In Proceedings of the International Conference Flight Simulation--The Next Decade, Royal Aeronautical Society, 10-12 May 2000, London, UK, pp. 23.1-23.7.

Simulator Fidelity – The Effect of Platform Motion

Dr J Bürki-Cohen, US Department of Transportation, E Boothe and Dr N Soja, Consultants, Dr R DiSario, Bryant College, Dr T Go, Massachusetts Institute of Technology and Dr T Longridge, Federal Aviation Administration, USA

CORRECTION

Due to a cell reference error, the number of subjects in Figures 5 and 6 on page 23.5 were reported incorrectly. The correct *n*'s for V_1 cut Transfer are 16 for the Motion group and 18 for the No-Motion group.

Simulator Fidelity – The Effect of Platform Motion

Judith Bürki-Cohen
Engineering Psychologist
Volpe Center, U.S. Department of Transportation, Cambridge, Massachusetts

Edward M. Boothe
Consultant, Flight Simulation and Training, Atlanta, Georgia

Nancy N. Soja
Consultant, Experimental Psychology
Brookline, Massachusetts

Robert DiSario
Assistant Professor
Bryant College, Smithfield, Rhode Island

Tiauw Go
Postdoctoral Associate
Massachusetts Institute of Technology, Cambridge, Massachusetts

Thomas Longridge
Manager, Advanced Qualification Program
Federal Aviation Administration, Washington, D.C.

Abstract

This research is part of the Federal Aviation Administration's (FAA) initiative towards promoting affordable flight simulators for U.S. commuter airline training. This initiative becomes even more important as the FAA is considering regulatory action that will mandate the use of simulators for all air carrier flight-crew training and qualification. Consequently, sound scientific data on the relationship between certain simulator features such as platform motion and their effect on the transfer of pilot performance and behavior to and from the respective airplane become very important. The present study examined the effect of platform motion (i.e., FAA qualified Level C six-degree-of-freedom synergistic motion) in the presence of a high-quality wide-angle visual system on 1) pilot performance and behavior for *evaluation* prior to any repeated practice or training, 2) *the course of training* in the simulator, and 3) the *transfer of skills* acquired during training in the simulator with or without motion to the simulator with motion as a stand-in for the airplane (quasi-transfer design). Every effort was made to avoid deficiencies in the research design identified in a review of prior studies, by measuring pilot stimulation *and* response, testing both maneuvers and pilots that are *diagnostic* of a need of motion, avoiding pilot and instructor *bias*, and ensuring sufficient statistical *power* to capture operationally relevant effects. Results of the analyses and their implications are presented in this paper.

Nomenclature

FAA	Federal Aviation Administration
PTS	Practical Test Standards
RTO	Rejected Take-Off
V_1	Take-off decision speed; the minimum speed in the take-off, following a failure of the critical engine, at which the pilot can continue the take-off and achieve the required height above the take-off surface within the take-off distance.
V_1 cut	Engine failure at or above V_1 with continued take-off
V_2	Take-off safety speed; a speed that will provide at least the gradient of climb required by the airplane certification rules with the critical engine inoperative.
PF	Pilot Flying
PNF	Pilot Not Flying
I/E	Instructor/Evaluator
n	Sample size
p	Probability of null hypothesis (i.e., no effect of motion)
r	Pearson correlation coefficient
STD	Standard Deviation

Introduction

This research effort is part of the Federal Aviation Administration's (FAA) initiative towards promoting the availability and affordability of flight simulators for U.S. commuter airline training. This initiative (Ref. 1) becomes even more important as the FAA is proposing a rule that would mandate the use of simulators for all air carrier training and qualification, limiting the use of the aircraft itself as a training option even for small regional airlines. However, there is a lack of sound scientific data on the relationship between certain key training device features, such as platform motion cuing, and their effect on the transfer of performance to and from the airplane. This project will develop a scientific basis to assure that FAA requirements promote full transfer of performance between simulator and airplane—anything less would compromise safety. The data will also help the FAA to evaluate air carrier proposals for the alternative use of other training equipment in lieu of full flight simulators. The first stage of this multi-year project was a state-of-the-art review of key aspects of flight simulation, involving both FAA and Industry subject matter expert workshops (Refs. 2 and 3) and an extensive literature review (Refs. 4 and 5). Based on this review, an empirical investigation of flight simulator requirements which seeks to correct deficiencies in the research design of prior studies has been initiated.

