Air Traffic Control Simulation Fidelity and Aircrew Training: A Field Study

Alfred T. Lee, Ph.D., CPE

March, 2003

BRI-TR-130303

BETA RESEARCH INC
18379 Main Blvd., Los Gatos, CA 95033 (408)353-2665 Fax: (408)353-6725 www.beta-research.com

Air Traffic Control Simulation Fidelity and Aircrew Training: A Field Study

Abstract

Air Traffic Control (ATC) communication is essential for safe and efficient flight operations. Simulation of these communications during line-oriented training and evaluation would be a logical requirement for all types of flight simulator training. However, a high degree of fidelity is currently not possible for simulators despite the substantial investment in flight simulation technology development which, in every other aspect of flight operations, currently provides significant operational realism. This lack of realistic ATC communications simulation fidelity in airline training simulators may affect crew performance and an instructor's ability to evaluate that performance. In any case, the utility of line-oriented simulations as valid crew assessment tools may be compromised. An assessment of the impact of variations in ATC fidelity on aircrew training and evaluation is needed in order to identify the potential benefits, if any, that may arise from improvements in existing communications simulation technologies. Toward this end, a field study was conducted to evaluate the effects of variations in ATC communications simulation fidelity on crew and instructor/evaluator behavior in Line-Oriented Flight Training (LOFT) at a major airline training organization. Results of the study indicate that variations in ATC fidelity do affect crew communications behavior. When a high level of ATC fidelity is present, a level comparable to that of line operations, management of communications both within and outside the cockpit is affected significantly. While there was no evidence that crews commit more communications procedural errors with high fidelity ATC simulations, the presence of high fidelity frequency chatter regularly produced call sign confusion errors. Crew and instructor comments reiterated the importance of high fidelity ATC communications simulation in LOFT environments. Implications of the findings for the role of ATC fidelity in line-oriented training are discussed.

Air Traffic Control Simulation Fidelity and Aircrew Training: A Field Study

Current airline aircrew flight training relies heavily on flight simulators for the training and evaluation of aircrews. At it's most advanced level of application, flight simulators are used to recreate the entire spectrum of flight operations. This full mission simulation capability or line-oriented simulation (LOS) is now common to airline training around the world. Recreating real life flight operations is complex and expensive. Simulators must provide a high degree of physical as well as perceived (or psychological) fidelity with the actual aircraft and its operating environment. The heavy reliance on these simulators for training and evaluation requires that the simulations have a high enough fidelity to elicit crew behaviors comparable to those that might be expected under actual aircraft operating conditions.

The technical complexity and cost associated with high fidelity flight simulation means that the goal of full fidelity may not always be met. One component of flight simulation, ATC communications, is far less developed then many other areas of flight simulation technology. A recent survey of U.S. air carrier instructor/evaluators (I/Es) indicates that virtually all ATC ownship[1] communications simulations are provided manually by the I/E (Burke-Cohen, Kendra, Kanki, and Lee, 2000). I/Es are not trained in ATC procedures and have only their flying experience to aid them in a task normally carried out by a trained professional. Significant differences in ATC ownship communications content, presentation rate, and communications procedures from those that crews experience in line operations are therefore inevitable in airline simulators. Frequency chatter, communications that are directed to other aircraft over a monitored radio frequency, are seldom provided. Part of the reason for the slow advancement of ATC communications simulation is the technical limitations inherent in current voice recognition and synthetic speech (Lee, 1999). An additional impediment is the absence of evidence that providing a high level of ATC communications simulation fidelity is worth the investment. There is little incentive to improve the existing flight simulation unless some clear benefits can be demonstrated in crew training and evaluation, I/E performance, or crew or I/E acceptance of the value of the technology in enhancing training.

