Proceedings of the 12th International Symposium on Aviation Psychology, April 2003

SIMULATOR FIDELITY REQUIREMENTS FOR AIRLINE PILOT TRAINING AND EVALUATION CONTINUED: AN UPDATE ON MOTION REQUIREMENTS RESEARCH

Judith Bürki-Cohen,[1] Tiauw H. Go,[2] William W. Chung,[3] Jeffery Schroeder,[4] Sean Jacobs,[1] Thomas Longridge[5]

[1]DOT/Volpe National Transportation Systems Center DTS-79
[2]Massachusetts Institute of Technology
Cambridge, Massachusetts
[3]Northrop Grumman Information Technology
Moffett Field, California
[4]NASA Ames Research Center
Moffett Field, California
[5]Federal Aviation Administration, Voluntary Safety Programs Branch, AFS-230
Washington, DC

Preliminary results are presented on the effect of enhanced hexapod motion on airline pilot recurrent evaluation, training, and transfer of training to the simulator with motion as a stand-in for the airplane (quasi-transfer). A first study, which tested "as is" motion in an FAA qualified full flight simulator, had not found any effect of motion. Under the enhanced motion conditions of the present study many effects of motion emerged that have not been previously shown in the airline-pilot training and evaluation context, indicating that motion may be required at least for pilot evaluation purposes. The implications of the results for recurrent training are also discussed.

Introduction

The FAA has proposed a rule that would establish regulatory requirements for simulator qualification based on existing advisory criteria, which are generally recognized as having been effective for that purpose based on several decades of experience. However, as these qualification criteria transition from advisory to regulatory status, it becomes increasingly important that any future modifications be based on sound scientific data. The existing qualification standards for simulator platform motion remain controversial in some circles. For example, there is a paucity of data supporting the hypothesis that an effect of motion on pilot-vehicle performance or pilot control behavior found in the simulator transfers to the airplane. Such effects have only been shown to a limited extent for quasi-transfer, but not in airline operations. Many of the studies addressing this issue, however, used non-diagnostic maneuvers or lacked the statistical power to find an effect (see Bürki-Cohen, Soja, & Longridge, 1998, and Boldovici, 1992, for a review).

A first study in the framework of the Volpe Center's Flight Simulator Fidelity Requirements Program (see Longridge, Bürki-Cohen, Go, & Kendra, 2001) investigated the role of motion in a typical FAA qualified Level C turboprop simulator on recurrent airline pilot qualification. It did not find a systematic effect of simulator motion on pilot control-input behavior or pilot-vehicle performance during evaluation, training, and transfer to the simulator with motion as a stand-in for the airplane (quasi-transfer). The presence or absence of motion also had no effect on pilots' opinions of the simulator. The same study also found, however, that the lateral acceleration stimulation provided by the simulator was substantially attenuated (Go, Bürki-Cohen, & Soja, 2000). A look at eight other simulators indicated that attenuated lateral acceleration might, however, be typical for the type of simulator used in airline training and evaluation.

A follow-up quasi-transfer study in collaboration with the National Aeronautics and Space Administration (NASA) tested whether it is possible to improve the motion realism of a typical hexapod platform such that the presence of motion does affect pilot control-input behavior and pilot-vehicle performance. For this purpose, the platform motion software of the NASA/FAA B747-400 simulator was modified to enhance its motion fidelity for the maneuvers tested. Preliminary results of this study are presented below.

Method

Participants

Forty current Boeing 747-400 Captains and First Officers volunteered as Pilots Flying (PF). Each flew from their usual seat. They were compensated for time and expenses. Each PF participated in either the Motion or the No-Motion group, resulting in 20 PFs per group. The Motion group was evaluated and trained with motion. The No-Motion group was evaluated and trained without motion. Both groups were subsequently tested for transfer of training in the simulator with motion as a stand-in for the airplane (quasi-transfer).

Two retired airline captains served as the Pilots Not Flying (PNF). They performed minimal functions as instructed by the PF. Only the PNFs knew about the purpose of the experiment. The motion status was concealed from both the PFs and PNFs. A retired Air Traffic Controller (ATC) provided ATC instructions and operated the simulator.

