Objective Motion Cueing Criteria Investigation Based on Three Flight Tasks

Peter M.T. Zaal*
San Jose State University
NASA Ames Research Center
Moffett Field, CA

Jeffery A. Schroeder†
Federal Aviation Administration
Moffett Field, CA

William W. Chung‡
Science Applications International Corporation
NASA Ames Research Center
Moffett Field, CA

Abstract—This paper intends to help establish fidelity criteria to accompany the simulator motion system diagnostic test specified by the International Civil Aviation Organization. Twelve airline transport pilots flew three tasks in the NASA Vertical Motion Simulator under four different motion conditions. The experiment used three different hexapod motion configurations, each with a different tradeoff between motion filter gain and break frequency, and one large motion configuration that utilized as much of the simulator's motion space as possible. The motion condition significantly affected 1) pilot motion fidelity ratings, and sink rate and lateral deviation at touchdown for the approach and landing task, 2) pilot motion fidelity ratings, roll deviations, maximum pitch rate, and number of stick shaker activations in the stall task, and 3) heading deviation after an engine failure in the takeoff task. Significant differences in pilot-vehicle performance were used to define initial objective motion cueing criteria boundaries. These initial fidelity boundaries show promise but need refinement.

1 Introduction

This paper addresses the question "what fidelity criteria should accompany the diagnostic that quantifies simulator motion cues?" Actuator hardware and software algorithms determine motion cues. Today, during a simulator qualification, engineers objectively measure only the hardware [1]. Pilot inspectors subjectively assess the software, but attributing fidelity weaknesses to that software is challenging in a simulator's integrated cueing environment. ICAO 9625 has an Objective Motion Cueing Test (OMCT) that evaluates the software and hardware together, but it lacks accompanying fidelity criteria [2]. Hosman has documented OMCT results for a statistical sampling of eight simulators [3], which is useful, but having validated criteria would be an improvement.

Sinacori suggested simple criteria [4], [5], which reasonably agree with much of the literature [6]–[15]. These criteria often necessitate motion displacements greater than what most training simulators can provide. While some of the previous work has used transport aircraft in their studies, the majority used fighter aircraft or helicopters. Those that used transport aircraft considered degraded flying characteristics. So, it would be fair to say that earlier criteria lean more towards being sufficient, rather than necessary, criteria. Considering the prevalence of 60-inch, six-legged hexapod training simulators, a more relevant question today is "what are the necessary criteria that can be used with the ICAO diagnostic?"

This study adds to the literature as follows. First, it examines well-behaved transport aircraft characteristics but in challenging tasks. Second, it uses the Vertical Motion Simulator (VMS), the world's largest vertical displacement simulator. This allows inclusion of a relatively large motion condition, much larger than what a typical training simulator can provide, so that variations in the conditions can be compared against something closer to flight. This makes sure some perspective is kept on potentially small motion differences across the configurations. Third, a sufficiently large pilot pool added statistical reliability to the results.

2 Flight Tasks

The study had three tasks: a sidestep approach and landing, a high-altitude stall recovery, and an engine-out after takeoff. The tasks were similar to three of the tasks used in [16]. The assumption was that varying the motion would affect how pilots perform these challenging tasks. Fig. 1 shows the flight cards for the three tasks, which the following sections describe.

2.1 Approach and Landing with Sidestep

This task began at an altitude of 1,250 ft on a -3 deg glideslope approach to RWY 28R at San Francisco International Airport. Moderate turbulence existed throughout. After breaking out of the cloud ceiling at 1,100 ft, air traffic control instructed the pilots to sidestep to RWY 28L. Pilots tried to maintain a -3 deg glideslope during the approach followed by landing within a 750-ft-long box with a sink rate of 6 ft/s or less. An audio call-out began at a main gear height of fifty feet and repeated in decrements of 10 ft until touchdown.

This task evaluated if motion affects 1) lateral-directional control in the sidestep, 2) speed and flightpath control along the glideslope, and 3) touchdown performance.

* *Research Engineer, Human Systems Integration Division, NASA Ames Research Center, Moffett Field, CA, 94035; peter.m.t.zaal@nasa.gov.*

† *Chief Scientific and Technical Advisor for Flight Simulation Systems, Federal Aviation Administration, Moffett Field, CA, 94035; jeffery.schroeder@faa.gov.*

‡ *Simulation Engineer, Simulation Laboratories, NASA Ames Research Center, Moffett Field, CA, 94035; william.w.chung@nasa.gov.*

(a) Approach and landing with sidestep

Task: approach to SFO RWY 28R with sidestep to landing on RWY 28L
Initial Condition: 1,250 ft altitude at V_{ref}+5 on GS and LOC for 28R
Configuration: gear down, flaps 30 deg
Ceiling/visibility: 1,100 ft ceiling, 10-mile visibility
Wind: 230/15 **Turbulence:** moderate **Gusts:** none

Procedure:
1) Track the GS and LOC to SFO RWY 28R maintaining 141 KIAS
2) Perform sidestep to RWY 28L at ATC command
3) Continue visual to RWY 28L maintaining GS
4) Flare and touchdown 750 –1,500 ft from the threshold

Desired performance:
1) Deviation from 141 KIAS within $\pm$5 kts until 200 ft altitude
2) Glideslope deviation within $\pm$ one dot until 200 ft altitude
3) Longitudinal touchdown 750 –1,500 ft from threshold
4) Lateral touchdown $\pm$14 ft from centreline
5) Sink rate at touchdown $\leq$6 ft/s

(b) High altitude stall recovery

Task: recover from a high altitude stall
Initial Condition: level at FL 410 and 210 KIAS
Configuration: clean
Ceiling/visibility: CAVU
Wind: none **Turbulence:** moderate **Gusts:** none

Procedure:
1) Retard throttle to idle
2) Roll left to a 20 deg bank angle
3) Pull up to decelerate at approximately 4 kt/s
4) Continue deceleration through stick shaker until a sink rate develops
5) Apply nose down pitch, roll as needed, power as needed to return to steady-state flight

Desired performance:
1) Maintain bank angle within $\pm$5 degs about 15 during entry
2) Proper stall recovery procedure (push, roll, thrust, stabilized flight)
3) Recovers smoothly to speed >210 KIAS and positive rate of climb
4) Recover without exceeding airplane's limitations
5) Oscillatory load factor peaks between 0.75 and 1.25 g
6) During recovery, no more than one additional stick shaker activation

(c) Engine out after takeoff

Task: recover from engine out on takeoff
Initial Condition: at takeoff position on RWY 28R
Configuration: gear down, flaps 10 deg
Ceiling/visibility: CAVU
Wind: none **Turbulence:** none **Gusts:** none

Procedure:
1) Advance throttles to takeoff thrust (90% N_1)
2) Maintain centreline
3) Rotate at V_r=128 kts, pitch to 10 degs and establish speed of V_2+10
4) Maintain heading and speed after single engine failure

Desired performance:
1) Desired heading $\pm$5 degs
2) Desired airspeed $\pm$5 kts
3) Bank approximately 5 degs towards operating powerplant

Fig. 1. Experiment flight cards.

2.2 High Altitude Stall Recovery

The task started during cruise at 210 kts ($M = 0.75$) at an altitude of 41,000 ft. Instructions were for the pilots to initiate a self-induced stall by setting the throttles to idle, rolling 20 degs left and pitching approximately 15 degs nose up, decelerating through the stall warning until a negative sink rate was developed (as the aircraft model did not have a pronounced pitch break). To recover from the stall, pilots had to follow the correct recovery sequence of reducing the angle-of-attack (by pitching to approximately 15 deg nose down), levelling the wings, and applying full throttle until establishing a safe positive climb rate. The task called for the pilots to pull the nose up gently and smoothly so as not to activate the stall warning during the recovery. Moderate turbulence was present.

This task evaluated if motion affects the recovery performance by helping a pilot damp the flight path response, as well as stabilize the progressively less-stable roll dynamics near stall.