In order to provide data on the value of high ATC communications simulation fidelity, a study was conducted which assessed crew and I/E behaviors in a FAA-certified flight research simulator when ATC fidelity was systematically varied (Lee, 2001). The study revealed significant differences in the content of intracrew communications as well as pilot-initiated communications to sources outside the cockpit (non-ATC as well as ATC). The purpose of the present study was to replicate this study in an actual airline training environment where crew and instructor motivation to perform would be much higher and where a larger number of number of I/Es could be included in the study. Additionally, if significant changes in crew performance can be demonstrated in an actual airline training

[1] Ownship communications are those directed by ATC to the simulated aircraft being flown by the trainees

environment there will be an additional impetus to develop and deploy improved ATC communications simulation technology in flight training simulators. The working hypothesis for the present study, as in the previous study, is that high ATC communications fidelity serves to provide the necessary secondary task loading for the crew that is normally provided in actual line operations. The monitoring of, and the responses to, ATC radio traffic required of the pilot-not-flying is necessary to avoid creating an artificially low workload for this pilot. This, in turn, might affect how other tasks, including tasks other than frequency monitoring, are performed by crews.

Method

Participants

A total of 10 two-person crews participated in the study. Five crews were assigned to each of two groups in a matched group design. Groups were matched for crew time in aircraft type (B-737). Time in type averaged 3,500 hrs for those in the low fidelity (LF) group and 3,800 hrs. for the high fidelity (HF) group, a difference which is not statistically significant. Each participant was enrolled in the Captain upgrade program at the participating major airline training organization. The program culminated in a LOFT session in the flight simulator. Only one of the crewmembers, the one serving in the Captain's role, was evaluated by an I/E during each LOFT, though the performance of both crewmembers was of interest in this study. To avoid any learning effects, a crew served only once in the study. The six I/Es in the study were all senior check airman for the airline, averaging about 14 yrs. in B-737 instruction.

Flight Simulator

A B-737, Level C , FAA-certified flight simulator was used in this study. The same flight simulator was used throughout the study to avoid the introduction of any simulator differences. For the LF crews, an I/E provided all ATC ownship communications and all other non-ATC communications (e.g., dispatch, flight attendant) to the simulated aircraft. These tasks were in addition to his normal duties of operating the simulator and observing/evaluating crew performance. For the HF crews, a professional air traffic controller provided all ATC ownship communications and frequency chatter. The frequency chatter included localized[2] airspace information including calls to other company traffic. In the high HF group, non-ATC communications were provided by the I/E who, as in the LF group, also operated the simulator and evaluated crew performance. The LOFT scenarios used in the simulator were part of the normal Captain upgrade LOFT training system and were approved for this use by the FAA. These scenarios were not altered in any way for the purposes of this study. Each scenario lasted approximately 90 min and required a diversion to an alternate airport. Scenarios typically included a system malfunction in addition to the diversion.

[2] Information about the airspace which is unique to the locale, e.g., specific departure route clearances

Procedure

All aircrews in this study had bid for, and served in, the six week Captain upgrade training program of the airline. The LOFT session served as the last event in the program and the last potential obstacle to promotion. All LOFT scenarios were conducted as they normally would have been by the I/Es. There was no attempt to alter the way in which I/Es conducted the LOFTs, with the exception of those in the HF condition where I/Es no longer provided ATC ownship communications. I/Es were briefed on the general goals of the study regarding ATC communications simulation, but were not aware of any of the specific hypotheses being tested. Aircrews were not aware of the purpose of the study and were assigned to the two groups based only on their aircraft time and availability. Following completion of the simulator sessions, all aircrews and I/Es completed individual questionnaires on the role of ATC communications, the fidelity of the simulation they had just completed, and workload ratings.

Results

In order to evaluate the effects of enhanced ATC communications simulation fidelity on crew and instructor behavior, three areas of crew performance were analyzed: Intra-crew communications[3] (communications between pilots), pilot-initiated communications to sources outside the cockpit, and subjective ratings by crews and I/Es. These measurements were selected as being those most likely to be sensitive to variations in communications workload[4].