Maneuvers

The test maneuvers satisfied the criteria described in the literature as diagnostic for a need of motion (see, e.g., Gundry, 1976; Hall, 1989):

1) skill- instead of procedure-based;
2) closed-loop to receive feedback from motion;
3) disturbances intruding on the control loop to highlight an early alerting function of motion;
4) asymmetric and high gain to magnify any motion effects and reduce the stability margins of the pilot-vehicle control loop;
5) high workload to increase the need for multiple cues.

The four maneuvers consisted of two failures with continued takeoff and two engine-out landing maneuvers with weather. All failures involved an outboard engine. The autothrottle was selected inoperative throughout.

For the takeoffs, the engine was failed either at V_1 to generate high asymmetry (V_1 cut) or after V_2 at 40 feet above ground level (AGL) to eliminate visual feedback from the runway, increasing reliance on motion (V_2 cut). The runway visible range was kept low at 600ft for both engine cuts.

The landings consisted of a Precision Instrument Approach (PIA) with shifting10-12 knots head-to-tail winds and a Sidestep Landing (SL) with vertical upgust. Both landings were hand-flown without a flight director with one engine out. The PIA was guided by the Instrument Landing System [ILS, localizer (LOC) and glide slope (GS)] and performed at low visibility (500ft cloud ceiling and 5200ft Runway Visual Range). The sidestep occurred at about 1000ft altitude from the left to the 1200ft apart right parallel runway. The gust was applied on the runway rollout and peaked at 25ft/s.

Simulator

The NASA FAA B747-400 simulator was qualified at the highest FAA level, i.e., Level D. Its high-brightness and high-resolution visual system provided a wide field-of-view panoramic out-the-window view with cross-cockpit viewing. The sound system provided direction and sound quality cues covering the entire operating range of the engines, including the failure simulated. The control-feel cues provided were equally inclusive.

The simulator met the FAA Level D Quarterly Test Guide requirements, and daily calibration was performed. The measured delays and bandwidths (BWs) for motion and visual cueing were well within requirements. The six simulator actuators provided a 54-inch stroke. To improve the motion cues compared to the first study, the lateral acceleration motion cues (and the heave cues) were enhanced. These enhancements, consisting of increasing the cue magnitude and decreasing the cue phase error, were in accordance with previous motion fidelity research (Bray, 1972; Schroeder, 1999; Mikula, Chung, & Tran, 1999).

Design and Procedures

The between-subjects design with Motion vs. No-Motion evaluation and training as the principal independent variable (IV) controlled for confounding effects of practice and sequence. PFs were counterbalanced across the two conditions with respect to seat, PNF, and experience.

PFs were briefed upon arrival that they would be flying challenging maneuvers to test different simulator configurations. Both pilot-vehicle performance and their control inputs would be recorded in addition to video and audio recording in the cockpit. They would complete extensive questionnaires on how they perceived the simulator and their workload. They were told that they would fly around the airport and were given airport, weather, and airplane information. All briefings were oral and written.

PFs were also told that they would be given a chance to practice the maneuvers with graphical feedback on their flight precision performance displayed on the navigation display screen. The feedback would be given with respect to the Practical Test Standards (FAA-S-8081-5D, 2001).

Each experiment run lasted up to seven hours including lunch. In the morning, PFs were evaluated and trained with or without motion dependent on group. In the afternoon, simulator motion was turned on for both groups to measure quasi-transfer of the skills and control strategies acquired in the simulator with or without motion. The experiment's three phases and sequence are given below.