The present study empirically examined the effect of platform motion (i.e., FAA qualified Level C six-degree-of-freedom synergistic motion) in the presence of a high-level visual system (i.e., wide-angle collimated cross-cockpit) on pilot training and pilot evaluation. It addressed the questions of whether the motion provided by an FAA qualified Level C simulator affects 1) First Look evaluation of pilot performance and behavior prior to any simulator practice, 2) the course of Training in the simulator, and 3) the Transfer of training acquired during training in the simulator with or without motion to the simulator with motion as a stand-in for the airplane. The analysis also examined whether the grading criteria used by the instructors/evaluators (I/Es) were affected by the presence or absence of motion. The statistical power of the experiment was also considered.

Method

An FAA qualified Level C flight simulator was used in the experiment. It represents a 30 passenger, three crew, turboprop airplane with wing-mounted twin engines and counter-rotating propellers. The six degree-of-freedom synergistic motion system with hydraulically actuated legs is capable of a 60 inch stroke. The high quality visual system provides wide angle collimated cross-cockpit viewing with a 150 degrees horizontal and 40 degrees vertical field of view available to each pilot.

The research was conducted using regional airline pilots in recurrent training. Data were collected from 42 crews. Two experiments were combined into one session to minimally disrupt the host airline's training and evaluation program, as well as to reduce pilot adaptation to a simulator configuration. "First Look evaluation" was designed to assess the effect of motion on the effectiveness of the simulator as a tool for evaluating the crew's aviating skills. In other words, it assessed the degree to which a pilot's existing skills transferred *from* the airplane *to* the simulator for each simulator configuration. This assessment needed to occur during the very initial exposure of the crew to the simulator, so that pilots' behavior and performance would reflect their actual skills in the airplane with as little contamination as possible from potential adaptation to a particular simulator configuration. The second experiment was designed to assess the effect of each simulator configuration on skill acquisition in the simulator, and, most importantly, on subsequent transfer of these skills to the airplane. This experiment was called "Training and Transfer testing." Training transfer was measured by comparing the effect of training received in the simulator, with and without motion, on performance and behavior in the simulator with motion (as a stand-in for the airplane, "quasi-transfer" design).

Two test maneuvers (i.e., pilot tasks) were chosen to maximize satisfaction of criteria described in the literature as diagnostic for the detection of a motion requirement, given the constraint that the experiment was conducted in the context of an FAA approved training program. These criteria included 1) closed loop, to allow for motion to be part of the control feedback loop to the pilot; 2) unpredictable and asymmetric disturbance, to highlight an early alerting function of motion (Ref. 6); 3) high gain and high thrust, to magnify any motion effects; 4) high workload with crosswind and low visibility, to increase the need for redundant cues such as provided by motion, out-the-window view, instruments and sound; and 5) short duration, to prevent pilots from adjusting to a lack of cues. Engine failures on take-off with either rejected take-off (RTO) or continued take-off (V_1 cut) were deemed as fulfilling most of these criteria, while requiring minimum disruption to the host airline's existing training program. To prevent bias, the state of the motion system was kept concealed from all participants.

A laptop computer was programmed to control the simulator and record events with minimal I/E intervention, eliminating the need for the presence of an experimenter that might have contaminated the regular training/evaluation environment and enabling the I/E to focus on behavior and performance of the crew. Even more importantly, this also eliminated any need to inform the I/E (or the crew) of the interest in motion and the motion state of the simulator for each maneuver, thus minimizing any bias.

The stimulation of the pilot by the simulator and the pilots' responses were measured by recording 78 simulator state and control input variables at a high sampling rate, resulting in a vast amount of objective data on simulator performance and pilot performance, behavior, and workload. Two forms of subjective data were also collected. First, at the conclusion of each maneuver the I/E provided a grade for the just-completed maneuver. Second, at the end of the training period and again at the end of the transfer period all participants were queried on PF performance and workload as well as simulator comfort and acceptability.

Motion Stimulation Provided by the Test Simulator

For the test simulator, the actually measured roll and longitudinal accelerations followed the airplane model fairly well given the limitations inherent to all simulators. For vertical acceleration, however, the motion system of the test simulator did not respond much to the command provided by the equations of motion. This is especially true for V_1 cut maneuvers. However, because the engine failures used in our experiment do not produce much vertical acceleration, the lack of vertical acceleration cuing may not be very important.