2.3 Engine Out after Takeoff

During the takeoff roll, the task asked for the pilots to rotate at 128 kts, establish a 10 deg nose up pitch attitude, and maintain a climb speed of 145 kts. Either the left or right engine failed randomly at a random altitude after liftoff but below 100 ft. After identifying the failed engine, the pilots needed to apply near full rudder pedal towards the good engine, roll approximately 5 deg into the direction of the good engine to maintain the desired track, and modulate the remaining thrust to maintain speed.

This task evaluated if motion helped pilots detect the failed engine promptly and recover using appropriate lateral-directional control.

2.4 Aircraft Model

The experimenters modified an existing mid-size twin-engine commercial transport aircraft model with a gross weight of 185,800 lbs for this study. The enhancements allowed for a more representative aircraft response in the stall task. Most significantly, modifications to the lift coefficient as a function of angle of attack allowed for typical transport post-stall characteristics. Modifications to the rolling moment coefficient due to roll rate allowed for satisfactory representation of reduced roll stability near stall.

The simulated aircraft had six degrees of freedom and a two-crew cockpit . The aircraft model included a landing gear model, allowing it to taxi, and takeoff and land. The aircraft could operate in the full flight envelope up to a maximum cruising altitude of 42,000 ft.

3 Motion Tuning

Pilots performed each flight task with four different motion configurations: three hexapod motion configurations (HEX-L, HEX-M, and HEX-H) and a large motion configuration (LMOT), see Table 1. The four motion configurations were optimized for each of the three tasks giving 12 motion configurations in total.

The experiment used the standard VMS motion algorithm and hardware for all motion configurations (Fig. 2).

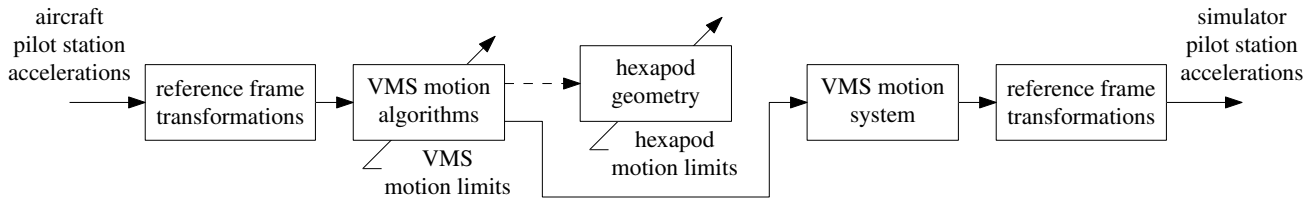

Fig. 2. Simulator motion block diagram.

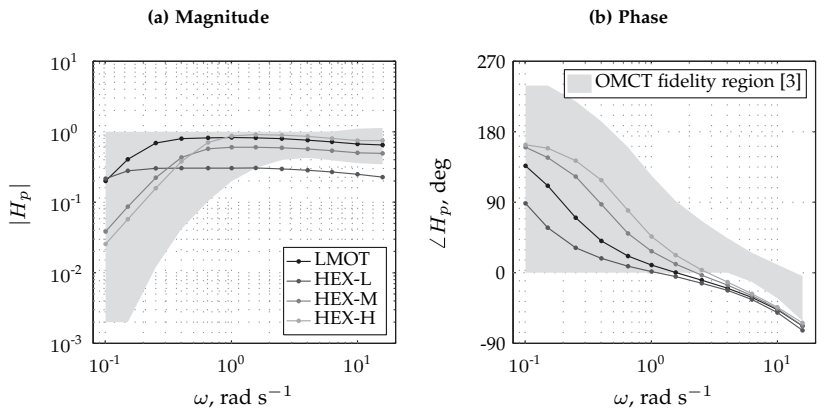

Fig. 3. OMCT roll responses for the stall task.

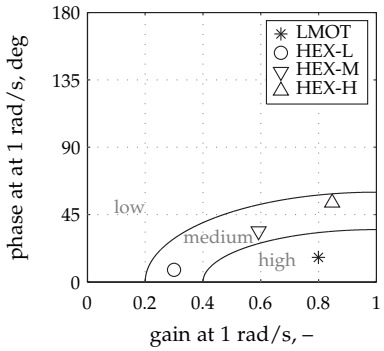

Fig. 4. Stall task roll Sinacori diagram.

TABLE 1. Motion configurations.

Label	Description
HEX-L	low-gain/low-break-frequency hexapod motion
HEX-M	medium-gain/medium-break-frequency hexapod motion
HEX-H	high-gain/high-break-frequency hexapod motion
LMOT	high-gain/low-break-frequency large motion

The equivalent time delays of the VMS motion system for the pitch, roll, yaw, longitudinal, lateral, and vertical axes are, 47, 68, 48, 50, 69, and 67 ms, respectively. More details about the motion algorithm are provided in [17]. A hexapod algorithm took the accelerations from the VMS motion algorithm and calculated the leg extensions of a typical hexapod motion platform with 60-inch legs. The algorithm accounted for the inherent constraints when trying to simultaneously move a hexapod in several degrees of freedom. Mass and inertia effects of the hexapod system were not taken into account.

A simulation model of the motion logic in Fig. 2 allowed the tuning of the motion configurations offline. The various OMCT fidelity regions (determined by a statistical sample of eight simulators) and the Sinacori criteria guided the motion tuning process [2], [4]. The main tuning goal was to cover a reasonable area of the OMCT fidelity region, while adhering to the hexapod motion limits. Tuning the parameters of motion washout filters constitutes a tradeoff between magnitude and phase error of the simulator motion with respect to the aircraft model motion. A higher gain (smaller magnitude error) generally requires a higher break frequency (larger phase error at lower frequencies). The three hexapod motion configurations employed this tradeoff. LMOT-L had a low gain and consequently a low break frequency, LMOT-H had a high gain and high break frequency, and LMOT-M had a gain and a break frequency approximately halfway in between the low and high gain and break frequency values. The large motion configuration, LMOT, had the highest possible gain and lowest break frequency, while adhering to the VMS motion limits.

As the three different tasks had different motions in each degree of freedom, the four motion configurations were optimized for each task. However, the tuning approach for each of the four motion configurations was the same for all tasks; for example, HEX-L was tuned to provide low-gain/low-break-frequency hexapod motion in all three tasks. The inputs and outputs of the block diagram in Fig. 2 define the OMCT responses for every degree of freedom.

The OMCT was performed for all 12 motion configurations of the experiment. Fig. 3 provides an example of the OMCT responses for the roll degree of freedom in the high-altitude stall task. Fig. 4 depicts the corresponding Sinacori diagram. Note that for this degree of freedom and task, all the hexapod motion configurations have the same predicted fidelity in the Sinacori diagram, in this case medium fidelity. The large motion configuration has the highest fidelity.

Fig. 5 shows time traces of specific forces and angular accelerations of important degrees of freedom for each task to illustrate the differences between motion configurations. The specific forces or angular accelerations from the aircraft model are also shown. Fig. 5a depicts the vertical specific force at the pilot station for a representative run of the sidestep task near touchdown. The differences in gain between the motion configurations can be seen just after touchdown. Fig. 5b depicts the yaw acceleration in the engine out on takeoff task near the engine failure. Note that large yaw accelerations occur after the engine failure. The simulator yaw accelerations for LMOT are closest to those from the aircraft model. HEX-H produces yaw accelerations with a high gain, but with large phase errors, while HEX-L produces yaw accelerations with a low gain, but with