Intracrew Communications

Communications between the pilot-flying and the pilot-not-flying were analyzed to determine whether enhanced ATC communications fidelity would have an effect on them. One measure of intracrew communications is the rate of exchange of communications between crewmembers. An index of this is the number of turns in communications per unit time. In any communication, each separate sender and recipient communication is considered a "turn". The number of these turns per unit time is the *turn rate* of the communications between crewmembers. High turn rates imply high rates of information exchange between crewmembers. Low rates imply low rates of exchange or that one or both members are engaged in extensive, explanatory communications as might occur during pre-departure briefings. In any case, a significant difference in mean turn rate between crews in LF and those in the HF condition would suggest that communications between crewmembers are being affected by ATC fidelity. Analyses of intracrew communications in the present study found a mean turn rate for crews in the LF condition of 2.79 while the turn rate for those in the HF condition was 2.36. The difference is statistically significant (Mann-Whitney $U = 11$, $p < .01$).

---

[3] Checklist tasks, briefings, scheduled or other procedural communications are not included these analyses
[4] Captain trainees were scored by I/Es on a pass/fail system, all trainees in this study passed the LOFT evaluation

Pilot-Initiated Extra Cockpit Communications

Pilot-Initiated Extra Cockpit Communications (PECs) were also analyzed to determine whether or not ATC fidelity would affect this type of pilot task. PECs are those communications which are initiated voluntarily by a pilot without prompting by an external source. These include ATC contacts as well non-ATC communications to company (e.g. dispatch) or cabin (e.g., flight attendant) sources. Figure 1 shows the number of PECs for communications in each of the two groups. While average number of total PECs (31) was approximately the same for crews in both groups, the distribution of the PECs differed for the two groups. While there were no significant differences in the percentage of non-ATC and ATC PECS for the LF group (49% vs. 51%), significant differences were found in the average percentage of PECs devoted to ATC communications (43%) when compared to non-ATC communications (57%) in the HF group ($t=2.31$, $p<.01$).

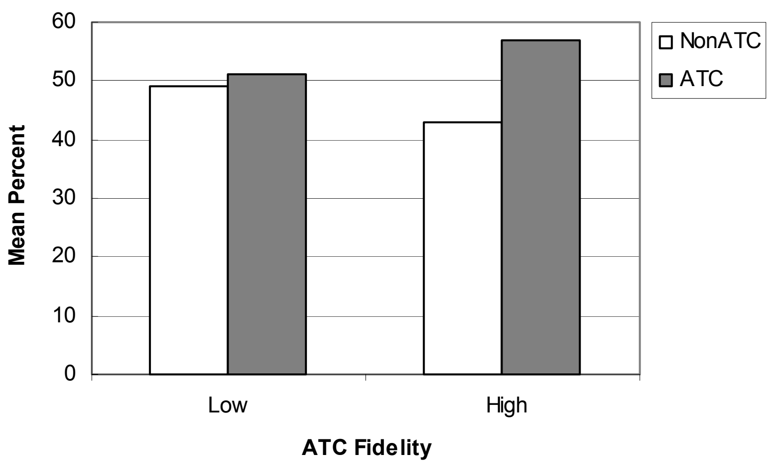

Figure 1. Mean percent Pilot-Initiated Communication as a function of ATC communications fidelity.

A further analysis of the PECs in the HF condition revealed that, while the percentage company communications remained essentially the same, a reduction in the percentage of cabin communications and an increase in ATC communications occurred for crews in this group (see Figure 2).