Phase I. Evaluation

1) Evaluate V_2 cut (Engine 1) followed by PIA on parallel runway (Scenario 1)
2) Evaluate V_1 cut (Engine 4) followed by SL from parallel to same runway (Scenario 2)
3) PF and PNF complete Questionnaire 1
4) Briefing on Feedback Displays using display copies printed during Scenario 2. Heading (HDG) and speed deviation, altitude and bank angle (bank) are displayed for takeoffs; glide path, LOC, and approach speed deviation are displayed for landings

Phase II. Training

5) Train three instances in a row of each maneuver, while failing the opposite engine from Evaluation. For the landing maneuvers, the simulator comes off freeze with an engine failed. The maneuver sequence is different for each pilot and counterbalanced across groups. Pilots know which maneuver will be trained and get feedback after each individual maneuver on the navigational display screen
6) PF and PNF complete Questionnaire 2, followed by off-site lunch

Phase III. Transfer Testing

7) Quasi-transfer to motion Scenario 1, Test 1
8) Quasi-transfer to motion Scenario 2, Test 1
9) PF and PNF complete Questionnaire 3
10) Quasi-transfer to motion Scenario 1, Test 2
11) Quasi-transfer to motion Scenario 2, Test 2
12) PF and PNF complete Final Questionnaire

Results

In this report, only control-input behavior and pilot-vehicle performance data are presented for the most critical phase of each maneuver. Moreover, the pilot-opinion data are restricted to the PF.

Pilot-Vehicle Performance and Control Behavior

Analysis. Only successful maneuvers were included in the analyses. As can be seen in Figure 1, the success rates of the two groups across maneuvers and phases are remarkably similar, with no significant group differences (all Fisher Exact p>.25). A successful maneuver is defined as one that was flown without crashing or scraping. To be considered a success, takeoff maneuvers must also be flown within four standard deviations (STD) of the mean maximum HDG deviation and bank. Landing maneuvers must be flown within four STDs of the mean maximum GS or LOC deviation. To calculate the success rate, missed approaches were excluded from the number of total maneuvers.

Each maneuver was analyzed separately using multivariate analyses of variance. MANOVA considers the set of interrelated dependent variables (DV) in combination. The number of DVs (and the loss of degrees of freedom and thus power) was reduced by choosing only one representative of DVs that correlated higher than r=.85.

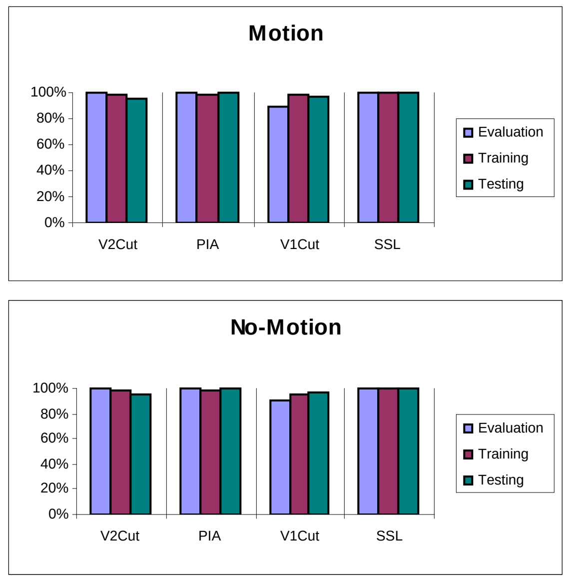

Figure 1. Success Rates by Phase and Maneuver

A two-way MANOVA examined the effect of the IVs Group (Motion vs. No-Motion) by Phase (Evaluation, Training, Transfer) on the remaining DVs. Interactions between Group and Phase were examined with two separate one-way MANOVAs on each group with Phase as the IV. A third set of MANOVAs examined the effect of Group and, where applicable, Trial separately for each phase, resulting in a one-way MANOVA for Evaluation and in two-way MANOVAs for Training (2 Groups by 3 Trials) and Transfer (2 Groups by 2 Trials). Because no effects of, or interactions with, Trial were found, no further analyses were required. Significant MANOVAs were followed up by univariate ANOVAs on the chosen variables. Differences between group means were analyzed with Bonferroni t tests. All analyses were performed in SAS™.

PIA. For the flight segment from Approach Fix to Decision Height, both overall Group and Phase effects were significant [Wilks' Lambda Λ=.71, $F(17,88)=2.10$ and Λ=.57, $F(34,176)=1.68$, respectively, $p<.05$]. There was no interaction between Phase and Group [Λ=.77, F(34,176)<1].