More important, however, is the finding that failure-induced lateral acceleration was not well represented by the motion system of the test simulator. Not only was it greatly attenuated, but visual inspection of the measured response does not lead to an easy distinction of failure-induced lateral acceleration, unlike the response derived from the equations of motion (relatively high peak shortly after engine failure). This may represent a significant deficiency in pilot stimulation, because lateral acceleration may act as a useful cue for proper failure recognition and for delivery of appropriate action. To the best of our knowledge, however, the importance of lateral versus other cues in failure recognition has not been systematically examined in the literature.

Analysis of I/E Grades

Figure 1 shows the grade distribution obtained by the two groups at First Look evaluation and Transfer. The possible grades were 1 (unsatisfactory), 2 (FAA Practical Test Standards, Ref. 7), 3 (company standards), and 4 (excellent). The experimental session appeared to have been effective in simulating a real training session in that the crews' performance improved across the session. Specifically, combining the two motion groups (or looking at them individually), the grades for RTOs and V_1 cuts improved across the training trials. This was even stronger for the V_1 cuts, which elicited lower grades than the RTOs during First Look, but caught up by Transfer.

Turning to the effect of motion, the presence or absence of motion had no effect on the grades for the RTOs at either First Look or Transfer. There was also no effect of motion for the V_1 cuts at First Look. Whether crews were trained with motion or without had also no effect on Transfer to the simulator with motion or improvement from First Look (or last training) to Transfer, at least not when comparing group means or number of grades of one and two vs. three and four. However, the motion-trained crews did receive more grades of two than the crews who had not previously had motion, and fewer grades of one (none actually). Additionally, there was no effect of motion on the course of Training or on the amount of training required before reaching the criterion needed to move onto Transfer for either of the maneuvers.

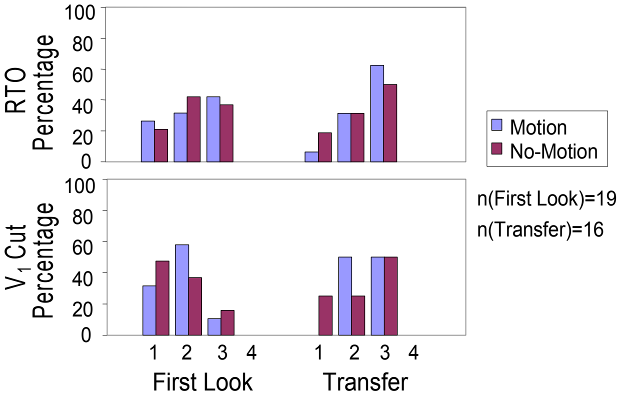

Fig. 1 Grade Distribution

Analysis Of Objective Data

From the 78 variables recorded in the experiment, a set of criterion measures was derived for determining whether or not motion had an effect on training and evaluation of the tested pilot task. These were categorized into performance and workload/behavior measures. Performance measures reflect a pilot's control precision and efficiency in handling the airplane by measurements such as flight path deviations and reaction time. Workload/behavior measures describe how a pilot uses the controls by measurement of control inputs. A guide to the determination of the measures was provided by the PTS and by the company standards of the host airline itself. An additional goal was to capture performance and workload immediately after the engine failure, because disturbance motion was expected to act as an alerting cue to the pilots that would enhance early performance. The list of the measures can be seen in Ref. 5. Most of the measures were computed over the 15 second time period following an engine failure. Exceptions include measures of reaction times and time to reach 400 ft altitude. In general, lower numerical values of the measures indicate better performance or lower workload.

The effect of motion on First Look evaluation, Transfer of training to the simulator, Training Progress, and improvement from last training trial to Transfer testing was examined. An attempt was also made to capture the criteria used by the I/Es when they were grading the pilots on the respective maneuvers by performing regression analyses between I/E grades and objective measures. These analyses showed whether the presence or absence of motion affected which measures I/Es considered for grading.

The objective measures that are discussed in this paper are the ones that are either listed in the PTS, were used by the instructors for grading, or showed an effect of motion. For each measure, the statistical power was determined (i.e., the smallest effect that could be detected given the idiosyncratic variability between crews with a probability of .80). The power of the experiment was found to be sufficient to capture any operationally relevant effects.