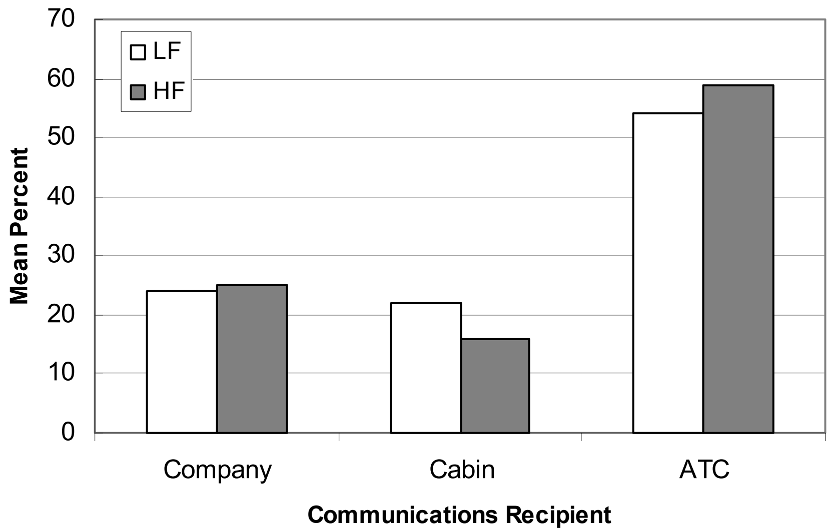

Figure 2. Pilot-initiated external cockpit communications as a function of communications recipient and ATC fidelity.

Communications Procedural Errors

Pilot communications with ATC were also examined for communications procedural errors in this study. These include readback errors, use of the wrong radio frequency, use of incorrect callsign number, wrong or no facility name when contacting an ATC facility, missed ATC calls, and failure to make a facility contact. Call sign confusions, where the pilot contacted ATC in the belief that the message was for his aircraft, were also included in the analyses. Call sign confusions could only occur in the HF condition as only during this condition did crews receive communications intended for other aircraft (i.e., frequency chatter). Figure 3 shows the mean percent communications procedural errors that occurred in each of the two groups. The percent communication procedural errors for each crew is based on the number of ATC ownship communications received by the crew. As with PECs, ownship communication events are defined here as those communications that are initiated by ATC and do not include those which are made in response to pilot transmissions. Although significant differences occur between the LF and HF crews in the percentage of communications procedural errors, the difference disappears almost entirely when call sign confusions are excluded from the analyses. With the exception of the call sign confusions associated with frequency chatter in the HF condition, there are no significant differences between the two groups in the number of communications procedural errors.

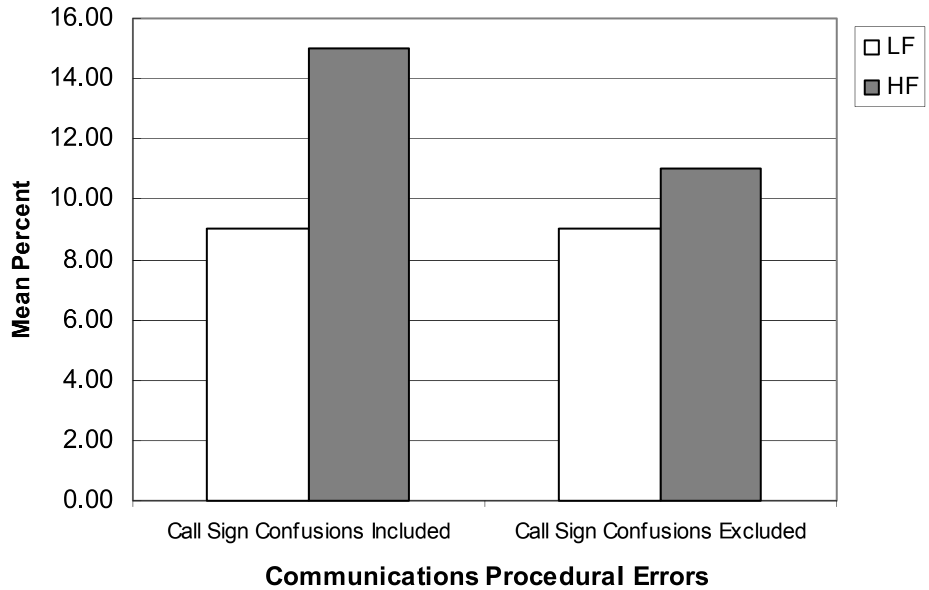

Figure 3. Communications procedural errors a function of ATC fidelity.