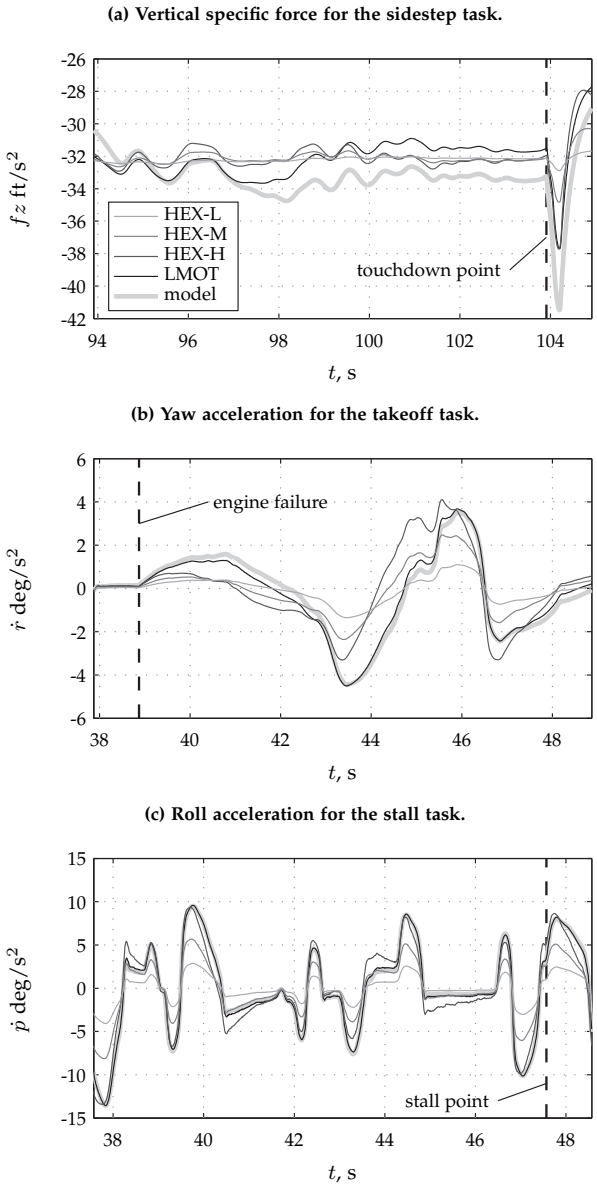

Fig. 5. Comparison of motion filter accelerations against model accelerations.

small phase errors. HEX-M is halfway in between HEX-L and HEX-H. Similar observations can be made for the roll acceleration in the high-altitude stall recovery task near the stall (Fig. 5c).

4 Experiment Design

4.1 Method

4.1.1 Independent Variable

The only independent variable was the motion condition. Each pilot flew all of the motion conditions for each task. Four different motion configurations were used for each of the three tasks: HEX-L, HEX-M, HEX-H, and LMOT.

4.1.2 Apparatus

The experiment used the VMS with the transport cab (T-CAB) (Fig. 6) [17]. Pilots flew from the left seat; the instructor occupied the right seat. A primary flight display (PDF) was positioned in front of each pilot (Fig. 7). The navigation displays were positioned next to the PFDs towards the centre of the cab. An additional display in the centre of the cab showed the engine parameters. This display changed to showing task performance after each run.

A column and wheel controlled aircraft pitch and roll, respectively. A thumb switch on the wheel controlled elevator trim. Conventional rudder pedals controlled the rudder deflection in the air and nose wheel rotation on the ground. Two throttle levers controlled the power of two engines. The instructor pilot configured flaps and gear with representative controls before each task.

The PFD symbology was similar to that on a Boeing 777 (Fig. 8). Speed and heading bugs indicated the desired speed and heading for each task. In addition, typical symbology on the speed tape indicated the minimum and maximum speeds, and V-speeds. The PFD also depicted conventional localizer and glideslope error indicators. A green speed trend vector originating from the indicated airspeed showed what the airspeed would be in 10 s. The control columns in the cockpit had stick shakers to warn pilots of an impending stall. The activation of the shakers occurred simultaneously when the minimum speed tape, also known as the barber pole, coincided with the indicated airspeed.

The collimated out-the-window display of T-CAB consisted of six cathode ray tube monitors casting a high-quality image on six spherical mirrors. The mirrors formed a dome-like section providing a continuous field-of-view image to both pilots. The out-the-window visual had a 220° horizontal field of view and a 28° vertical field of view (10° up and 18° down). A Rockwell-Collins EPX5000 computer image generator created the out-the-window visual scene. The visual system equivalent time delay was 62 ms [17]. This was in line with the equivalent time delays of the motion system (Section 3).

4.1.3 Procedures

The pilot pre-briefing explained the purpose of the experiment, the procedures, conditions, and performance criteria for each flight task. Pilots were informed they would perform each task with four different motion configurations and that the configurations would be presented randomly. However, no specifics about the motion configurations were given.

Pilots performed all three tasks, each with the four different motion configurations. Each motion configuration was repeated eight times for a total of 32 runs per task. The motion configurations were presented randomly. The first 12 runs were used as training runs in which the four motion conditions were presented three times each. The remaining 20 runs were used for data analysis. A randomized Latin square determined the order of the tasks and motion configurations. All runs for a particular task were performed in a single session. Each task session lasted one-and-a-half to two hours, and breaks were taken in between sessions.

Before each task, the instructor pilot reviewed the procedures, performance criteria, the relevant controls, and displays. The instructor evaluated a pilot's performance after each run using task performance information displayed in the cab. An experiment observer in the control room verified the evaluation. After completing each run, participant

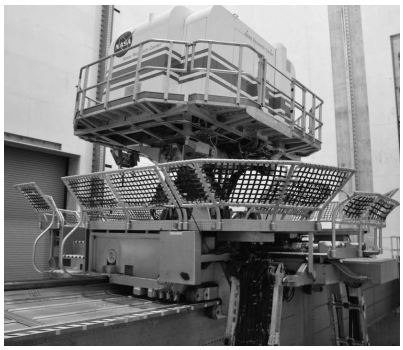

Fig. 6. Vertical Motion Simulator.

Fig. 7. Cockpit setup.

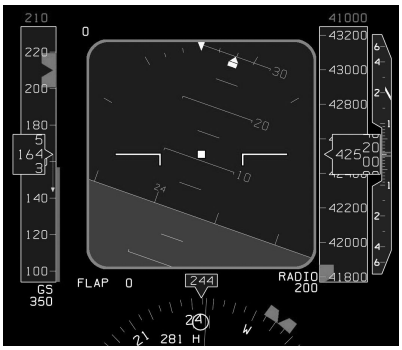

Fig. 8. Primary flight display.

pilots rated the motion with two different rating scales (Section 4.1.5).

4.1.4 Participants

Twelve experienced airline transport pilots participated. Each had a B757/B767 type rating, as the aircraft model used in the simulation was of an aircraft similar to a B757 in terms of configuration, size, and weight. All pilots had experience on many other aircraft types. The average number of flying hours pilots had on a B757/B767 was 4,503 with a standard deviation of $\pm$4,463. The average number of flying hours on other aircraft types was 8,579 with a standard deviation of $\pm$4,458. All pilots were male and eight were captains.

4.1.5 Dependent Measures

Six subjective dependent measures and 11 objective dependent measures were recorded and analysed. Pilots rated the simulator motion in each run with two different ratings in all three tasks. For the first rating, pilots assigned a specific fidelity level out of three levels on a motion fidelity scale (MFS):

High (H): motion sensations like those of flight.

Medium (M): motion sensations are noticeably different from flight, but not objectionable.

Low (L): motion sensations are noticeably different from flight and objectionable.

For the second rating, pilots answered the following question: "How would you rate motion fidelity for this condition with respect to what you would expect from the visual motion?" by drawing a vertical marker on an analogue fidelity scale (AFS). The analogue scale ranged from 0% to 100%.

Several objective measures determined the effect of motion on task performance. Many measures related directly to the performance criteria (Fig. 1). For the approach and landing with sidestep, the RMS of the glideslope, RMS_{gs}, and speed deviation in the approach, RMS_V, were calculated. Calculations for these variables used data from the last 60 s before reaching the decision altitude of 200 ft. When the main gear touched the runway, data captures occurred for the sink rate, $\dot{h}_{td}$, and the longitudinal and lateral deviations, Δx_{td} and Δy_{td}, from the desired touchdown point.

For the stall recovery, the RMS roll error, RMS_ϕ, applied to the 30 s before the start of the stall recovery. The maximum pitch rate deviation, q_{max}, and the number of stick shakers, N_s, applied to the stall recovery segment. Pitch rate was substituted as a partial surrogate to analyse load factor oscillations in the stall recovery, as it was more suitable for analysing oscillatory behaviour. Finally, for the engine out on takeoff task, the RMS of the heading and speed deviations after the engine failure, RMS_ψ and RMS_V, used data from the last 15 s after the engine failure. The reaction time of the initial pedal input after the engine failure, t_p, was from the time of the engine failure to the time when the pedal input was 10% of the maximum input.