Follow-up univariate ANOVAs showed that the motion variable significantly affected seven of the 17 DVs (Table 1). The No-Motion group flew more precisely, with lower STDs around the desired HDG and LOC and lower bank STD. Interestingly, it achieved this performance with steadier wheel control inputs, i.e., lower root mean square (RMS) and reversals (number of times the wheel exceeds a ten-degree band around the neutral position). The No-Motion group, however, used higher pedal response BW (frequency below which the area under the pedal power spectrum density curve constitutes half of the total area). Group and Phase did not interact, showing that these effects persisted when the No-Motion group transferred to motion.

Variable	**Mean**		**Stats**
	Motion	**No-Motion**	F(1,104)>5.6
STD HDG	3.77 deg	2.84 deg	p=.0008
STD bank	3.35 deg	2.92 deg	p=.02
STD LOC	.55 dot	.36 dot	p=.0003
LOC exceedance	.25 dot	.09 dot	p<.0001
Wheel reversals	8.93	6.68	p=.02
RMS Wheel	2.39 deg	2.08 deg	p=.005
Pedal BW	.015 Hz	.025 Hz	p=.01

Table 1. PIA Univariate Results for Group

Ten of the DVs were significantly affected by Phase (Table 2). Both groups improved pilot-vehicle performance (HDG, bank, pitch, and LOC STD) and reduced control inputs (wheel and column reversals, RMS, and BW) progressively with phase.

Variable	**Mean**	**Differences**			**Stats**
		I-II	**II-III**	**I-III**	F(2,104)>3.5
STD HDG (deg)	3.32	1.27*	-.15	1.13*	p=.0006
STD bank (deg)	3.15	.66*	-.25	.41	p=.02
STD pitch (deg)	1.21	.28*	-.004	.27*	p=.02
STD LOC (dot)	.46	.21*	.004	.21*	p=.003
Wheel reversals	7.84	2.61	.94	3.55*	p=.008
Column reversals	4.57	2.03	1.20	3.22*	p=.03
RMS wheel (deg)	2.24	.46*	-.04	.42*	p=.001
Wheel BW (Hz)	.12	-.004	.03*	.02	p=.02
RMS column (in)	.51	.10*	.03	.13*	p=.003
Column BW (Hz)	.093	-.01	.03*	.02	p=.03

* indicates significant difference (p<.05)

Table 2. PIA Univariate Results for Phase

SL. Between the upward gust and touchdown, both overall Group and Phase effects were highly significant [Wilks' Lambda Λ=.62, F(20,95)=2.93 and Λ=.37, F(40,190)=3.08, p<.0005]. There was no interaction [Λ=.66, F(40,190)=1.10, p=.32].

Group (Table 3) significantly affected three of the 20 individual variables. The two groups appear to use different touchdown (TD) strategies regardless of phase: The Motion group landed softer, but longer (yet within the landing box). The No-Motion group again had higher pedal BWs.

Variable	**Mean**		**Statistics**
	Motion	**No-Motion**	F(1,114)>6.0
Pedal BW (Hz)	.04	.08	p=.0003
TD distance (ft)	1660	1435	p=.0007
TD descent rate (ft/min)	285	327	p=.02

Table 3. SL Univariate Results for Group

Both groups significantly improved on nine variables across phases (Table 4). Pilot-vehicle performance improved only for GS (deviation STD and GS deviation exceeding reference by .5 dot). Pilots significantly reduced their yaw activity (absolute yaw rate), wheel reversals, wheel and pedal RMS, and wheel, pedal, and column response BWs.

