We examined the potential vulnerabilities of pilots flying a mixed fleet of two different aircraft types. A “worst case” scenario was evaluated in which a pilot, flying one type exclusively, would need to fly the different type, after 6 months without any recurrency training on the latter. These circumstances invite negative transfer of habits in the “old” aircraft, to performance in the “new” aircraft. Documents of both aircraft were evaluated to establish those aspects of design and procedures differences that invite such negative transfer; a list of 36 such “vulnerabilities” were identified. Then 40 active commercial airline pilots from a US carrier participated in an evaluation of such negative transfer between two different types within the fleet. The sample was divided into 2 groups each of which normally flew one of the types and not the other. After training on the “new” type, each pilot returned to either 3 or 6 months of flying exclusively with their “old” type, and then returned for simulator evaluations on the “new” type that were targeted to reveal the 36 vulnerabilities. Even with power-sensitive statistical analyses, only slight evidence for negative transfer was found. Those areas where such transfer did emerge were targeted for recommendations of either procedural harmonization or minor design changes.

**INTRODUCTION**

Mixed fleet flying (MFF) is defined as the operation of two or more variants or types of airplanes by the same pool of pilots. The training program for these pilots is usually as one full program that teaches the pilots both airplanes and their similarities and differences. Under an FAA-approved mixed-fleet flying program requirements are given to the airline for the maximum amount of time that a pilot can go without flying one of the types or variants. This currency period is typically 3 months, which means that under a “worst case scenario” a MFF trained pilot may fly one aircraft for just under three months, and then, unexpectedly, be called upon to fly the other aircraft the next day. To qualify for approval of such program, it is required that careful scrutiny be given to the possibility of negative transfer between the two aircraft (Holding, 1987; Lyall, 1990). Such negative transfer is experienced in every day life, as, for example, we switch from one keyboard to another, where critical function keys are located in different places. Classic analysis of transfer (Holding, 1987) reveals that the “red flag” inviting negative transfer results when the similar displays and circumstances between the “old” and “new” system, and also similar, but not identical actions, whereby the latter have very different consequences in the old from the new systems (Braune, 1989; Wickens & Hollands, 2000).

The goal of the current study was to examine the extent to which there may be safety consequences due to negative transfer related to the mixing of two aircraft, the Boeing 767-400, and the Boeing 777. Selective evidence for negative transfer (i.e., in some skills and procedures) would signal the need for specific steps to be taken prior to adopting a safe MFF regime, including selective training on those vulnerable procedures, efforts to harmonize the procedures, or possibly selective re-design. The study was conducted in two phases. The first phase was to carefully analyze the differences in procedures between the two aircraft types, to identify where particular vulnerabilities to negative transfer might occur. The second phase was a longitudinal study that empirically examined
transfer between aircraft under a “worst case” scenario.

**Phase 1: Differences Vulnerability Study**

A full set of pilot tasks was used as the foundation of the analysis. The method was to document for each task the situations and actions required on each airplane then analyze the similarities and differences between the airplanes identifying potential vulnerabilities for pilot performance that could result from mixing the airplanes. The documentation began by gathering information from the flight and training manuals for both airplanes. These manuals include normal and non-normal procedures and checklists (approximately 1200 pages of material for each airplane). The information from the manuals was then expanded using interviews with pilots and instructors qualified on one or both of the airplanes and, if necessary, conducting directed sessions in simulators and training devices to understand how the pilots may need to interact differently in the two airplanes. The result of the analysis was a list of possible pilot performance vulnerabilities. A vulnerability was identified if the required performance to accomplish a task was different in the two airplanes. The vulnerabilities were categorized as possibly major and needing to be verified in the longitudinal study to measure the criticality if the behavior could result in a safety consequence. Minor vulnerabilities were also identified related to performance that would likely be different in the two airplanes, but would not have safety consequences. Possible vulnerabilities that needed to be further verified in the longitudinal study were identified for 19 tasks. Minor vulnerabilities that should be considered when updating training and procedures were identified for 17 tasks.

**Phase 2: Longitudinal Study: Methods**

The participants were 40 line pilots (20 captains and 20 first officers) who currently fly one of the airplane types and were available to be trained in the other type. The pilots were to remain in their current airplane for six months after receiving training in the other type of airplane. Pilots were not considered for participation if he or she had previously flown the “other” airplane, had an advanced entitlement award, or was older than 58 years.

A 777 fixed-based simulator and a 767-400 full-flight simulator (without motion) were used for two systems training modules. Full-flight simulators for each airplane were used for all other training and evaluation events.

All Instructor/Evaluators (I/Es) were calibrated in accordance with state-of-the-art methods and the number of I/Es involved in the study have been kept to a minimum.