4.2 Hypotheses

The null hypothesis was assumed. Consistent with this hypothesis, the three hexapod motion configurations were designed to deliberatively trade off gain and naturally frequency while trying to stay within the same fidelity region as predicted by the Sinacori criteria. The large motion configuration used as much of the VMS motion space as possible, generally leading to a higher predicted motion fidelity than the hexapod conditions. This higher-fidelity condition was included as a quasi-truth case to lessen the tendency to subjectively exaggerate small differences between hexapod conditions.

5 Results

This section presents the mean results of the 12 pilots for the approach and landing with sidestep and engine out on takeoff tasks. The mean results of 11 pilots are presented for the high-altitude stall recovery task, as one pilot did not complete this task due to technical difficulties. For every pilot, data from the last five runs for each task and motion configuration were averaged. Error-bar plots present the continuous-interval dependent measures, with means and 95% confidence intervals for each motion condition. Bar plots present the ordinal dependent measures, with the number of occurrences for each dependent measure level and the median for each motion condition. Wherever appropriate, grey dashed lines depict the task performance criteria.

A repeated-measures analysis of variance (ANOVA) detected possible statistically significant differences in the continuous interval dependent measures. A Friedman test was used for the ordinal dependent measures. For the repeated-measures ANOVA to produce accurate results, the data must meet three assumptions: 1) there should be no significant outliers, 2) data should be approximately normally

TABLE 2. Summary of statistical test results.

Measure	df	F/χ^2	p	η^2	Sig.
Sidestep Task					
MFS	3.0	16.399	0.001	–	**
AFS	3.0,33.0	1.746	0.177	0.137	–
RMS_{gs}	3.0,33.0	2.352	0.090	0.176	–
RMS_V	3.0,33.0	1.490	0.235	0.119	–
$\dot{h}_{td}$	3.0,33.0	12.881	0.000	0.539	**
Δx_{td}	3.0,33.0	0.995	0.407	0.083	–
Δy_{td}	3.0,33.0	2.265	0.099	0.171	*
Stall Task					
MFS	3.0	8.117	0.044	–	**
AFS	3.0,30.0	1.428	0.254	0.125	–
RMS_ϕ	1.6,15.6	3.350	0.070	0.251	*
q_{max}	1.4,14.3	3.052	0.091	0.234	*
N_s	3.0	9.111	0.028	–	**
Takeoff Task					
MFS	3.0	1.200	0.753	–	–
AFS	1.6,17.7	0.105	0.957	0.009	–
RMS_ψ	3.0,33.0	4.642	0.008	0.297	**
RMS_V	3.0,33.0	1.316	0.286	0.107	–
t_p	3.0,33.0	2.136	0.114	0.163	–

** = significant ($p < 0.05$)
* = marginally significant ($0.05 \leq p < 0.1$)
– = not significant ($p \geq 0.1$)

distributed for each level of the independent variable, and 3) variances of the differences between all combinations of levels of the independent variable must be equal (i.e., assumption of sphericity).

Box plots identified outliers, normality was checked using the Shapiro-Wilk test, and Mauchly's test checked the assumption of sphericity. Few outliers were detected, so they were kept in the data analysis. Data were generally normally distributed. In the few cases they were not, we did not correct for it, as a repeated-measures ANOVA is fairly robust to deviations from normality. When the assumption of sphericity was violated, the degrees of freedom of the ANOVA were corrected using the Greenhouse-Geisser adjustment.

Post-hoc tests with Bonferroni adjustment were performed for pairwise comparisons of motion conditions when the ANOVA indicated an overall significant difference. All statistical tests had a significance level of 0.05. Table 2 gives a summary of the statistical test results for all the repeated measures (Section 4.1.5). In this table, df are the degrees of freedom, F or χ^2 is the test statistic for the ANOVA or Friedman test, respectively, p is the probability of observing an effect, and η^2 is the partial eta-squared, indicating sample effect size. Next, details of the motion fidelity rating results are presented, followed by the task performance results.

5.1 Motion Fidelity Ratings

Fig. 9 shows the MFS ratings. The bars indicate the number of times pilots gave a certain rating. Note that 60 ratings (12 pilots $\times$ 5 runs) should be given for each motion condition in the sidestep and takeoff tasks, and 55 ratings (11 pilots $\times$ 5 runs) for each motion condition in the stall task. However, HEX-L for the sidestep task and LMOT for the stall task contain only 59 and 54 ratings, respectively, as pilots forgot to rate one run in each condition. For the approach and landing with sidestep task, the MFS ratings were significantly different between the different motion conditions, as indicated by the Friedman test. The post-hoc analysis revealed a marginal difference between the ratings for HEX-L and HEX-H ($p = 0.083$). More lower ratings were given for HEX-H compared to HEX-L (Fig. 9a).

Statistically significant differences were found in MFS ratings between motion conditions for the high-altitude stall recovery task. Pairwise comparisons did not reveal significant differences between specific motion condition pairs. However, Fig. 9b shows that more low fidelity ratings were given for the HEX-H and LMOT motion configurations. No significant differences were found in MFS ratings between motion conditions for the engine out on takeoff task (Fig. 9c). Comparing the different tasks, pilots generally provided higher MFS ratings for the high-altitude stall and engine out on takeoff tasks.

Fig. 10 depicts the AFS rating results for each task. No statistically significant differences were found in AFS ratings between motion conditions for any of the tasks. However, it can be observed that ratings for the stall and takeoff tasks are higher (around 80%) compared to the sidestep task (around 60%). This is consistent with the MFS ratings .

5.2 Task Performance Results

5.2.1 Approach and Landing with Sidestep

Fig. 11 provides the task performance results for the approach and landing with sidestep task. Figs. 11a and 11b depict the RMS of the glideslope and speed deviation in the approach, respectively. No statistically significant differences between motion conditions were found for both variables . Note that pilots were able to meet the glideslope-deviation performance requirement easily, as they deviated around 0.2 dots from the glideslope on average. However, the average airspeed deviation was above 5 kts, indicating that pilots had difficulties meeting the speed-deviation performance requirement. This was mainly caused by the significant airspeed perturbations introduced by the atmospheric turbulence.

The repeated-measures ANOVA revealed that the sink rate at touchdown was significantly different between motion conditions. A post-hoc analysis with Bonferroni adjustment revealed that $\dot{h}_{td}$ in the LMOT condition was significantly different from $\dot{h}_{td}$ in the HEX-L ($p < 0.001$) and HEX-H ($p = 0.009$) motion conditions, and marginally different from $\dot{h}_{td}$ in the HEX-M condition ($p = 0.052$). No effect of sinkrate at touchdown was found between the hexapod motion conditions. The average sinkrate at touchdown in LMOT was well below the maximum desired value of 6 ft/s; however, the average value in the hexapod motion conditions was around the maximum desired value.

The longitudinal deviation from the desired touchdown point was not significantly different between motion conditions. Pilots generally touched down long, close to the maximum allowable 1,500 ft from the runway threshold. The ANOVA found that the lateral deviation from the desired touchdown point was marginally different between motion conditions. Post-hoc analysis found that Δy_{td} was statistically different between the HEX-H and LMOT motion conditions only ($p = 0.080$).