Variable	**Mean**	**Differences**			**Stats**
		I-II	**II-III**	**I-III**	F(2,114)>4.0
Yaw activity (deg/s)	.41	.07*	-.01	.06*	p=.01
STD GS (dot)	.56	.05	.04	.09*	p=.01
GS exceedance (dot)	.23	.10	.03	.12*	p=.02
Wheel reversals	8.07	1.84*	.82	2.66*	p=.002
RMS wheel (deg)	2.93	.46*	-.06	.40	p=.01
Wheel BW (Hz)	.15	.02	.07*	.09*	p<.0001
Column BW (Hz)	.10	.05*	.04	.08*	p<.0001
RMS pedal (in)	.40	.12*	-.04	.07	p=.0007
Pedal BW (Hz)	.06	-.03	.04*	.02	p=.001

* indicates significant difference (p<.05)

Table 4. SL Univariate Results for Phase

V_2 Cut. The segment analyzed for both takeoff maneuvers is between engine failure and 800ft above ground. Both maneuvers were highly affected by

Group and Phase, with a significant interaction.

The overall statistics for the V_2 cut are, Wilks' Lambda Λ=.66, F(15,99)=3.35 (Group) and Λ=.31, F(30,198)=5.23 (Phase), p<.0001; interaction Λ=.65, F(30,198)=1.59, p=.03.

The effect of Group on three of the 15 variables interacted with Phase: The Motion-trained group activated the pedal .76s slowerin response to the engine failure than the No-Motion group, which reacted within 2.34s [F(1,71)=8.69, p=.004], but this effect emerged only at Transfer. Also only during Transfer, the Motion group had a .29in higher column RMS than the group trained without motion, whose RMS was .86in [F(1,71)=10.22, p=.002]. Finally, the Motion group reversed the pedal .45 times more often than the No-Motion group reversing 1.05 times [F(1,38)=9.68, p=.004], but this effect disappeared during Training and did not reemerge.

Group regardless of Phase affected three variables (Table 5). The Motion groups demonstrated higher wheel activity (RMS, reversals) and lower pedal BW.

Variable	**Mean**		**Stats**
	Motion	**No-Motion**	F(1,113)>4.8
Wheel reversals	3.27	2.53	p=.0002
RMS wheel (deg)	6.97	5.44	p=.0001
Pedal BW (Hz)	.04	.05	p=.03

Table 5. V_2 Cut Univariate Results for Group

Seven variables were affected by Phase regardless of Group (Table 6). HDG STD and average failure-induced HDG deviation in the direction of the failed engine improved during Training, but the improvement didn't transfer. This was true also for reduced bank STD and wheel RMS. A pedal RMS decrease during training transferred, but incompletely. Increased wheel and pedal BWs during Training were transferred for pedal only.

Variable	**Mean**	**Differences**			**Stats**
		I-II	**II-III**	**I-III**	F(2,113)>3.1
STD HDG (deg)	3.66	.85*	-.96*	-.11	p=.006
Failure-induced HDG (deg)	5.47	5.40*	-4.03*	1.37	p<.0001
STD bank (deg)	5.69	1.54*	-1.71*	-.16	p=.01
RMS wheel (deg)	6.20	1.22*	-1.15*	.07	p=.01
Wheel BW (Hz)	.063	-.02 *	.01	-.01	p=.05
RMS pedal (in)	1.07	.19*	-.11	.08	p=.0008
Pedal BW (Hz)	.04	-.02*	.001	-.02*	p<.0001

* indicates significant difference (p<.05)

Table 6. V_2 Cut Univariate Results for Phase

V_1 Cut. The overall statistics for the V_1 cut are Wilks' Lambda Λ=.47, F(19,92)=5.47 (Group) and Λ=.41, F(38,184)=2.74 (Phase), p<.0001; interaction Λ=.56, F(38,184)=1.63, p=.02.

For six of the 19 DVs, the IVs Group and Phase interacted with each other (one only as a trend). As explained below, this shows that although the Motion group did have an advantage during Evaluation and frequently even during Training for these variables, this advantage disappeared at Transfer because the No-Motion group was immediately able to avail itself of the motion cues at Transfer.

Most importantly, a .39s faster pedal RT to the failure of the Motion compared to the No-Motion group (reacting after 1.53s) at Evaluation does point to an early alerting function of motion [F(1,34)=5.02, p=.03]. A significant RT difference of .29s persisted during practice [F(1,110)=14.72, p=.0002], when pilots were told which failure to anticipate, but disappeared when all pilots transferred to motion [F(1,74)=2.18, p=.14].