I/Es' Grading Criteria Linear and logistic regression analyses on the relationship between the grades and the objective measures were used to infer the I/Es' grading criteria and whether the platform motion had an affect on these criteria. Although the logistic regression was considered to be more appropriate for cases involving ordinal data (like the grading system used here), the results of both regression analyses were quite similar. The regression models obtained were not meant to model I/E's decision process in determining the grades, which is actually very complex. They were only used to examine whether any available measures contributed to the I/E's grading criteria.

For RTOs, regardless of whether the platform motion was on or off, the measures of lateral and heading deviations played an important role in predicting I/E grades. For V_1 cuts, the results of the regression analyses suggest that the platform motion status may affect grading. In both motion-on and motion-off conditions, some (but not the same) lateral measures seemed to affect I/E grades. However, the level of importance of other types of measures in the I/Es' grading criteria depended on the status of the platform motion. Notably, longitudinal measures appeared to matter mainly when the platform motion was on.

-

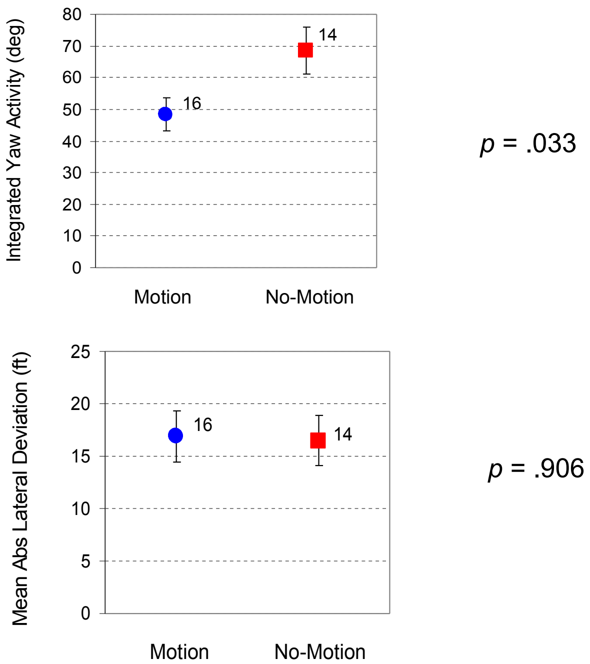

Fig. 2 RTO First Look: Directional Performance

First Look Evaluation, RTOs Motion did not affect performance in heading deviation, lateral deviation, power lever reaction time, nor did it affect any workload measures. It did, however, improve Integrated Yaw Activity (integral of absolute yaw rate for 15 seconds after engine failure, see Fig. 2 [i]), a measure which was not found to be important in the I/Es' grading criteria. This suggests that for First Look evaluation, the presence of motion may improve performance, but not to the extent of affecting grades.

First Look Evaluation, V_1 Cuts Motion did not affect bank angle or heading control variables (and these, especially bank angle, are important for grades) or reaction time. Interestingly, there is a marginal chance that motion may have improved the pitch angle standard deviation (*p*<.10) (Fig. 3). This effect was, however, physically small and not accompanied by any other performance or workload effects. This, together with the fact that there was practically no simple correlation between STD Pitch Angle of Motion pilots and grades (r^2 = .01), and even the stepwise regression model selecting three more longitudinal measures accounts for no more than 30 percent of the variance in the grades, suggests that the platform motion would not affect pilot grades during First Look evaluation. This result also validates the subjective grade results presented earlier.

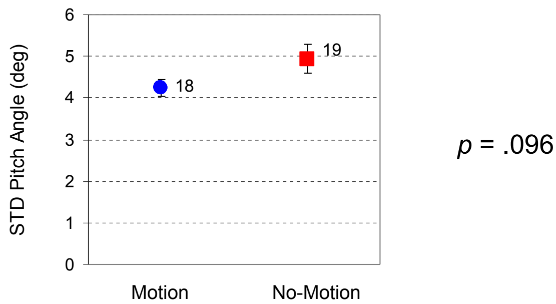

Fig. 3 V_1 Cut First Look: STD Pitch Angle

Training Transfer, RTOs Training Transfer was tested for all crews on the simulator with motion activated as a stand-in for the airplane. Despite the fact that the Motion crews were trained and tested on the same simulator configuration, they did not do any better than the No-Motion crews with any RTO performance and workload measure. Additionally, the power of the experiment was generally higher after training, and still no effects of prior motion were found. One *caveat* is that for heading control, although there was no difference between the two groups, more No-Motion crews improved than Motion crews between the last training and the Transfer testing (see Fig. 4). This may indicate that the addition of motion was beneficial, although during Transfer testing the two groups performed at the same level (as just described).