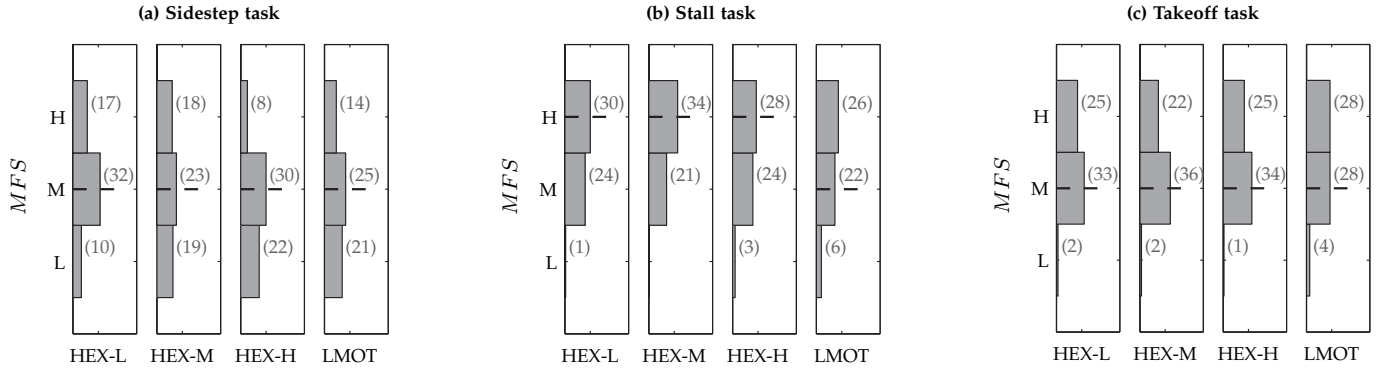

Fig. 9. Pilot motion fidelity scale ratings.

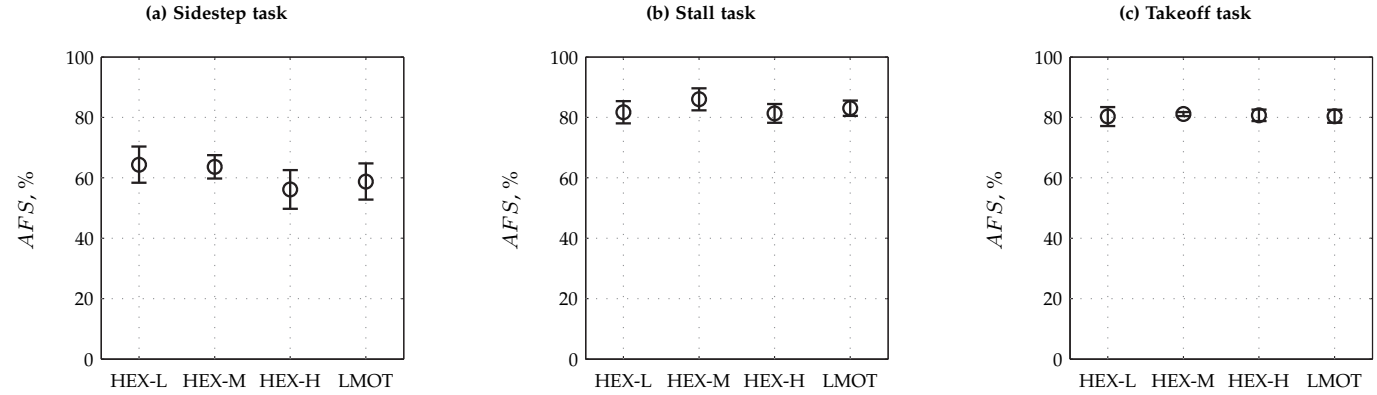

Fig. 10. Pilot analogue motion fidelity scale ratings.

5.2.2 High-Altitude Stall Recovery

Fig. 12 presents the task performance results for the high-altitude stall recovery task. Fig. 12a depicts the RMS roll deviation during the approach to stall. The RMS roll deviation was marginally different between the motion conditions. Post-hoc analysis with Bonferroni adjustment revealed that the RMS roll deviation was significantly different between HEX-H and LMOT ($p = 0.029$). The RMS roll deviation was not significantly different between the hexapod motion configurations. Note that pilots met the roll deviation performance requirement without difficulty.

Fig. 12b depicts the maximum pitch rate deviation during the stall recovery. Overall, this variable was marginally different among motion configurations. Post-hoc analysis found a marginal difference between HEX-L and HEX-M ($p = 0.097$) and a significant difference between HEX-M and HEX-H ($p = 0.008$). No significant differences were found between other motion condition pairs.

The bars in Fig. 12c indicate the number of times a pilot activated 0, 1, or 2 additional stick shakers during the stall recovery. A Friedman test revealed that the number of additional stick shakers during the stall recovery was significantly different between motion conditions. Post-hoc analysis did not reveal significant differences between specific motion condition pairs. However, Fig. 12c indicates that LMOT is the only motion configuration where none of the pilots activated additional stick shakers, while in the hexapod motion conditions seven or eight times at least one additional shaker was triggered.

5.2.3 Engine Out on Takeoff

Fig. 13 provides the task performance results for the engine out on takeoff. The RMS heading deviation after the engine failure was significantly different between the motion configurations (Fig. 13a). Post-hoc analysis with Bonferroni adjustment revealed a significant difference between HEX-L and LMOT ($p = 0.019$), and a marginal difference between HEX-H and LMOT ($p = 0.073$). No statistical differences were found between other motion condition pairs.

The RMS speed deviation after the engine failure is depicted in Fig. 13b. The ANOVA did not detect a statistically significant difference between motion conditions. Note that pilots were able to meet the heading and airspeed performance criteria with ease. Fig. 13c depicts the reaction time of the initial pedal input after the engine failure. This measure is not significantly different between motion conditions. However, although not significant, a slight decrease in reaction time might be observed going from the hexapod motion conditions to the large motion condition.

6 Discussion

6.1 Effects of Motion on Task Performance

Our previous quasi-transfer-of-training study evaluated whether different levels of training motion fidelity affect

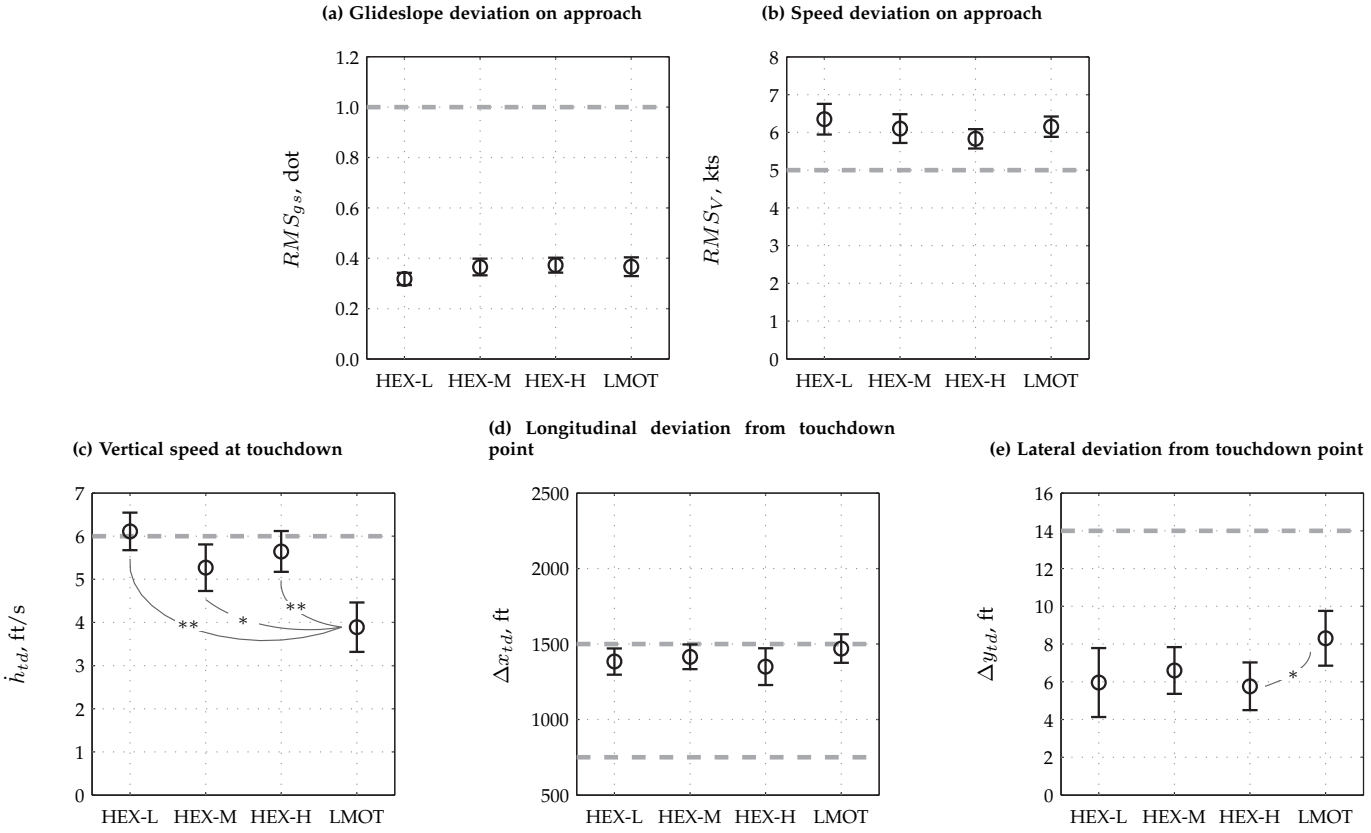

Fig. 11. Performance data of the approach and landing with sidestep task.