The faster pedal RT of the Motion group resulted in lower pedal RMS compared to the No-Motion group during Evaluation [.62 vs. .77in; F(1,34)=21.53, p<.0001] and Training [.60 vs. .70in, F(1,110)=51.64, p<.0001]. Pedal response BW of the Motion group was higher through Evaluation [.08 vs. .11HZ; F(1,34)=9.42, p=.004] and Training [.09 vs. .11HZ; F(1,110)=17.58, p<.0001]. Lower for the Motion group during evaluation only were yaw activity [.55 vs. .79 deg/s; F(1,34)=7.26, p=.01], pitch STD [5.63 vs. 6.40 deg; F(1,34)=7.21, p=.01] and perhaps HDG STD [2.28 vs. 3.04 deg; F(1,34)=4.34, p=.04].

Three variables showed Group differences regardless of Phase (Table 7). The Motion group controlled the wheel more aggressively (more reversals, higher RMS), but had fewer pedal reversals throughout.

Variable	**Mean**		**Statistics**
	Motion	**No-Motion**	F(1,110)>5.6
Wheel reversals	5.72	4.49	p=.0009
RMS wheel (deg)	3.99	3.41	p=.02
Pedal reversals	1.16	1.45	p=.0009

Table 7. V_1 Cut Univariate Results for Group

Two bank variables improved across Phase regardless of Group (Table 8). Failure-induced bank increased during training, but decreased at Transfer. Roll activity decreased at Transfer.

Variable	Mean	Differences			Stats
		I-II	II-III	I-III	F(2,110)>3.1
Failure-induced bank (deg)	1.20	-.44*	.54*	.10	p=.0001
Roll activity (deg/s)	1.36	.11	.10	.20*	p=.05

* indicates significant difference (p<.05)

Table 8. V_1 Cut Univariate Results for Phase

Power to Find an Effect

Many previous studies had few subjects or confounding within-group variables, reducing the power to find an effect (Boldovici, 1992). The effect sizes given in Table 9 show a sample of the smallest differences between Group (and Phase) detectable at a reasonable power of 1-β=.8 in the current study.

Maneuver	Measure	Effect Size	
		Group	Phase
V_2 Cut	STD bank	1.45 deg	1.72 deg
	STD HDG deviation	.75 deg	.89 deg
	Pedal RT	.43 s	.51 s
V_1 Cut	STD bank	.52 deg	.61 deg
	STD HDG deviation	.51 deg	.59 deg
	Pedal RT (s)	.18 s	.22 s
PIA	STD GS deviation	.07 dot	.08 dot
	STD LOC deviation	.14 dot	.17 dot
Sidestep Landing	STD GS deviation	.07 dot	.08 dot
	STD LOC deviation)	.11 dot	.13 dot

Table 9. Detectable Group and Phase Effect

Pilot Questionnaires

Pilots completed detailed questionnaires comparing the test simulator with the airplane or their company's simulator, as appropriate. The questions were asked separately for each control and maneuver. Pilots often took over half an hour to complete one questionnaire.

Acknowledging that pilots may not have experienced all maneuvers in the airplane, they were asked to base their comparisons on their expectation of how the airplane would respond in an identical situation. When considering the results, however, keep in mind the difficulty of the test maneuvers and the unusually light weight of the simulated airplane (550,000lbs).

The questionnaires were adapted from those designed for the first study and reviewed by six NASA pilots after flying the full test sequence. Scales ranged from 1 ("much worse than") to 7 ("much better than"), or from 1 (very different) to 4 (the same), as appropriate. Adverbs were adapted to the questions (worse, higher, less, harder, etc.). Many pilots volunteered additional comments in the space provided. A sample question is reproduced in Figure 2.

Acceptability

Compare the NASA 747-400 simulator to
the **last 747-400 SIMULATOR you have flown**
in terms of your **acceptance** based on your perception of the
presence or absence of deficiencies that might affect your flying.

Acceptability of the NASA 747-400 simulator was...