[i] In this and subsequent figures, numbers next to data points refer to sample size

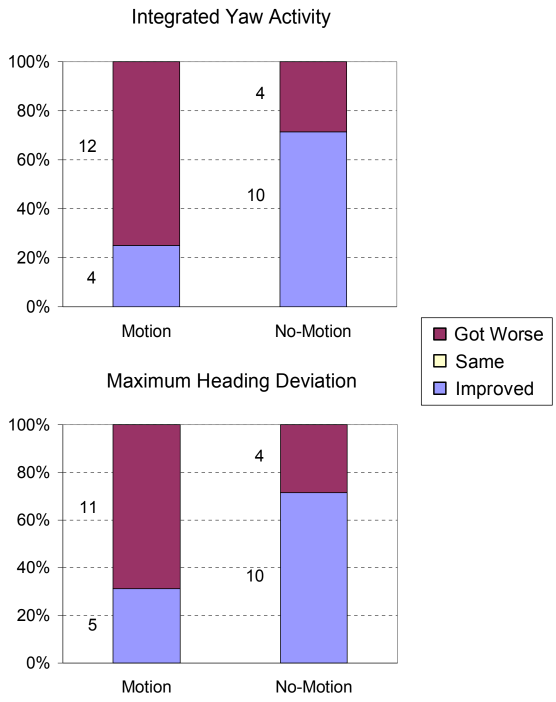

Fig. 4 RTO Last Training vs. Transfer: Directional Control Performance

Training Transfer, V_1 Cuts Having been trained with motion did improve speed control (*p*=.006), i.e., Integrated Airspeed Exceedance (integral of absolute airspeed deviation outside (0,+5 knots) band from the recommended V_2) following V_1 cuts during Transfer (Fig. 5). It came at the price of increasing pitch angle standard deviation (*p*=.025), but was still advantageous because of the critical role speed plays in aircraft control and safety, e.g. for clearing obstacles and maintaining a margin above stall speed. However, there was also an increase in Integrated Yaw Activity by the motion-trained group, although it did not appear to affect heading.

With regard to workload during V_1 cuts, the Motion group had fewer wheel reversals than the No-Motion group (*p*=.059), whereas the No-Motion group had fewer pedal reversals than the Motion group (*p*=.008) (see Fig. 6). The increased number of Wheel Reversals of the No-Motion group was not accompanied by any lateral performance differences. The increased number of pedal reversals of the Motion group, however, was accompanied by an increase in Integrated Yaw Activity, as was discussed earlier. The difference was not apparent at First Look, nor did a combined Analysis of Variance (ANOVA) of Motion/No-Motion by First Look vs. Transfer find a significant interaction, probably due to the variability in number of pedal reversals for the Motion group during First Look. The questionnaire data indicated that the Motion group felt the pedal was less like the airplane than the No-Motion group did.