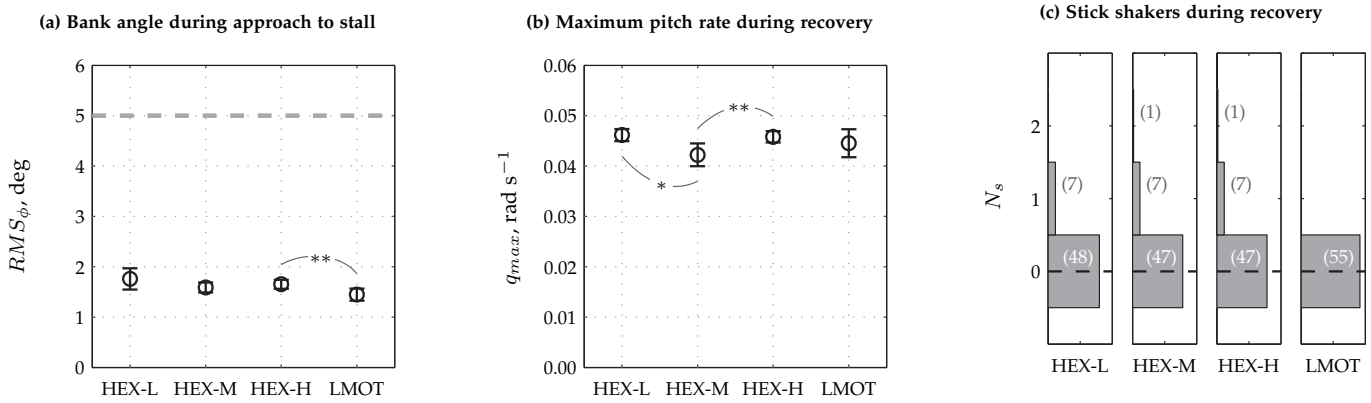

Fig. 12. Performance data of the high altitude stall recovery task.

the initial training of commercial transport pilots [16]. The study found only a limited number of significant effects of training motion, with some in the direction not predicted. We hypothesized that this was mainly caused by the fact that the group of general aviation pilots who participated in the experiment had no prior experience with the aircraft type used. This made it difficult for pilots to meet all task performance criteria. Furthermore, the experiment had a between-subjects design with pilots training for each task with only a single motion configuration, reducing the statistical power to detect differences between configurations.

The experiment here used three of the four tasks of the previous experiment with minor modifications. However, this time, we used experienced commercial airline transport pilots, representing pilots receiving recurrent training, in a within-subjects experiment in which pilots performed each task with all motion configurations. Contrary to the first experiment, these pilots met virtually all performance criteria for each run. More significant differences between motion configurations were found here compared to the previous experiment, and the effects found were more in line with what was expected based on previous research investigating the effects simulator motion on pilot control behaviour and performance.

Pilots used two different motion fidelity ratings for subjective evaluation of the simulator motion after each run.

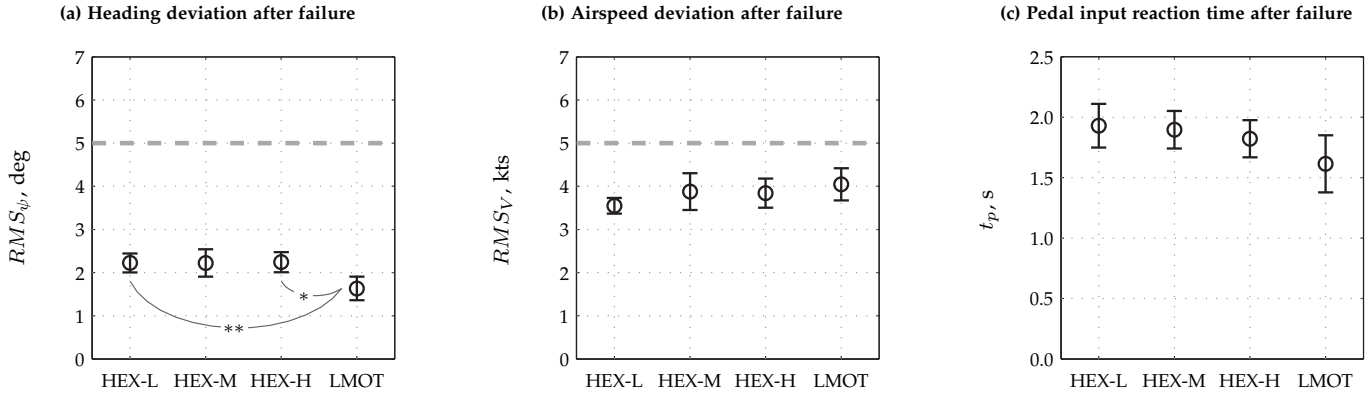

Fig. 13. Performance data of the engine out on takeoff task.

One compared the simulator motion to real flight (MFS ratings) and the other to visual motion in the out-the-window view (AFS ratings). A large spread in motion ratings was found for each task, a common issue with subjective ratings. Despite the different nature of the two ratings, they produced similar results overall. For the sidestep task, the MFS ratings showed the largest variation in fidelity ratings, and pilots gave more lower ratings across all conditions compared to the other tasks. This result was also found for the AFS ratings. Pilots tended to rate motion fidelity lower in the high-gain motion conditions (HEX-H and LMOT). This effect was significant in the MFS ratings and most likely caused by the more abrupt changes in simulator accelerations in these conditions due to atmospheric turbulence. A similar significant effect in MFS ratings was found in the stall task, which was the only other task where atmospheric turbulence was present. A plausible reason for the overall lower ratings in the sidestep task is that false tilt motion cues due to rolling were increasingly apparent closer to the ground as increasingly larger roll motions were induced at lower altitudes due to the increase in atmospheric turbulence intensity.

In the sidestep landing task, the main objective measure affected by the motion condition was the sink rate at touchdown. Pilots touched down with a significantly lower sink rate in the large motion condition (LMOT) compared to the hexapod motion conditions (HEX-L, HEX-M, HEX-H). This effect has been observed in previous research [9]. Pilots generally touched down long in all motion configurations due to the strong ground effect of the simulated aircraft. The lateral deviation from the desired touchdown point was marginally greater in the LMOT condition compared to HEX-H. This was most likely also an effect of the more pronounced false tilt motion cues due to rolling deliberately designed into the LMOT condition. The false tilt cues introduced a motion sensation similar to a lateral wind gust, causing pilots to use large rudder inputs just before touchdown.

Roll deviation RMS during the approach to the stall was significantly affected by the motion condition in the stall recovery task. Pilots deviated significantly less from the desired roll angle of 20 deg in LMOT compared to HEX-H. This is an effect generally observed in compensatory control tasks when the fidelity of motion feedback is increased [13]. The maximum pitch rate during the recovery from the stall was significantly lower in HEX-M compared to HEX-L and HEX-H. This result does not seem to be consistent with changes in the motion configurations. Finally, motion introduced a significant difference in the number of additional stick shakers during the stall recovery. No specific differences were found between motion configurations; however, in the large motion condition, no additional shakers were triggered by any of the pilots, compared to seven to eight additional shakers in the hexapod motion conditions. This is in line with our previous experiment [16]. Pilots may have been able to use the higher-fidelity pitch motion in LMOT to provide lead compensation in pitch, thereby resulting in less angle-of-attack oscillations during the recovery and, in turn, less stick shaker activations.