**Design and Procedure**

20 flight crews were selected for each aircraft type (i.e., 777 “old” → 767-400 “new” and 767-400 “old” → 777 “new”). Each crew consisted of a Captain and a First Officer. Half of the crews for each fleet were given a simulator evaluation in their current aircraft before any training on the new aircraft began. This evaluation served as the “control” group for the subsequent analyses. Following this evaluation all participants were trained in the “new” aircraft and administered an end-of-course check-ride. All the crews then returned to the line to fly their normal aircraft and were instructed not to read or otherwise attend to anything in relation to flying the new aircraft. At the 90-day interval, half of the crews from each fleet were tested on a simulator evaluation in their newly trained aircraft. At the 180-day interval, all the crews were tested on a simulator evaluation in their newly trained aircraft. Thus half the crews would demonstrate the negative transfer associated with flying for 90 days on one “old” aircraft type, and the other half would demonstrate the negative transfer after 180 days. The evaluation flights were tailored specifically to examine those vulnerabilities identified in the “Differences” phase of the study, as well as to assess tasks for which skill degradation
from not performing in the aircraft for the specified time periods was a possibility.

The protocol and content of the evaluations to be conducted at the end-of-course, at the 90-day interval, and at the 180-day interval were the same and were developed with input by the Boeing members of the MFF study team and FAA AEG inspectors. The evaluation consisted of a flight scenario performed by the crew and maneuvers performed by each pilot individually. It was determined that the maneuvers to include in the evaluations should be those in the standard Appendix F required maneuvers list. An instructor pilot served as the pilot-not-flying during the maneuvers segment to control for the possible confounding effects of having a non-qualified MFF participant serving that role.

Statistical Analysis

In contrast to much statistical analyses in the social science, which adopts traditional null hypothesis significance testing (usually creating a bias toward confirming the null hypothesis of no difference), the current experiment, with its safety implications, took two steps to bias our analysis toward finding differences, where such differences might impair the safety of mixed-fleet flying. First, all tests were one tailed, examining only cases where the simulator tests revealed a degradation of performance relative to the control tests which the pilots received just prior to being trained on the “new” aircraft. Second, a relatively liberal alpha level was set at 0.20 to report any safety-compromising differences that might have emerged.

RESULTS

A series of 36 comparisons were made between items at the control point, the 90-day point and the 180-day point. We note here the following vulnerabilities observed in the 777 tests (by pilots who normally flew the 767-400) when compared to the “control” group:

- Did the pilot initially set the appropriate power setting? Reliable differences were found for the 180-day group.
- Did the pilot apply the appropriate procedure during the “engine failure after V1” maneuver? Reliable differences were found for the 180-day group.
- Delayed response time to activate take-off/go-around (TOGA) switch during rejected landing. Reliable differences were found for the 180-day groups.
- Greater frequency of non-compliance, and poor situation awareness during non-ILS approach and the resulting manual landing. Reliable differences were found for the 180-day group.
- Poorer coupled ILS missed-approach procedures. Reliable differences were found for the 180-day group.

There were actually very few errors in the opposite direction (i.e., errors shown on the 767-400 by pilots who normally flew the 777), the most important of which was a degradation of performance during the coupled missed approach procedures (reliable differences found between the control and the 180-day groups).

DISCUSSION

The current data suggest that mixed-fleet flying is quite feasible with these two types of aircraft, not surprising, given the common manufacturer, and the great number of similarities in design and procedures. While a few examples of negative transfer were revealed (and more, after the longer delay period), these were small in number, given the large number that were actually sampled, (and given that this sample itself was drawn selectively, on the basis of the differences analysis, from a much larger sample that could have been assessed). Thus, assessing under what we might consider a “worst case” scenario, in which the pilot would have little time to practice the new aircraft, and evaluating
those unusual circumstances (e.g., missed approaches, engine outs), where our analyses revealed negative transfer most likely to occur, we still found only a small proportion of instances in which the prior frequent piloting of the “old” aircraft produced a performance decrement. We should note too that had conventional statistical conventions been employed, many of these differences would have been hidden.

Our general assessment of the feasibility of MFF for these aircraft types does not negate the importance of considering those differences that were observed, and our conclusions spawn a tailored set of recommendations regarding how each of these can be mitigated. For example, it was observed that the engine failure during takeoff procedures differed between these two aircraft categories. This difference led, in turn, to a difference in the missed approach procedure for both normal and engine-out situations. By simply harmonizing these procedures, it is believed that most of the variability found during the maneuvers portion of the study would be eliminated. Similarly, by harmonizing the limitations (i.e., static takeoff in icing conditions, crosswind limitations, and autopilot disengagement altitude) it is believed that additional sources of variability (read: pilot confusion) would be eliminated.

In conclusion, we note that this appears to be the first study of its kind, in which actual negative transfer predictions have been evaluated in a full mission simulator with qualified line pilots, sensitivity to sample size (indeed we nearly exhausted the full population of relevant pilots in the airline fleet), and fully realistic tasks.

ACKNOWLEDGMENTS

We would like to note that the work on this project by Research Integrations was funded by the FAA under contract DTFAWA-04-C-00035. This contract was monitored by Dr. Eleana Edens. We would also like to thank Harley Lever for his help with video analysis.

REFERENCES