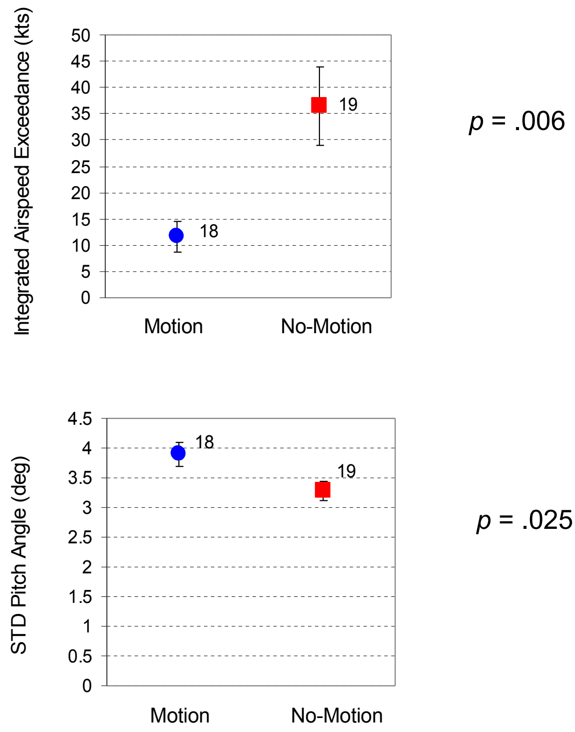

Fig. 5 V_1 Cut Transfer: Longitudinal Performance

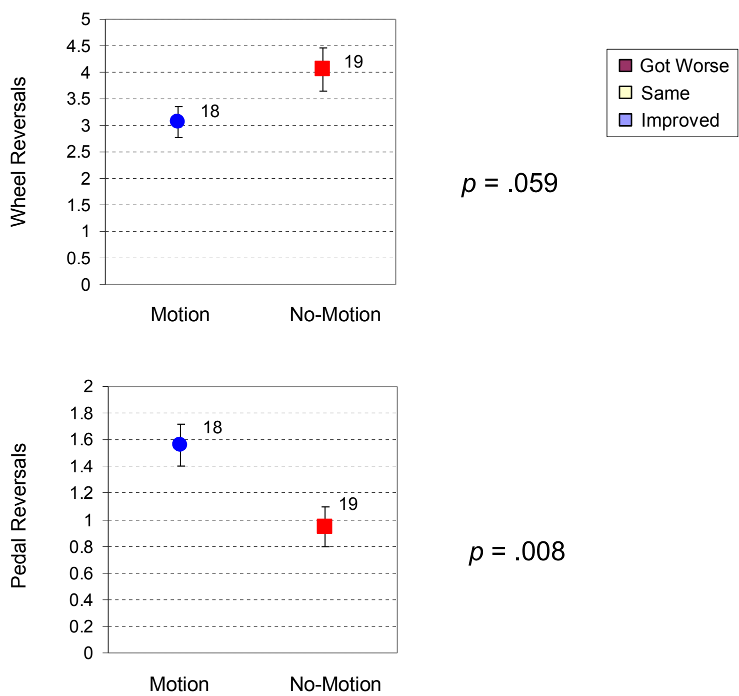

Fig. 6 V_1 Cut Transfer: Wheel and Pedal Reversals

Training Progress, RTOs No statistically significant differences in improvement from first to last training trial were found between groups for any of the measures (all p>.2). This suggests that the platform motion did not affect the training progress of the pilots.

Also, the overall number of crews (Motion and No-Motion) improving in lateral performance and workload measures was significant for most measures, with the exception of Integrated Yaw Activity with no overall improvement and pedal reversals, which actually increased after training. When looking at the groups separately for these two measures, neither of the groups shows any improvement or deterioration. This confirms that the pilots generally did improve during training regardless of the motion status of the simulator.

Fig. 7 V_1 Cut First vs. Last Training: Longitudinal Performance

Training Progress, V_1 Cuts The course of training for V_1 cuts reflected the Transfer results. Motion did improve training progress for Integrated Airspeed Exceedance, but hindered training for STD Pitch Angle (the price for reduced Integrated Airspeed Exceedance) (Fig. 7) and also for the heading variables. For workload variables, there were no differences between the two motion groups.

The data indicate that the No-Motion group improved on more measures than the Motion group. While Motion crews improved in Integrated Airspeed Exceedance and STD Column Position only, the No-Motion crews improved in Integrated Bank Angle Exceedance, Heading Deviation, Time to Reach 400 ft Altitude, and STD Pitch Angle. During Transfer, however, the No-Motion group surpassed the Motion group only with steadier pitch angle and yaw activity; and the actual size of these differences was very small.

The above discussion indicates that the training without motion was at least as effective as the training with motion, and the earlier results on Transfer show that although some differences were found in training progress between the two groups, they did not translate into operationally relevant differences during Transfer.

Analysis Of Questionnaire Data

Each of the PFs and PNFs was given two questionnaires (i.e., one after Training and one after Transfer) that each had six questions (i.e., control precision, control strategy and technique, workload, gaining proficiency, simulator comfort and acceptability). Each I/E was also given two questionnaires, each with five questions (i.e., the same questions as above, but without acceptability). PFs responded always with reference to themselves. PNFs and I/Es referred to the PFs, with the exception of comfort and, for the PNF, acceptability.

Regarding motion, given all of these questions, only four differences were found between the Motion and No-Motion crews. 1) After Training, the PNFs from the No-Motion crews rated the control precision of the PFs better than the PNFs from the Motion crews did. 2) The PFs from the No-Motion crews, once transferred to the simulator with motion, rated their control precision higher than their motion-trained counterparts. This is possibly because of the contrast between the added motion and the lack of motion they had been experiencing. 3) In contrast, after Transfer, the I/Es gave higher ratings for performance to the PFs from the Motion group than to the PFs from the No-Motion group. 4) Looking across both questionnaires, the PFs from the No-Motion crews gave better ratings to the simulator for training ("gaining proficiency") than the PFs from the Motion crews.