For the takeoff task, only the heading deviation RMS was significantly affected by the motion condition. The deviation from the desired heading after the engine failure was significantly lower in LMOT compared to HEX-L and HEX-H. In the large motion condition, pilots were able to stabilize the heading after the engine failure more effectively using the higher-fidelity yaw and lateral motion cues.

In summary, we found significant effects of motion on task performance for all of the flight tasks. Contrary to the first experiment, higher motion fidelity as predicted by the Sinacori diagram or OMCT criteria always helped pilots in performing the task. Overall, the biggest differences in performance were between the large motion condition and the high-gain/high-break-frequency hexapod motion condition. Performance in the three hexapod motion configurations was similar. Note that, for the motion tuning strategy we adopted in this experiment, this result was predicted by the Sinacori criteria; that is, all hexapod motion conditions had the same predicted fidelity and the large motion condition had a higher predicted fidelity (Fig. 4).

6.2 Objective Motion Cueing Criteria

Based on results to date, interim criteria are proposed here. More data are needed to draw precise criteria, so the interim criteria should be considered a work in progress. The approach to developing them was twofold. First, statistically significant results from this experiment, if found and applicable to a particular degree of freedom, defined

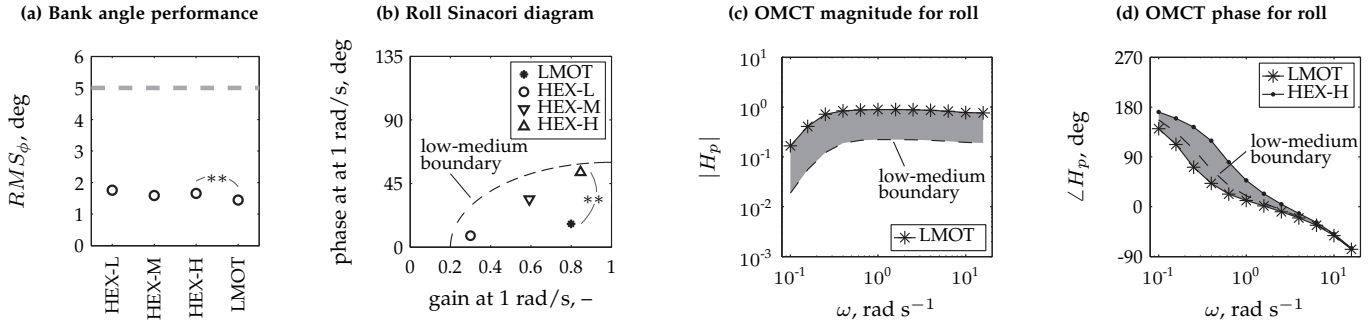

Fig. 14. Development of the OMCT roll fidelity boundaries.

the boundaries. Second, if such results were not found or applicable, then information from the literature defined the boundaries. Specifically, the boundary between low and medium fidelity from the Sinacori criteria was used [4], [5], as those criteria have assimilated and been used with many other motion studies [18].

As an example for applying this experiment's results, consider the vertical speed at touchdown. We hypothesize that vertical motion cues are most important for that measure. Then, using significant differences in task performance between different motion configurations, fidelity boundaries can be defined using the OMCT responses of those motion configurations. If a significant difference in vertical speed exists between the hexapod motion configurations and the large motion configuration (Fig. 11c), an OMCT fidelity boundary likely exists between the hexapod and large motion OMCT responses.

The HEX-L and LMOT motion configurations mainly differed in motion filter gain; that is, the motion filter break frequency between the two conditions was similar (Fig. 4). Hence HEX-L was used to suggest fidelity boundaries for magnitude only. On the other hand, HEX-H and LMOT mainly differed in motion filter break frequency. Therefore, HEX-H was used to suggest boundaries for phase only. HEX-M was used in both magnitude and phase plots when significant differences arose in the objective measures.

For clarity, Fig. 14 shows how the roll fidelity boundaries were developed using the following steps:

1) Roll deviation RMS in the stall task was used to develop roll OMCT fidelity boundaries. As indicated in Fig. 14a, roll deviation RMS was significantly different between HEX-H and LMOT only.
2) The roll Sinacori diagram is depicted in Fig. 14b, showing the gain/break-frequency tradeoff between the hexapod motion configurations. Again, the only significant difference was found between HEX-H and LMOT, which mainly differed in motion filter break frequency.
3) The Sinacori low-to-medium boundary and the LMOT configuration set the roll gain criteria, while the difference between the HEX-H and LMOT configurations set the roll phase criteria (Fig. 14c and Fig. 14d, respectively).

Fig. 15 presents fidelity boundaries based on the approach discussed above for primary responses in the six degrees of freedom. The Sinacori low-to-medium boundary is shown on all plots, but was used as a criteria boundary only when no differences were found among our motion configurations. The figure also depicts the OMCT fidelity regions based on the statistical sampling of eight simulators for reference. For the pitch OMCT response, significant differences in the number of additional stick shaker activations in the stall recovery (Fig. 12c) have been used. Significant differences in roll deviation RMS (Fig. 12a) defined the fidelity boundaries for the roll response. Significant differences in heading deviation RMS in the takeoff task (Fig. 13a) were used for both the sway and yaw OMCT responses. For the heave OMCT response, significant differences in sink rate at touchdown (Fig. 11c) were used to define the task-performance fidelity boundary. No experimental differences were found for defining the surge OMCT fidelity boundary. The responses defining the cross coupling between translational and rotational degrees of freedom have been omitted for brevity.

Note again that the fidelity regions in Fig. 15 (the dark-grey areas) are believed to contain the fidelity boundary, but are not necessarily a hard boundary yet. Results from this experiment allow us to draw relatively wide fidelity boundaries only. Future experiments will refine these fidelity boundaries.

7 Conclusions

This study used the Vertical Motion Simulator (VMS) with three flight tasks and four different motion configurations per task for an initial investigation into objective motion cueing criteria based on task performance. Three motion configurations were tuned to fit a hexapod motion platform, and the fourth was tuned to the large motion space of the VMS. Twelve experienced airline transport pilots participated, performing each task with all four motion configurations.

Both subjective and objective experimental measures were affected by the motion condition. For the subjective measures, pilots rated motion fidelity lower in the high-gain motion configurations in the tasks with atmospheric turbulence. The task performance results indicate that pilots touched down on the runway softer in the large motion condition compared to the hexapod motion conditions, and had a higher lateral deviation from the centreline in the large motion configuration compared to the high-gain hexapod configuration. Pilots deviated less from the desired roll

