation of theories of failure-induced change and threat-rigidity." In L. L. Cummings and B. M. Staw (eds.), Research in Organizational Behavior, 17: 287–331. Greenwich, CT: JAI Press.

Pasztor, A., and A. W. Mathews
1999 "Why more plane-crash probes end in doubt." Wall Street Journal, March 22, p. B1.

Perrow, C.
1984 Normal Accidents. New York: Basic Books.
1994 "Accidents in high-risk systems." Technology Studies, 1: 1–19.

Phillips, D.
2000 "French Concordes to stay grounded as crash probe continues." Washington Post, August 4, p. A30.

Reason, J.
1987 "The Chernobyl errors." Bulletin of the British Psychological Society, 40: 201–206.
1997 Managing the Risks of Organizational Accidents. Sydney, Australia: Ashgate.

Roberts, K.
1991 "Some characteristics of one type of high reliability organization." Organization Science, 2: 160–176.
1993 New Challenges to Understanding Organizations. New York: Macmillan.

Rose, N. L.
1990 "Profitability and product quality: Economic determinants of airline safety performance." Journal of Political Economy, 98: 944–964.

Sagan, S.
1993 The Limits of Safety: Organizations, Accidents, and Nuclear Weapons. Princeton, NJ: Princeton University Press.

Seidel, M.
1997 "Competitive realignment in the airline industry: A dynamic analysis of generalist and specialist organizations under different power network structures." Unpublished doctoral dissertation, University of California, Berkeley.

Sitkin, S. B.
1992 "Learning through failure: The strategy of small losses." In B. M. Staw and L. L. Cummings (eds.), Research in Organizational Behavior, 14: 231–266. Greenwich, CT: JAI Press.

Swaminathan, A.
1995 "The proliferation of specialist organizations in the American wine industry." Administrative Science Quarterly, 40: 653–680.

Tamuz, M.
1987 "The impact of computer surveillance on air safety reporting." Columbia Journal of World Business, Spring: 69–77.

Teachman, J. D.
1980 "Analysis of population diversity." Sociological Methods and Research, 8: 341–362.

Torres, J. C.
1995 "The dynamics of the UK motor industry: An ecological analysis." Unpublished doctoral thesis, Stanford University.

Vaughan, D.
1999 The *Challenger* Launch Decision: Risky Technology, Culture, and Deviance at NASA. Chicago: University of Chicago Press.

Weick, K.E.
1987 "Organizational culture as a source of high reliability." California Management Review, Winter: 112–127.

Zimmerman, M. B.
1982 "Learning effects and the commercialization of new energy technologies: The case of nuclear power." Bell Journal of Economics, 13: 297–310.

APPENDIX A: Descriptions of Accident Cause Codes

1. Weather—turbulence*
2. Weather—other*
3. Problematic ground conditions (not under control of facilities)*
4. Pilot error
5. Flight crew error
6. Flight attendant error
7. Ground crew error
8. Equipment malfunction
9. Maintenance error
10. Manufacturer error*
11. FAA supervision/procedures inadequate*
12. Airline supervision/procedures inadequate
13. Airline failure to incorporate correction
14. Air traffic control error*
15. Inadequate weather assistance*
16. Passenger error*
17. Pilot of other aircraft—error*
18. Inadequate group coordination
19. Birds*
20. Facilities problem (e.g., runway maintenance)*
21. Flight into terrain
22. Unknown, undetermined
23. Other—nonclassifiable

* These are causes seen to be relatively less under the airline's control than other causes in the above list (based on interviews with airline executives) and were excluded from the dependent variable in the analysis reported in table 4, models 4 and 5.