	1	2	3	4	5	6	7
	much worse than last simulator flown	moderately worse	slightly worse	just like the last simulator flown	slightly better	moderately better	much better than last simulator flown
Overall acceptability							
Please elaborate if acceptability is different from last simulator							

Figure 2. Sample PF Question

General Findings. Not all No-Motion pilots mentioned motion. Thirteen commented on the motion during Evaluation, but not all of them mentioned that motion was completely absent. The last six pilots all commented on the motion, when before no more than two in a row had made any comments. Three realized that motion was reduced during Training. Four never mentioned anything about motion throughout the experiment.

All pilots found the acceptability of the test simulator less than "slightly worse" than the one of their company simulator, but there were no effects of Group during any of the three phases [F(1,38)=.41, p=.53 (Evaluation); F(1,38)=.05, p=.82 (Training); F(1,38)=.05, p=.83 (Transfer)].

Physical and mental workload, although consistently perceived as less than "slightly higher" than in the airplane, remained also unaffected by Group across phases, with one exception: at Transfer, the Motion group perceived the mental workload as higher than the No-Motion group did [F(1,38)=4.29, p=.05]. This mainly was due to higher workload ratings by the Motion group compared to the earlier phases.

Physical comfort in the test simulator was rated as almost identical or "slightly higher" than in pilots' company simulator, with one notable trend [F(1,38)=3.73, p=.06]: the No-Motion pilots apparently didn't always like the transfer to motion.

When pilots were asked, for each maneuver, whether there were any "other cues" that were different from the airplane, the Motion group generally found "other cues" less different from the airplane than the No-Motion group [F(1,38)=3.54, p=.07] during Evaluation and [F(1,38)=4.78, p=.04] during Training. Neither group ever found them more than slightly different. As would be expected, this effect disappeared at Transfer to all motion.

Opinions on Control. Regardless of Group or Phase, pilots found their control strategy to be only less than "slightly different" from the one they adopt in the airplane. Both groups also found that the controls were less than "slightly more sensitive" than in the airplane. The control feel was generally perceived as less than "slightly lighter," more "lighter" by the No-Motion group than the Motion group during Training [F(1,38)=3.99, p=.05] and even at Transfer [F(1,38)=4.01, p=.05]. Handling qualities were consistently rated as less than slightly worse than in the airplane, however more "worse" by the Motion pilots during Training [F(1,38)=4.82, p=.03].

Discussion and Conclusion

All maneuvers showed significant effects of Group and Phase. For the landing maneuvers, Group effects occurred regardless of Phase (and phase effects regardless of Group), indicating that whatever differences existed, they persisted even when the No-Motion group transferred to motion. Consistent with previous work, the No-Motion group generally had higher control-input BWs than the Motion group, but lower control activity (Hess, Malsbury, & Atencio, 1993). Both groups improved equally across phases, reducing their control inputs and, mainly for the PIA, improving pilot-vehicle performance.

For the PIA, the motion stimulation may have had a distracting effect, enticing the Motion group to higher control-input behavior to the detriment of pilot-vehicle performance. For the SL, there was a potential landing strategy difference between groups, with the Motion group using the vertical acceleration cues to arrest sink rate, resulting in softer but longer landings than the ones of the No-Motion group. For both maneuvers, the No-Motion group did not change strategy when exposed to motion cues.

For the takeoff maneuvers, motion effects were contingent upon phase. Curiously, for the V_2 cut, the majority of Group differences emerged only at Transfer, such as the slower pedal RT and the higher column RMS of the Motion compared to the No-Motion group. This may be due to fatigue, which for the No-Motion group was counteracted by the appearance of motion stimulation (see also lower workload perception of the No-Motion group at Transfer). Across phases, the No-Motion group again displayed lower control activity but higher BWs, although this effect was lost during training for pedal reversals. For both groups, none of the pilot-vehicle performance improvements and only one of the control-input reductions during Training transferred, again pointing to fatigue.