Crew and Instructor Ratings

Following the completion of their LOFT session, each crewmember was requested to complete a questionnaire concerning ATC communications simulation. Figure 4 shows the average ratings received for two questions regarding ATC fidelity. The first question asked crewmembers to compare the workload of routine events (e.g., takeoff) in the simulator to those they experienced in the actual operational environment. While there is an increase in rated routine workload for those in the HF condition when compared the LF condition, the difference was not statistically significant. Routine workload does not appear to be affected by ATC fidelity. When asked to compare their simulator experience to those of prior LOFTs, crews in the HF condition rated their simulator experience as significantly greater in operational realism than crews in the LF condition (t=3.0, p<.01). While not experiencing greater overall crew workload, the perceived fidelity as well as crew communications workload was increased by improved ATC simulation fidelity in this study.

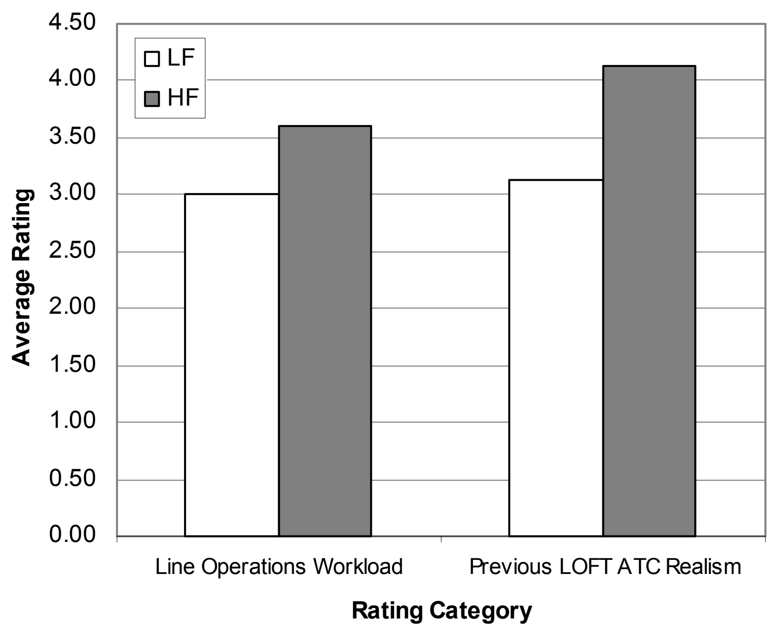

Figure 4. Crew ratings of ATC fidelity and realism

Instructor ratings were also collected in the study. The six I/Es in this study provided feedback on the workload impact as well as the importance of ATC communications simulation in LOFT sessions. The six I/Es rated the effects of providing ATC ownship communications as being low to moderate, a rating in line with previous studies on the impact of ATC simulation task on I/E workload (Burke-Cohen, et al., 2000). When asked to rate the importance of realistic ATC communications simulation in Captain upgrade training, instructor ratings averaged from very important to essential.

Discussion

The simulation of ATC communications has long been an issue for LOS training and evaluation where the goal has been to provide a level of operational realism comparable in every respect to that which a trainee might experience in real life. However, the limited ATC fidelity that can be provided by the I/E for ownship and non-ownship communication, is far from operational realism. The degree of ownship communications that can be provided is necessarily limited due to the workload capacity of the I/E and the fact that it ATC simulation is a task of secondary importance to the I/E. Furthermore, I/Es are not trained air traffic controllers. They are pilots whose skill and knowledge is limited to those ATC procedures experienced during flight operations. The simulation of frequency chatter, common to all flight operations, is rarely simulated in the airline industry despite its important role as a secondary monitoring task for aircrews, particularly for the pilot-not-flying.

The present study was conducted to investigate the effects, if any, of high fidelity ATC communications simulation on aircrews undergoing LOFT evaluation in an actual airline training environment. While a previous study conducted in a research environment allowed for considerable control over simulation variables, the present study provided the opportunity for the assessment of ATC fidelity effects on crew behavior under the more motivating environment of a Captain upgrade LOFT evaluation. The study also allowed feedback from a larger sample of I/Es than was possible in a research environment.