Conclusions and Recommendations

The results of this study indicate that the motion provided by the test simulator, which may or may not be typical of other FAA qualified Level C flight simulators, does not, in an operationally significant way for the tasks tested, affect either First Look evaluation, Training Progress, or Transfer of training acquired in the simulator with or without motion to the simulator with motion. It also doesn’t consistently affect the PFs’, PNFs’, and I/Es’ subjective perception of the PFs’ performance, workload, and training, or of their own comfort in the simulator. Neither does it affect the acceptability of the simulator to the PF and the PNF.

Two *caveats* have to be kept in mind, however. First, the simulator used in this study may not have provided sufficient motion to be effective. The measurements indicate that the simulator may have failed to provide lateral acceleration cuing representative of the aircraft for the test maneuvers (RTO and V_1 cut).

A second *caveat* is that the current study used the simulator with motion as a stand-in for the airplane. Although some may believe that this quasi-transfer design needs to be validated, others may say that high-level simulators have been validated as a stand-in for the airplane by many years of use of the simulator for total flight training. Also, given that the motion-trained group transferred to the same simulator configuration that they had been trained in, whereas the No-Motion group transferred to a configuration that was new to them (i.e., the motion configuration), the Motion group should have had an advantage. Based on the quasi-transfer results, it is unlikely that it would have had a greater advantage transferring to an airplane.

Clearly additional steps must be taken to determine the extent to which it may or may not be appropriate to draw generalizations from these results. These should include a comparison of the objective measures from the motion system used in this experiment with such measures taken from other FAA qualified Level C simulators to determine whether or not the motion used in the present study is representative. This should be followed by an investigation on whether operational relevant effects of motion would be found with a simulator where the motion is manipulated to assure that it is representative of the airplane for the maneuvers selected. Additional maneuvers that may be diagnostic and a different pilot population should be tested as well. Ideally, some validation of the quasi-transfer design with a real airplane would also be undertaken.

Acknowledgments

This work was funded by the Office of the Chief Scientist for Human Factors of the Federal Aviation Administration, AAR-100. The authors would like to thank the FAA Program Manager Dr. Eleana Edens for her guidance throughout the project. Many thanks also to Mr. Paul Ray, Manager of the National Simulator Program Office AFS-210, and Dr. Ed Cook of the same office, for their continued interest and advice.

The opinions expressed are those of the authors and not necessarily those of the Department of Transportation, the Federal Aviation Administration, or the U.S. Government.

References

1. Longridge, T., Ray, P., Boothe, E.M., and Bürki-Cohen, J., Initiative towards more affordable flight simulators for U.S. commuter airline training, paper presented at the Royal Aeronautical Society Conference on Training—Lowering the Cost, Maintaining the Fidelity, London, May 1996.
2. Transcript [ii] of the Joint FAA/Industry Symposium on Level B Airplane Simulator Aeromodel Validation Requirements, Washington Dulles Airport Hilton, March 13-14, 1996.
3. Transcript [ii] of the Joint FAA/Industry Symposium on Level B Airplane Simulator Motion Requirements, Washington Dulles Airport Hilton, June 19-20, 1996.
4. Bürki-Cohen, J., Soja, N., & Longridge, T., Simulator Platform Motion — The Need Revisited, The International Journal of Aviation Psychology, Vol. 8, No. 3, 1998, pp. 293-317._
5. Bürki-Cohen, J., Soja, N.N., Go, T.H., Boothe, E.M., DiSario, R., and Jo, Y.J., Simulator Fidelity: The Effect of Platform Motion, Report No. DOT/FAA/RD-00/XX, in prep aration._
6. Gundry, J., Man and Motion Cues, paper presented at the Third Flight Simulation Symposium, London, April, 1976.
7. Federal Aviation Administration, Airline Transport Pilot and Type Rating Practical Test Standards, FAA-S-8081-5B, U.S. Government Printing Office, Washington, D.C., July 1995.

[ii] Available in electronic format from Dr. Thomas Longridge, Advanced Qualification Program Manager, AFS-230, tel. (703) 661-0275