APPENDIX B: Sample NTSB Accident Report

General Information

Data Source:	NTSB AVIATION ACCIDENT/INCIDENT DATABASE
Report Number:	ATL86LA090
Local Date:	03/12/1986
Local Time:	717 CST
State:	TN
City:	MEMPHIS
Airport Name:	MEMPHIS INTL
Airport Id:	MEM
Event Type:	ACCIDENT
Injury Severity:	SERIOUS
Report Status:	FINAL
Mid Air Collision:	NO

Operations Information

Category of Operation:	SCHEDULED, PART 121
Aircraft Type:	AIRPLANE
Aircraft Damage:	NONE
Phase of Flight:	530 CLIMB
Aircraft Make/Model:	BOEING B-727-22
Operator Name:	AMERICAN AIRLINES, INC.
Operator:	AMERICAN AIRLINES, INC—AALA

Narrative

AMERICAN AIRLINES FLIGHT 502 EXPERIENCED A FIRE WARNING LIGHT ILLUMINATION ON NO. 1 ENGINE AS THE AIRCRAFT WAS CLIMBING THROUGH 6000 FEET. THE PILOT DECLARED AN EMERGENCY AND RETURNED TO MEMPHIS. FOLLOWING THE LANDING THE PILOT ORDERED A PASSENGER EMERGENCY EVACUATION. DURING THE EVACUATION THE REAR SLIDE DEFLATED AFTER THE SLIDE MATERIAL WAS PUNCTURED. THE EXAMINATION OF THE MATERIAL DISCLOSED THAT THE PUNCTURE WAS THE RESULT OF A SHOE HEEL. THE EXAMINATION OF THE FIRE WARNING SYSTEM DISCLOSED THAT A DUCT CONNECTOR IN THE SYSTEM HAD FAILED. THE INJURY RECEIVED BY THE INJURED PASSENGER RESULTED FROM JUMPING OFF THE LEFT WING SURFACE DURING THE EVACUATION PHASE.

Sequence of Events

Occurrence #: 1 430 MISCELLANEOUS/OTHER
Phase of Operation: 530 CLIMB

Findings

Subject—Modifier—Personnel Cause/Factor

1a. 12401(S)—1135(M) Factor
FIRE WARNING SYSTEM, POWERPLANT—FAILURE, TOTAL

2a. 17106(S)—1135(M)—Factor
MISC EQPT/FURNISHINGS, SLIDES—FAILURE, TOTAL

4b. 24545(S)—3124(M)—4127(P) Cause
EMERGENCY PROCEDURE—NOT FOLLOWED—PASSENGER

Aircraft Information

Number of Seats:	150
Type of Operation:	14 CFR 121
Domestic/International:	DOMESTIC
Passenger/Cargo:	PASSENGER
Registration Number:	877AA
Air Carrier Operating Certificates:	FLAG CARRIER/DOMESTIC (121)
Aircraft Fire:	NONE

	Injuries			
	Fatal	Serious	Minor	None
Crew	0	0	0	7
Pass	0	1	1	61
Other	0	0	0	0
Invlvd	0	1	1	68

Landing Gear:	TRICYCLE-RETRACTABLE
Certificated Maximum Gross Weight:	191500
Engine Make:	P&W
Engine Model:	JT8D-9A
Number of Engines:	3
Engine Type:	TURBO FAN

Environment/Location Information

Basic Weather Conditions:	VISUAL METEOROLOGICAL CONDITIONS
Wind Direction (deg):	0
Wind Speed (knots):	0
Visibility (sm):	7
Visibility RVR (ft):	0
Visibility RVV (sm):	0
Cloud Height Above Ground Level (ft):	0
Visibility Restrictions:	NONE
Precipitation Type:	NONE
Light Condition:	DAYLIGHT
Departure City:	MEMPHIS
Departure State:	TN
Destination City:	NASHVILLE
Destination State:	TN
Flight Plan Filed:	INSTRUMENT FLIGHT RULES (IFR)
ATC Clearance:	IFR
VFR Approach/Landing:	PRECAUTIONARY LANDING
Event Location:	ON AIRPORT

Pilot-in-Command

Certificates:	AIRLINE TRANSPORT		
Ratings:			
Plane:	SINGLE ENGINE LAND, MULTIENGINE LAND		
Non-Plane:	NONE		
Instrument:	AIRPLANE		
Had Current BFR:	YES		
Months Since Last BFR:	3		
Medical Certificate:	CLASS 1		
Medical Certificate Validity:	VALID MEDICAL—NO WAIVERS OR LIMITATIONS		
Flight Time (Hours)			
Total:	11670	Last 24 Hrs:	8
Make/Model:	5450	Last 30 Days:	0
Instrument:	0	Last 90 Days:	117
Multi-Engine:	0	Rotorcraft:	