The V_1 cut results confirm that it may be one of the most diagnostic maneuvers for a need of motion, as assumed in the first study. Although the motion advantages found were small, they do demonstrate the predicted early alerting function of motion. The disappearance of any Motion group advantage except fewer pedal reversals at Transfer, however, would suggest that recurrent training with motion may not be required to take advantage of this function. In fact, the V_1 cut is the only maneuver where the overall Group effect found in separate MANOVAs for Evaluation and Transfer disappeared at Transfer. Higher wheel activity of the Motion group compared to the No-Motion group persisted throughout all phases, however. Improvements across phases regardless of Group were restricted to bank variables.

Many of the motion effects found in this study using a B747-400 simulator are on pilots' control behavior, but they don't always translate into pilot-vehicle performance differences. The operational relevance of the effects that were found on pilot-vehicle performance needs to be further examined, especially for airplanes that are more agile than the B747-400.

With the exception of the different strategies for the SL, none of the evaluation or recurrent training advantages of motion found persisted when both groups had motion, but some of the evaluation/training advantages of the absence of motion stimulation did (especially for the PIA).

Contingent upon discussion of the operational relevance of the effects found, a preliminary conclusion based on the Group differences found during Evaluation and Transfer is that motion may be required even for recurrent evaluation of airline pilots. No beneficial effect of motion on recurrent training, however, was found. In fact, for maneuvers where motion cues do not serve an alerting function, training without motion may lead to steadier control strategy resulting in better pilot-vehicle performance in the airplane than motion training. The differential effects of motion on the test maneuvers confirm that the effect of motion depends on the maneuver characteristics, and that the criteria found in the literature may be valid predictors for which maneuvers may or may not require motion.

This research has found many effects of hexapod motion that have not been previously shown in the airline pilot training and evaluation context. The emergence of an early alerting effect of motion during the V_1 cut with enhanced lateral acceleration cues shows the importance of the quality of motion cues. Results of this and the previous hexapod motion research should assist the FAA in determining future research directions in the effort to develop improved motion standards.

Acknowledgments

The FAA Human Factors Division, AAR-100, supported this work. We greatly appreciate the insights of our FAA Program Manager Dr. Eleana Edens. We also thank the excellent experiment team at NASA Ames and the participating pilots who volunteered their expertise.

References

Boldovici, J.A. (1992). *Simulator Motion* (ARI Technical Report 961). Alexandria, VA: US Army Research Institute.

Bray, R.S. (1972). Initial operating experience with an aircraft simulator having extensive lateral motion. NASA TM X-62155.

Bürki-Cohen, J., Soja, N. N., Longridge, T. (1998). Simulator platform motion—The need revisited. *International Journal of Aviation Psychology, 8 (3),* 293-317.

Go, Tiauw H.; Bürki-Cohen, J., Soja, N.N. (2000): *The Effect of Simulator Motion on Pilot Training and Evaluation* (AIAA 00-4296). American Institute of Aeronautics and Astronautics.

Gundry, J. (1976). *Man and motion cues*. Paper presented at the Third Flight Simulation Symposium, London, UK.

Hall, J.R. (1989). *The need for platform motion in modern piloted flight training simulators* (Technical Memorandum FM 35). London, UK: Royal Aerospace Establishment.

Hess, R.A., Malsbury, T., Atencio, Jr., A. (1993). Flight Simulator Fidelity Assessment in a Rotorcraft Lateral Translation Maneuver. *Journal of Guidance, Control, and Dynamics, 16 (1),* 79-85.

Longridge, Thomas, Bürki-Cohen, J., Go, T.H., Kendra, A.J. (2001). Simulator fidelity considerations for training and evaluation of today's airline pilots. *Proceedings of the 11th International Symposium on Aviation Psychology*. Columbus, OH: The Ohio State University Press.

Mikula, J., Chung, W, and Tran, D. (1999). *Motion fidelity criteria for roll-lateral translational tasks* (AIAA 99-4329). American Institute of Aeronautics and Astronautics.

Sinacori, J.B. (1977). The determination of some requirements for a helicopter flight research simulation facility. NASA CR-152066.

Schroeder, J.A. (1999) Helicopter flight simulator motion platform requirements. NASA TP-208766.