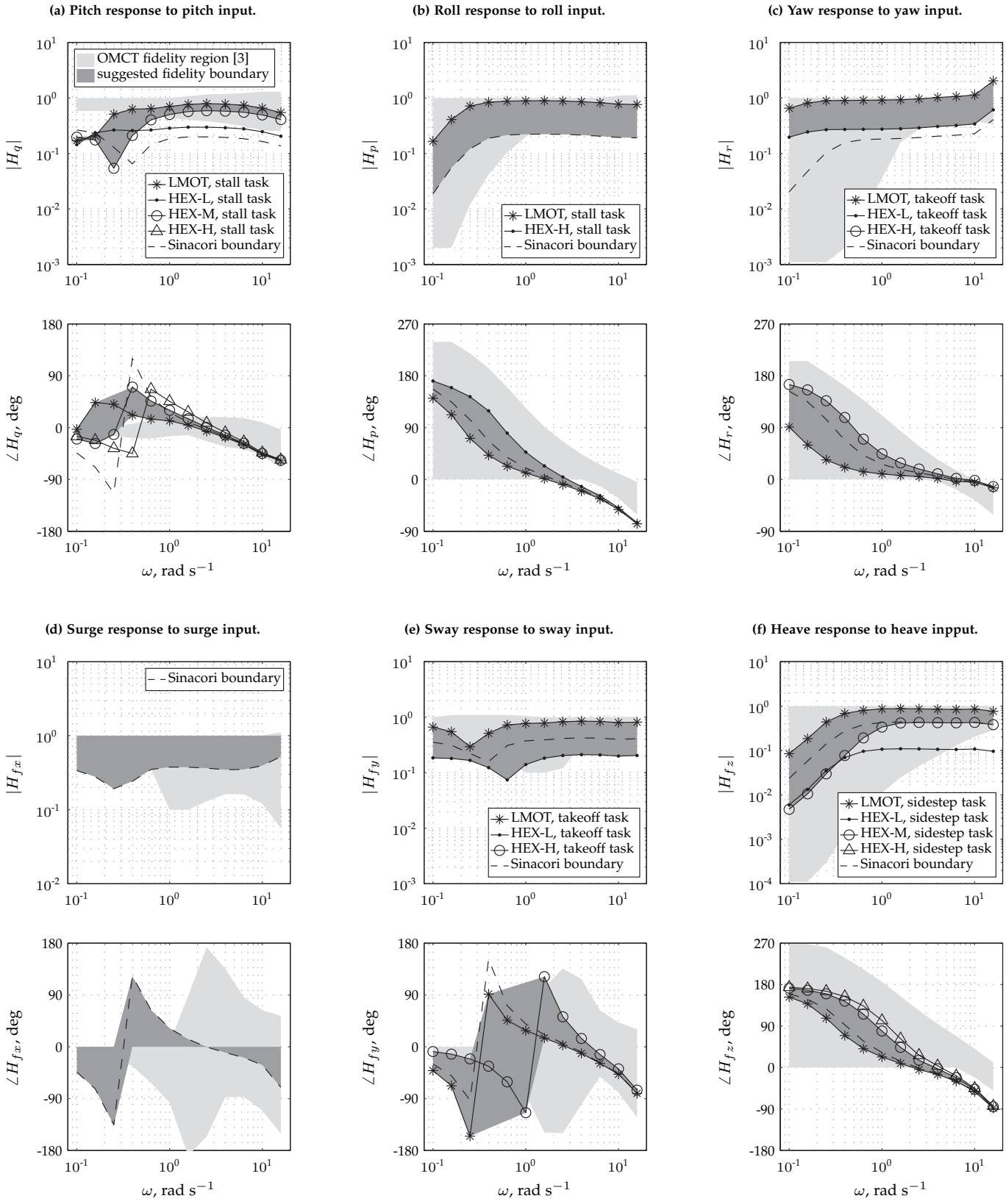

Fig. 15. Objective motion cueing test results for the takeoff task.

angle in the stall task in the large motion condition compared the high-gain hexapod condition. In the stall recovery, pilots had a lower maximum pitch rate in the medium-gain hexapod motion condition compared to the low- and high-gain hexapod conditions. Pilots had fewer stall warnings when recovering from the stall with the large motion configuration compared to the hexapod configurations. Finally, pilots deviated less from the desired heading after an engine failure in the large motion condition compared to the low and high gain hexapod conditions.

The significant effects in the objective measures were used to suggest initial objective motion cueing fidelity boundaries. Future experiments will refine these fidelity boundaries. In conclusion, the study found: 1) significant differences in task performance that were introduced by the motion condition, with the main significant differences found between the hexapod en large motion conditions, and 2) initial objective motion cueing criteria boundaries based on these significant differences in pilot performance.

ACKNOWLEDGMENTS

The authors thank the Vertical Motion Simulator simulation engineers who contributed to the experiment. We especially thank Emily Lewis and Nick Riccobono for their valuable contributions in setting up and running the experiment. We also thank Gordon Hardy for being our test pilot and instructor pilot. Finally, we would like to thank Dr. Andrew Cheng from the FAA Technical Center for the continued program support.

REFERENCES

[1] *Flight Simulation Training Device Initial and Continuing Qualification and Use*, Code of Federal Regulations, title 14, Part 60.

[2] *ICAO 9625: Manual of Criteria for the Qualification of Flight Simulation Training Devices. Volume 1 – Aeroplanes*, International Civil Aviation Organization, 2009, 3rd edition.

[3] R. J. A. W. Hosman and S. K. Advani, "Are Criteria for Motion Cueing and Time Delays Possible? Part 2." in *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Boston (MA)*, no. AIAA-2013-4833, 19–22 Aug. 2013.

[4] J. B. Sinacori, "The Determination of Some Requirements for a Helicopter Research Simulation Facility," Systems Technology, Inc., Tech. Rep. NASA CR-152066, Sep. 1977.

[5] J. A. Schroeder, "Helicopter Flight Simulation Motion Platform Requirements," NASA, Tech. Rep. NASA/TP-1999-208766, Jul. 1999.

[6] R. L. Stapleford, R. A. Peters, and F. R. Alex, "Experiments and a Model for Pilot Dynamics with Visual and Motion Inputs," NASA, Tech. Rep. NASA CR-1325, 1969.

[7] H. P. Bergeron, J. J. Adams, and G. J. Hurt, Jr., "The Effects of Motion Cues and Motion Scaling on One- and Two-Axis Compensatory Control Tasks," NASA Langley Research Center, NASA Technical Note NASA TN D-6110, Jan. 1971.

[8] R. S. Bray, "Initial Operating Experience with an Aircraft Simulator Having Extensive Lateral Motion," Ames Research Center, Moffett Field (CA), Technical Memorandum NASA TM X-62,155, 1972.

[9] ——, "Vertical Motion Requirements for Landing Simulation," Ames Research Center, Moffett Field (CA), Technical Memorandum NASA TM X-62,236, Feb. 1973.

[10] ——, "Visual and Motion Cueing in Helicopter Simulation," Ames Research Center, Moffett Field (CA), Technical Memorandum NASA TM-86818, Sep. 1985.

[11] D. E. Cooper and J. J. Howlett, "Ground Based Helicopter Simulation," in *Proceedings of the American Helicopter Society Symposium on Status of Testing and Model Techniques for V/STOL Aircraft, Essington, PA*, 1973.

[12] M. F. C. van Gool, "Influence of Motion Washout Filters on Pilot Tracking Performance," in *Piloted Aircraft Environment Simulation Techniques*, no. AGARD-CP-249. AGARD, 1978, pp. 19–1 – 19–5.

[13] H. R. Jex and R. E. Magdaleno, "Roll Tracking Effects of G-vector Tilt and Various Types of Motion Washout," in *Fourteenth Annual Conference on Manual Control*. Los Angeles (CA): University of Southern California, 25–27 Apr. 1978, pp. 463–502.

[14] H. R. Jex, W. F. Jewell, and R. E. Magdaleno, "Effects of Various Lateral-Beam-Motion Washouts on Pilot Tracking and Opinion in the "Lamar" Simulator," in *Fifteenth Annual Conference on Manual Control*. Dayton (OH): Wright State University, 20–22 Mar. 1979, pp. 244–266.

[15] D. K. Shirachi and R. S. Shirley, "Visual/Motion Cue Mismatch in a Coordinated Roll Maneuver," Computer Sciences Corporation, Contractor Report NASA CR-166259, Nov. 1981.

[16] P. M. T. Zaal, J. A. Schroeder, and W. W. Chung, "Transfer of Training on the Vertical Motion Simulator," *Journal of Aircraft*, 2015, accessed May 7, 2015. [Online]. Available: http://arc.aiaa.org/doi/abs/10.2514/1.C033115

[17] S. D. Beard, S. E. Reardon, E. L. Tobias, and B. L. Aponso, "Simulation System Optimization for Rotorcraft Research on the Vertical Motion Simulator," in *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Minneapolis (MN)*, no. AIAA-2012-4634, 13–16 Aug. 2012.

[18] J. A. Schroeder and P. R. Grant, "Pilot Behavioral Observations in Motion Flight Simulation," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit, Toronto (ON), Canada*, no. AIAA-2010-8353, 2–5 Aug. 2010.