The effects on communications workload management were most evident in the present study. The reduced rate of communications exchange between the two pilots and the change in pilot external communications which resulted from the higher ATC fidelity supports the hypothesis that high ATC fidelity produces crew communications behavior which is significantly different than that currently provided by low ATC fidelity in airline training environments. Additional evidence for this assertion was found in the previous study (Lee, 2001). Under high ATC fidelity, pilot's also tended to shift external cockpit communications to higher priority ATC sources and away from cabin recipients (flight attendants and passengers). These differences are likely due to the overall increase in communications workload demanded by the high ATC fidelity condition, a workload level more like that of line operations. Higher ATC fidelity, with its concomitant demands on pilot communications management, would provide the I/E with the opportunity to observe how aircrews prioritize communications and as well as their ability to manage distractions in high workload conditions.

More realistic ATC communications simulation did not, however, affect the overall rate of communications procedural errors committed by aircrews in this study - with the exception of call sign confusions. Call sign confusions occurred as a direct result of the frequency chatter provided in the high ATC fidelity condition. Confusing one's own call sign with that of another aircraft on a radio frequency is not uncommon in airline flight operations. Call sign confusion, typically the result of similarities between ownship and other aircraft call signs on the same frequency, reflect the lack of full attention paid to ATC communications by some aircrews. This crew performance deficiency can only be observed in LOS when the frequency chatter component of high fidelity ATC is available. It would be important for the I/E to be able to identify this deficiency and correct it in training in order to reduce its occurrence in actual operations (see Sumwalt, Morrison, Watson, and Taube, 1997). As such, LOS simulations should contain frequency chatter and that chatter should allow for the evaluation of call sign confusions which might be committed by aircrews.

The increase in I/E workload resulting from the need to provide even a low level of ATC fidelity is evident from the ratings and comments received by I/Es in this study. Clearly, there is an impact on I/E workload, but it is not clear whether or how this affects the ability of the I/E to evaluate crew performance. It is possible that the low level of ATC fidelity provided is such as to not affect I/E evaluation performance, though any impact on I/E workload which is not relevant to the task of training and evaluation is inherently undesirable.

In the present study, as in the previous study, high ATC fidelity was defined by the employment of a professional air traffic controller and the presence of frequency chatter localized for the routes simulated. This provided a level of ATC fidelity which was measurably higher in physical and perceived fidelity than the low fidelity condition where only limited ownship communication was provided by the I/E. However, high ATC communications simulation included both ownship communications and non-ownship communications (frequency chatter). While certain effects, such as call sign confusions, were clearly caused by the presence of frequency chatter, changes in communications management could be the result of either of these components or of a combination of them. In order to separate the effects of frequency chatter from ownship communications, research will be needed to assess these components separately in LOS. It should be noted that the answer to these questions are not simply of academic interest. The cost of providing a high degree of realistic ownship communications is likely to be considerably higher than providing realistic frequency chatter. Therefore, identifying their individual contributions to enhanced LOS training is an important goal for developing cost-effective ATC communications simulation requirements.

References

Burke-Cohen, J., Kendra, A.J., Kanki, B.G., and Lee, A.T. (2000). *Realistic Radio Communications in Pilot Simulator Training*. (Tech. Rpt. DOT/FAA/AR-00/62). Washington, D.C.: Office of Aviation Research.

Lee, A.T. *Radio Communications Simulation Fidelity and Aircrew Training.* (2001). (Tech. Rpt. BRI-TR-130901). Los Gatos,CA: Beta Research, Inc.

Lee, A.T. (1999). *Advanced Technology Applications in Flight Simulation and Training: Voice Generation and Recognition*. (Tech. Rpt. BRI-TR-121099). Los Gatos, CA: Beta Research, Inc.

Sumwalt, R, Morrison, R, Watson, A. and Taube, E. (1997). What ASRS data tell about inadequate flight crew monitoring. In *Proc. of the Nonth International Symposium Aviation Psychology,* pp. 977-983.